# Mutations in STAT3 and diagnostic guidelines for hyper-IgE syndrome

**Cristina Woellner, MSc,[a] E. Michael Gertz, PhD,[b] Alejandro A. Schäffer, PhD,[b] Macarena Lagos, MD,[d] Mario Perro, MSc,[a] Erik-Oliver Glocker, MD,[a] Maria C. Pietrogrande, MD,[e] Fausto Cossu, MD,[f] José L. Franco, MD, PhD,[g] Nuria Matamoros, MD,[h] Barbara Pietrucha, MD, PhD,[i] Edyta Heropolitańska-Pliszka, MD,[i] Mehdi Yeganeh, MD,[j] Mostafa Moin, MD,[j] Teresa Español, MD, PhD,[k] Stephan Ehl, MD,[l] Andrew R. Gennery, MD,[m] Mario Abinun, MD, PhD,[m] Anna Bręborowicz, MD,[n] Tim Niehues, MD,[o] Sara Sebnem Kilic, MD,[p] Anne Junker, MD,[q] Stuart E. Turvey, MD, PhD,[q] Alessandro Plebani, MD,[r] Berta Sánchez, PhD,[s] Ben-Zion Garty, MD,[t] Claudio Pignata, MD,[u] Caterina Cancrini, MD,[v] Jiri Litzman, MD,[w] Özden Sanal, MD,[x] Ulrich Baumann, MD,[y] Rosa Bacchetta, MD,[z] Amy P. Hsu, BA,[c] Joie N. Davis, CRNP,[c] Lennart Hammarström, MD,[aa] E. Graham Davies, MD,[bb] Efrem Eren, MD,[cc] Peter D. Arkwright, MD, PhD,[dd] Jukka S. Moilanen, MD, PhD,[ee] Dorothee Viemann, MD,[ff] Sujoy Khan, MRCP,[gg] László Maródi, MD,[hh] Andrew J. Cant, MD,[m] Alexandra F. Freeman, MD,[c] Jennifer M. Puck, MD,[ii] Steven M. Holland, MD,[c] and Bodo Grimbacher, MD[a]** *London, Newcastle upon Tyne, Southampton, Manchester, and Scunthorpe, United Kingdom, Bethesda, Md, Valparaíso, Chile, Milan, Cagliari, Brescia, Naples, and Rome, Italy, Medellín, Colombia, Palma de Mallorca, Barcelona, and Seville, Spain, Warsaw and Poznan, Poland, Tehran, Iran, Freiburg, Düsseldorf, Hannover, and Münster, Germany, Bursa and Ankara, Turkey, Vancouver, British Columbia, Canada, Petah Tiqva, Israel, Brno, Czech Republic, Stockholm, Sweden, Tampere, Finland, Debrecen, Hungary, and San Francisco, Calif*

From [a]the Department of Immunology and Molecular Pathology, Royal Free Hospital, University College London; [b]the National Center for Biotechnology Information and [c]the Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, Bethesda; [d]Cátedra de Inmunología Escuela de Medicina, Universidad de Valparaíso; [e]the Department of Pediatrics, University of Milan, Fondazione Policlinico IRCCS; [f]the Bone Marrow Transplant Unit, Ospedale Microcitemico, Cagliari; [g]the Group of Primary Immunodeficiencies, University of Antioquia, Medellín; [h]the Immunology Service, Son Dureta Hospital, Palma de Mallorca; [i]the Gastroenterology, Hepatology and Immunology Clinic, Children's Memorial Health Institute, Warsaw; [j]Immunology, Asthma and Allergy Research Institute, Children Medical Centre, Tehran University of Medical Sciences; [k]the Immunology Unit, Hospital Vall d'Hebron, School of Medicine, Barcelona; [l]the Department of Pediatrics and Adolescent Medicine, University Hospital Freiburg; [m]the Children's Bone Marrow Transplant Unit, University of Newcastle upon Tyne; [n]the Department of Pediatric Pulmonology, Allergy and Clinical Immunology, 3rd Department of Pediatrics, Poznan University of Medical Sciences; [o]the Immunodeficiency and Pediatric Rheumatology Centre, HELIOS Klinikum Krefeld, Heinrich Heine University of Düsseldorf; [p]the Department of Pediatric Immunology, Faculty of Medicine, Uludag University, Bursa; [q]the Department of Pediatrics, British Columbia's Children's Hospital and University of British Columbia; [r]the Department of Pediatrics and Institute of Molecular Medicine A. Novicelli, University of Brescia; [s]the Immunology Service, University Hospital, Virgen del Rocıo, Sevilla; [t]the Department of Pediatrics, Schneider Children's Medical Center, Petah Tiqva; [u]the Department of Pediatrics, "Federico II" University of Naples; [v]the Division of Immunology and Infectious Disease, Bambino Gesù Children's Hospital, University of Rome Tor Vergata; [w]the Department of Clinical Immunology and Allergology, Faculty of Medicine, Masaryk University, St Anne's University Hospital, Brno; [x]the Immunology Division, Hacettepe University Children's Hospital, Ankara; [y]the Department of Pediatric Pulmonology and Neonatology, Medical School Hannover; [z]the San Raffaele Telethon Institute for Gene Therapy (HSR-TIGET), Milan; [aa]the Division of Clinical Immunology, Department of Laboratory Medicine, Karolinska Institute at the Karolinska University Hospital, Stockholm; [bb]the Department of Immunology, Great Ormond Street Hospital, London; [cc]the Immunology Department, Southampton General Hospital; [dd]Paediatric Immunology, University of Manchester; [ee]the Department of Clinical Genetics, University of Oulu and Medical School, University of Tampere; [ff]the Institute of Immunology and Department of Pediatrics, University of Münster; [gg]Path Links Immunology, Scunthorpe General Hospital; [hh]the Department of Infectious and Pediatric Immunology, Medical and Health Science Center, University of Debrecen; and [ii]the Department of Pediatrics, University of California, San Francisco.

Supported by in part by the European grant MEXT-CT-2006-042316 to B.G., a grant of the Primary Immunodeficiency Association provided by GSK, the European consortium grant EURO-PADnet HEALTH-F2-2008-201549, the grant OTKA49017 to L.M., the foundation C. Golgi from Brescia to A.P., the Intramural Research Program of the National Institutes of Health, NLM (E.M.G., A.A.S.), the National Institute of Allergy and Infectious Diseases (A.P.H., A.F.F., J.N.D., S.M.H.), and federal funds from the National Cancer Institute, National Institutes of Health, under contract N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Disclosure of potential conflict of interest: B. Grimbacher has received research support from the European Union and the Primary Immunodeficiency Association, and is secretary of the European Society for Immunodeficiencies. J. L. Franco is vice president of Fundación Diana García de Olarte. T. Niehues has received consulting fees from Talecris Biotherapeutics, Octapharma, Behring ZLB, and Baxter. S. S. Kilic has received research support from Uludag University. A. Junker has received research support from the BS Children's Hospital Foundation, the Michael Smith Foundation, and the Canadian Institutes of Health Research. L. Hammarström has received research support from the National Institutes of Health, the European Union, and the Swedish Research Council. J. S. Moilanen has received research support from the Academy of Finland. A. J. Cant has received research support from the Bubble Foundation, UK, and the Medical Research Council, and has served as an expert witness on cases of bacterial meningitis and herpes simplex encephalitis. J. M. Puck has received research support from the National Institutes of Health, NICHD, and the Jeffrey Modell Foundation, and is on the Steering Committee of the Immune Deficiency Foundation and the Expert Committee on Primary Immunodeficiencies for IUIS. The rest of the authors have declared that they have no conflict of interest.

Received for publication November 13, 2008; revised October 2, 2009; accepted for publication October 8, 2009.

Reprint requests: Bodo Grimbacher, MD, Dept of Immunology and Molecular Pathology, Royal Free Hospital and University College London, Pond Street, London NW3 2QG, United Kingdom. E-mail: b.grimbacher@ucl.ac.uk.

0091-6749/$36.00



doi:10.1016/j.jaci.2009.10.059

**Background: The hyper-IgE syndrome (HIES) is a primary immunodeficiency characterized by infections of the lung and skin, elevated serum IgE, and involvement of the soft and bony tissues. Recently, HIES has been associated with heterozygous dominant-negative mutations in the signal transducer and activator of transcription 3 *(STAT3)* and severe reductions of $T_H17$ cells.**
**Objective: To determine whether there is a correlation between the genotype and the phenotype of patients with HIES and to establish diagnostic criteria to distinguish between *STAT3* mutated and *STAT3* wild-type patients.**
**Methods: We collected clinical data, determined $T_H17$ cell numbers, and sequenced *STAT3* in 100 patients with a strong clinical suspicion of HIES and serum IgE >1000 IU/mL. We explored diagnostic criteria by using a machine-learning approach to identify which features best predict a *STAT3* mutation.**
**Results: In 64 patients, we identified 31 different *STAT3* mutations, 18 of which were novel. These included mutations at splice sites and outside the previously implicated DNA-binding and Src homology 2 domains. A combination of 5 clinical features predicted *STAT3* mutations with 85% accuracy. $T_H17$ cells were profoundly reduced in patients harboring *STAT3* mutations, whereas 10 of 13 patients without mutations had low (<1%) $T_H17$ cells but were distinct by markedly reduced IFN-γ–producing $CD4^+$T cells.**
**Conclusion: We propose the following diagnostic guidelines for STAT3-deficient HIES. Possible: IgE >1000IU/mL plus a weighted score of clinical features >30 based on recurrent pneumonia, newborn rash, pathologic bone fractures, characteristic face, and high palate. Probable: These characteristics plus lack of $T_H17$ cells or a family history for definitive HIES. Definitive: These characteristics plus a dominant-negative heterozygous mutation in *STAT3*. (J Allergy Clin Immunol 2010;125:424-32.)**

***Key words:*** *Hyper-IgE syndrome, HIES, Job syndrome, $T_H17$ cells, STAT3 mutations, diagnostic guidelines*

*Abbreviations used*
HIES: Hyper-IgE syndrome
NIH: National Institutes of Health
SEB: *Staphylococcus* enterotoxin B
SH2: Src homology 2
STAT3: Signal transducer and activator of transcription 3
SVM: Support vector machine
UPN: Unique patient number

The hyper-IgE syndrome (HIES) is a multisystem disorder characterized by eczema, skin abscesses, recurrent staphylococcal infections of the skin and lungs, pneumatocele formation, candidiasis, eosinophilia, and elevated serum levels of IgE. Nonimmunologic features of HIES include characteristic facial appearance, scoliosis, retained primary teeth, joint hyperextensibility, bone fractures after minimal trauma, and craniosynostosis.[1-3] In 1999, the National Institutes of Health (NIH) clinical HIES scoring system based on 19 clinical and laboratory findings was introduced.[4] A point scale was developed: more specific and objective findings were assigned more points. Scores of at least 40 points suggested HIES, whereas a score below 20 made the diagnosis unlikely. For intermediate values, no firm conclusion could be reached.

In 2007, mutations in signal transducer and activator of transcription 3 *(STAT3)* were shown to be a cause of HIES.[5,6] STAT3 plays a key role in the signal transduction of a broad range of cytokines.[7,8] After cytokine binding and Janus kinase activation, STAT3 is phosphorylated, dimerizes, and translocates to the nucleus, where it controls transcription of its target genes.[9] STAT3 is crucial for the IL-6–mediated regulation of $T_H17$ cells that are significant producers of IL-17, a proinflammatory cytokine involved in the host defense of *Staphylococcus aureus* and *Candida*.[10-12]

There are 7 publications on *STAT3* mutations reporting on 155 patients with HIES. In 141 of these patients, heterozygous mutations in *STAT3* were identified.[5,6,13-17] Therefore, we addressed the following question: how common is a diagnosis of HIES without a *STAT3* mutation? We also asked the following question: do some features of the HIES phenotype make it more likely that a patient with HIES has a *STAT3* mutation, and can any features of the HIES phenotype predict the location of mutations within *STAT3*—that is, is there any phenotype-genotype correlation?

Because *STAT3* has 24 exons and 3 splice variants, predicting which patients are likely to have a *STAT3* mutation could save sequencing resources.

In a multicenter cohort of 100 patients with suspected HIES, we evaluated 17 of the clinical and laboratory features used in the original scoring method,[4] the reported laboratory feature of a very low $T_H17$ $CD4^+$ T-cell count, and the genetic diagnosis to develop a new scoring system aimed to discern those patients with HIES with *STAT3* mutations from those without mutations.[14-16]

On the basis of our analysis of 100 unrelated patients evaluated worldwide, we propose guidelines for a clinical assessment before a confirmation of the suspected diagnosis by laboratory and molecular analysis.

## METHODS

### Patients and controls

Over the period of the last 8 years, we have collected a cohort of 228 patients with the suspected diagnosis of HIES in a worldwide collaboration; 55 of these patients have been published elsewhere.[6,13,17,18] Of the remaining patients, 100 unrelated patients fulfilled inclusion criteria for this study: signed consent form, complete NIH scoring sheet, a strong clinical suspicion of HIES according to the referring immunologist, an IgE >1000 IU/mL, and an available sample of genomic DNA or RNA. To promote uniformity of documentation across the 33 different participating centers, a scoring sheet listing the original NIH clinical symptoms was used.[4]

Of the 100 patients with the clinical suspicion of HIES, 61 were male and 39 female; the age of the patients at the time of clinical evaluation ranged between 1 and 58 years. Seventy-two patients came from Europe, 20 from the Middle East, 7 from South America, and 1 from North America.

Eighty patients had HIES scores ≥40, suggesting that these patients probably had HIES, whereas 20 patients had scores below 40, suggesting a diagnostic uncertainty or a variant of HIES.[19] Only 2 of these patients (unique patient numbers [UPNs] 133 and 134) have been described in published *STAT3* mutation reports.[17] Detailed information on patients, including clinical scores and detected *STAT3* mutations, are summarized in this article's Table E1 in the Online Repository at www.jacionline.org and in Moin et al.[20]

We applied the diagnostic guidelines, developed by using the clinical scores of the cohort of 100 patients to a replication set of 50 unrelated patients all scored by a consistent team of clinicians at the NIH. Of these 50, 33 had a

mutation in *STAT3* and 17 did not; the 33 patients with a *STAT3* mutation were from a previously published cohort.[6] In addition, 28 patients with severe atopic dermatitis and an IgE level >1000 IU/mL were scored.

Control DNA was isolated from 100 healthy Caucasian subjects according to approved protocols. *STAT3* was sequenced in all controls to verify that the sequence changes seen in patients were not frequent polymorphisms. In addition, 12 controls were studied for their lymphocyte phenotype. All patients and controls or their parental or legal guardians provided written consent for the conducted studies following local ethics committee requirements.

## PCR and sequencing

Genomic DNA and RNA of controls and patients was isolated from either whole blood or PBMCs by using the RNeasy Kit (Qiagen, United Kingdom) according to the manufacturer's instructions. RNA was reverse-transcribed by using Omniscript reverse transcriptase (Qiagen).

Coding genomic sequences and cDNA of *STAT3* were amplified and purified by using the QIAquick PCR purification kit (Qiagen). Primer sequences are available on request. Purified PCR products were sequenced with the ABI PRISM BigDye Terminator cycle ready reaction kit V3.1 (Applied Biosystems, Foster City, Calif) by using the PCR primers as sequencing primers. The sequencing was performed on a 3130xl Applied Biosystems Genetic Analyzer, and the data were analyzed with DNA Sequencing Analysis software, version 5.2 (Applied Biosystems) and Sequencher version 4.8 (Gene Codes Corp, Ann Arbor, Mich).

## Cell culture and TNF-α ELISA

Monocyte-derived macrophages were generated as previously described,[21] kept in Opti-Mem I serum free medium (Invitrogen, United Kingdom), and differentiated by adding 50 ng/mL M-CSF (PeproTech, United Kingdom) per day. After 5 days of culture, the monocytes/macrophages were preincubated with 25 ng/mL IL-10 (R&D, United Kingdom) for 1 hour and then stimulated overnight with 50 ng/mL *Escherichia coli* LPS (Sigma, United Kingdom). TNF-α release was measured by ELISA (PeproTech, United Kingdom) according to manufacturer's instructions.

## Lymphocyte stimulation and flow cytometry

Blood was collected from 30 patients and 12 healthy donors. PBMCs were isolated by using Lymphoprep (Axis-Shield, United Kingdom) and viably frozen. Freshly thawed cells resuspended at a concentration of $10^6$/mL in RPMI (Lonza, United Kingdom), supplemented with 10% FCS (Gibco, United Kingdom), were exposed to *Staphylococcus* enterotoxin B (SEB; Sigma) 1μg/mL and brefeldin A (Sigma) 2.5 μg/mL 16 hours at 37°C and 5% $CO_2$. The cells were then surface-stained for CD4 (perinidin chlorophyll protein) and CD45RO (phycoerythrin-Cy7), followed by an intracellular staining for IL-17 (Alexa Fluor 647) and IFN-γ (fluorescein isothiocyanate). Intracellular staining was performed by using the BD Cytofix/Cytoperm Fixation/Permeabilization Kit (BD Biosciences). All antibodies were from BD Biosciences, except for IL-17, which was from eBioscience. Data were acquired on an LSR II flow cytometer and analyzed by using FACSDiva software (BD Biosciences). Wilcoxon rank-sum tests were performed with Prism 5.01 software (Prism); *P* values are 2-sided.

## Statistical analysis

Univariate statistical analyses were done by using GNU R.[22] For tests for which we had a previous hypothesis (eg, higher clinical score is associated with a *STAT3* mutation), *P* values were 1-sided; otherwise, *P* values were 2-sided. A Bonferroni correction was applied to each series of univariate tests.

The 100 patients in the cohort form our data set for univariate analysis and our training set for the machine learning methods described in this subsection. The gold standard for the diagnosis of *STAT3*-deficient HIES was the discovery of a heterozygous mutation (most mutations proven to be dominant-negative) in *STAT3* by using the DNA sequencing techniques described under PCR and sequencing.

Table I shows the prevalence of each feature in our HIES cohort. On the basis of preliminary experiments, we found that support vector machines (SVMs) are an effective way to classify this data set. To compare feature sets,[23] we took the scores for 17 of the 19 features[4] and calculated the leave-1-out accuracy for SVMs generated from each subset of size at least 2 and no more than 7. Because there are 41,208 such subsets, we used the software package OOQP (http://pages.cs.wisc.edu/~swright/ooqp),[24] which contains specialized code to generate linear SVMs. We used a prerelease version of OOQP (available from E.M.G.) that provides leave-1-out estimation and that has been optimized for speed.

We chose as candidate feature sets those 344 sets for which OOQP reported leave-1-out accuracy of better than 80%. Leave-1-out testing was performed with OOQP and SVMlight (http://svmlight.joachims.org),[25] as described in the subsection Statistical analyses in the Online Repository. For several feature sets, a classifier was trained by using SVMlight on the scores from our cohort of 100 patients and then applied to predict whether each of the patients in the replication set had a *STAT3* mutation.

# RESULTS

## *STAT3* mutations

Of the 100 unrelated patients with suspected HIES, 64 carried heterozygous mutations within *STAT3*. Thirty-six patients did not show any mutation in the coding regions of *STAT3* or their flanking intronic sequences. Overall, we found 31 distinct mutations: 46 patients carried previously described mutations,[5,6,13-17] and 18 patients harbored distinct, novel mutations. Of these 18 mutations, 8 affected the DNA-binding domain, 5 the Src homology 2 (SH2) domain, 4 the transactivation domain, and 1 the coiled-coil domain (Table II).

Five mutations, seen in 6 patients, were heterozygous mutations at the intron-exon boundaries of exons 12 or 22, affecting the DNA binding and transactivation domains, respectively. Each of these intronic mutations led to an incorrect splicing of its neighboring exon (data not shown). Three patients had 2 distinct heterozygous mutations at the 3′ splice site before exon 12, and 2 patients had distinct mutations in at the 5′ splice site after exon 12. Two of these 4 mutations were described by Renner et al,[13] and the other 2 are novel. All 4 mutations cause exon 12 to be skipped, suggesting an in-frame deletion of amino acids 371 to 380.

None of 100 healthy controls carried any of the mutations listed in Table II.

To analyze the effects of the newly observed *STAT3* mutation *in vitro*, we used a TNF-α release assay as described previously by Minegishi et al.[5] We preincubated monocytes/macrophages of 10 patients and 11 healthy controls with IL-10 for 1 hour and then stimulated with LPS. Eight of the 10 patients tested had novel mutations, 1 was wild-type for *STAT3*, and 1 had the previously reported mutation R382Q. In healthy controls and a *STAT3* wild-type patient, IL-10 reduced TNF-α release by approximately 90%, whereas in the R382Q patient, TNF-α release was reduced by only 40%. All 8 novel mutations showed a reduction in IL-10–mediated downregulation of TNF-α secretion. We compared the effect of the 8 novel mutations collectively with 11 healthy controls by a rank-sum test and reached significance ($P < .0001$). However, not all mutations had the same effect, suggesting that some mutations may lead to STAT3 proteins with a hypomorphic function and possibly to a milder phenotype (Fig 1).

A substantial portion (36 patients) of our cohort did not have mutations in the coding or splice regions of *STAT3*. To exclude

**TABLE I.** Number and percentage of patients positive for the 17 evaluated features

| | all HIES | | HIES STAT3 wild-type | | HIES STAT3 mutated | |
|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % |
| Recurrent pneumonia | 85/100 | 85 | 24/36 | 66.7 | 61/64 | 95.3 |
| Eczema | 90/100 | 90 | 32/36 | 88.9 | 58/64 | 90.6 |
| Recurrent skin abscesses | 86/100 | 86 | 28/36 | 77.8 | 58/64 | 90.6 |
| Characteristic face | 82/99 | 82.8 | 24/35 | 68.6 | 58/64 | 90.6 |
| Failure to shed deciduous teeth | 60/86 | 68.9 | 16/31 | 51.6 | 44/55 | 80.0 |
| Lung cyst formation | 61/97 | 62.9 | 14/34 | 41.2 | 47/63 | 74.6 |
| Eosinophilia | 68/94 | 72.3 | 27/36 | 75.0 | 41/58 | 70.7 |
| Newborn rash | 52/86 | 60.5 | 15/29 | 51.7 | 37/57 | 64.9 |
| Other unusual infections | 47/94 | 50 | 13/34 | 38.2 | 34/60 | 56.7 |
| Increased interalar distance | 37/83 | 44.6 | 10/31 | 32.3 | 27/52 | 51.9 |
| Cathedral palate | 41/84 | 48.8 | 12/31 | 38.7 | 29/53 | 54.7 |
| Hyperextensibility | 37/87 | 42.5 | 8/32 | 25.0 | 29/55 | 52.7 |
| Pathologic bone fractures | 32/94 | 34.0 | 5/35 | 14.3 | 27/59 | 45.8 |
| Recurrent upper respiratory infections | 41/92 | 44.6 | 14/33 | 42.4 | 27/59 | 45.8 |
| Candidiasis | 37/91 | 40.6 | 12/33 | 36.4 | 25/58 | 43.1 |
| Scoliosis | 20/83 | 24.1 | 7/33 | 21.2 | 13/50 | 26.0 |
| Midline anomaly | 12/86 | 14.0 | 5/34 | 14.7 | 7/52 | 13.5 |

Features were considered as present if more than 1 point of the NIH score[4] was achieved. The 5 cardinal clinical features (recurrent pneumonia, newborn rash, pathologic bone fractures, characteristic face, and cathedral palate) are *shaded*.

**TABLE II.** Among the 100 patients studied, 64 patients had 31 distinct mutations

| No. of patients | Protein domain | Site of mutation | DNA sequence change | Predicted amino acid change |
|---|---|---|---|---|
| 1 | Coiled-coil | Exon 3 | c.172C>T | H58Y |
| 1 | DNA-binding | Exon 10 | c.982_990dupTGCATGCCC | C328_P330dup |
| 1 | DNA-binding | Exon 10 | c.1025G>A | G342D |
| 2 | DNA-binding | Intron 11 | c.1110–2A>G | D371_G380del |
| 1 | DNA-binding | Intron 11 | c.1110–1G>T | D371_G380del |
| 1 | DNA-binding | Intron 12 | c.11391 1G>T | D371_G380del |
| 1 | DNA-binding | Intron 12 | c.11391 2insT | D371_G380del |
| 14 | DNA-binding | Exon 13 | c.1144C>T | R382W |
| 2 | DNA-binding | Exon 13 | c.1145G>T | R382L |
| 9 | DNA-binding | Exon 13 | c.1145G>A | R382Q |
| 2 | DNA-binding | Exon 13 | c.1150T>C | F384L |
| 1 | DNA-binding | Exon 13 | c.1166C>T | T389I |
| 1 | DNA-binding | Exon 14 | c.1268G>A | R423Q |
| 1 | DNA-binding | Exon 16 | c.1387_1389delGTG | V463del |
| 1 | DNA-binding | Exon 16 | c. 1396 A>G | N466D |
| 1 | DNA-binding | Exon 16 | c.1397A>G | N466S |
| 1 | DNA-binding | Exon 16 | c.1397A>C | N466T |
| 1 | DNA-binding | Exon 16 | c.1398C>G | N466K |
| 1 | DNA-binding | Exon 16 | c.1407G>T | Q469H |
| 1 | SH2 | Exon 20 | c.1771A>G | K591E |
| 1 | SH2 | Exon 20 | c.1865C>T | T622I |
| 1 | SH2 | Exon 21 | c.1907C>A | S636Y |
| 10 | SH2 | Exon 21 | c.1909G>A | V637M |
| 1 | SH2 | Exon 21 | c.1910T>C | V637A |
| 1 | SH2 | Exon 21 | c.1915T>C | P639S |
| 1 | SH2 | Exon 21 | c.1970A>G | Y657C |
| 1 | SH2 | Exon 21 | c.2003C>T | S668F |
| 1 | Transactivation | Exon 22 | c.2124C>G | T708S |
| 1 | Transactivation | Exon 22 | c.2129T>C | F710C |
| 1 | Transactivation | Exon 22 | c.2141C>G | T714A |
| 1 | Transactivation | Intron 22 | c.21441 1G>A | p.? |

*p.?*, Unknown effect on protein level.
Thirteen of these mutations were reported previously.[5,6,13-17] The other 18 distinct mutations, which are *shaded*, have not been previously reported. The cDNA sequence positions reported are from isoform NM_139276.2. The amino acid positions are from the isoform NP_644805.1.

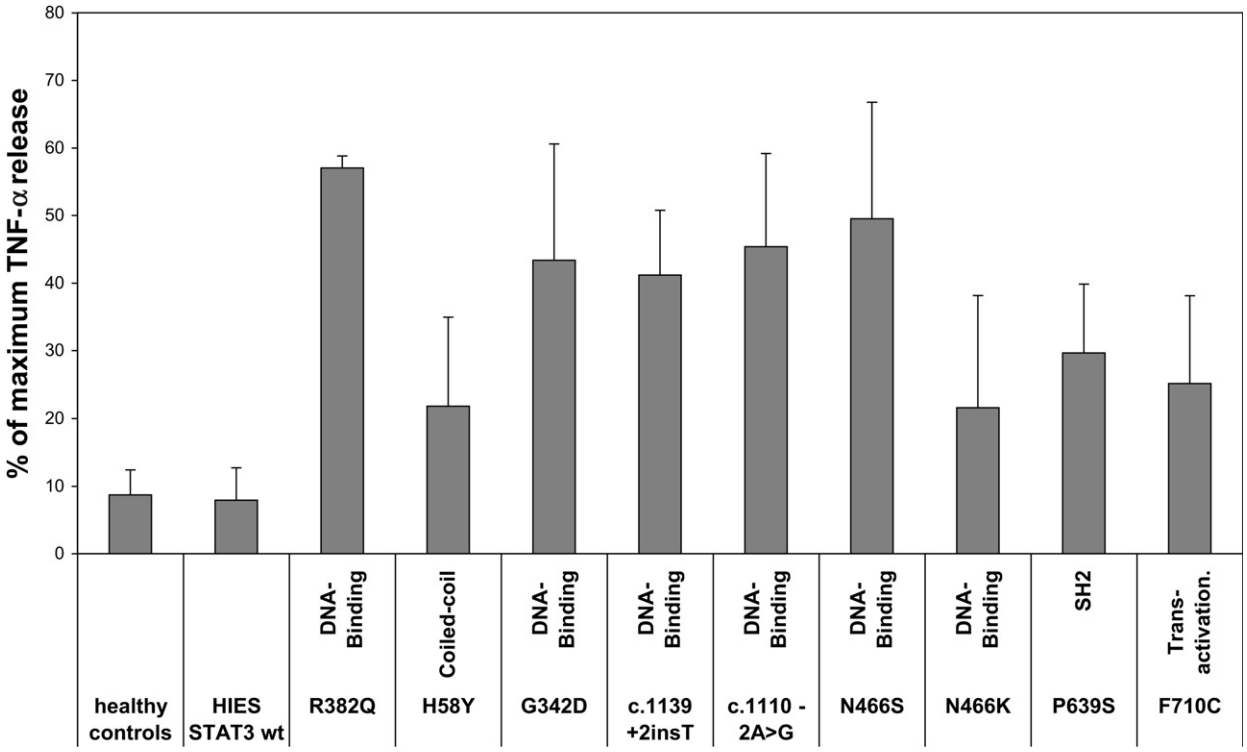


**FIG 1.** Inhibition of TNF-α release in LPS-activated macrophages by IL-10. Macrophages of nine *STAT3*-mutated patients, 8 of whom had novel mutations, one *STAT3* wild-type *(wt)* patient, and 11 healthy controls were pretreated with IL-10 and then stimulated with LPS. Supernatants were examined for the presence of TNF-α. To ensure better comparability of data from different healthy donors, the impact of IL-10 on TNF-α release is shown as percentage of maximum TNF-α release on LPS stimulation.

undetected splice site variants not described in healthy individuals, we sequenced the cDNA from 25 of these patients, none of whom revealed any mutation or exon skipping. Because of the limited availability of samples, the cDNA of 11 patients could not be tested (UPNs 21, 28, 35, 38, 58, 78, 80, 82, 88, 105, and 112). Although some patients lacking *STAT3* mutations also had NIH-HIES scores below the likely diagnostic cutoff of 40 points, 20 (56%) had scores ≥40 (Table E1).

## Evaluating the value of the NIH-HIES clinical scoring system

To evaluate the NIH-HIES scoring system,[4] we performed 3 Wilcoxon rank-sum tests to calculate the significance of the association of total clinical HIES score with mutation status. High total HIES score was strongly associated with having a mutation in *STAT3* (*P* value 3.9e-07). Among those patients with a *STAT3* mutation, a high total HIES score may carry some information about whether the mutation is located in the DNA binding domain ($P = .094$) but was not significantly associated with having a mutation in the SH2 domain rather than elsewhere in *STAT3* (see this article's Table E2 in the Online Repository at www.jacionline.org).

We evaluated whether any cutoff value for the NIH-HIES score would be useful for predicting whether a patient has a *STAT3* mutation. Using the rule "choose the largest threshold T that gives the smallest error," leave-1-out testing gave 23% error, 75% sensitivity, and 80.1% selectivity. The optimal cutoff was T = 49.5. On the replication set of 50 patients, the cutoff T = 49.5 has a sensitivity of 97.0% but a selectivity of 58.8%. Therefore we sought a smaller and more precise set of features than the NIH-HIES score. These seemingly competing objectives could be achieved with a linear support vector machine.

## Set of clinical features best predicting the genotype

In search of a minimal set of clinical features associated with *STAT3* mutation-positive HIES, we prospectively scored the 100 patients according to the NIH-HIES method[4] and performed leave-1-out error estimates for all feature sets tested.

There were 12 sets with a leave-1-out accuracy of 85% (see this article's Table E3 in the Online Repository at www.jacionline.org). By logistic regression of the individual features occurring in at least 2 feature sets, pneumonia ($P = .002$), lung cysts ($P = .001$), and characteristic face ($P = .002$) were positively associated with a *STAT3* mutation. Of the 12 sets, 1 had 5 features, and all other sets had more than 5 features. These 5 features are (1) pneumonia, (2) newborn rash, (3) pathologic fractures, (4) characteristic face of Job syndrome, and (5) cathedral palate. Notably, a patient with HIES does not need to have severe scores for all 5 features to get a total score above the threshold of 30 points for predicting a *STAT3* mutation (see this article's Table E4 in the Online Repository at www.jacionline.org). More information about the 5 features chosen for the STAT3 score is provided in a subsection of the same name in the Online Repository.

Although some clinical features may take years to develop, the misclassification rate on young children (5/37 for age 11 years and younger; 2/17 for age 7 years and younger) was comparable

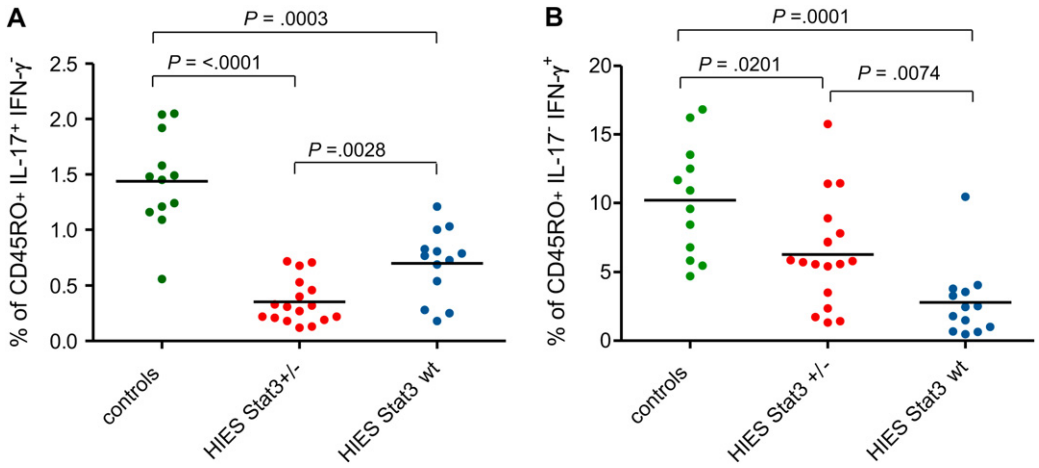


**FIG 2.** Percentage of IL-17 **(A)** and IFN-γ **(B)** expressing $CD4^+$ memory T cells, determined by intracellular cytokine expression after overnight stimulation with SEB. Each *symbol* represents the value from an individual donor or patient. Statistical significance was determined with a Wilcoxon rank-sum test. *P* values are 2-sided. Median values are shown as *horizontal bars*. *wt,* Wild-type.

**TABLE III.** Percentages of IL-17 and IFN-γ expressing $CD4^+$ memory T cells in 30 analyzed patients

| UPN | Sex | Age of scoring | Origin | DNA sequence change | Predicted amino acid change | Domain | NIH score* | STAT3 score† | % of $CD45RO^+$ $IL17^1$ IFN-$\gamma^-$ cells | % of $CD45RO^+$ $IL17^-$ IFN-$\gamma^+$ cells | Candidiasis |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 110 | F | 28 | EU | c.1025G>A | G342D | DNA-binding | 79 | 63.33 | 0.27 | 11.44 | Yes |
| 114 | M | 22 | EU | c.11102A>G | D371_G380del | DNA-binding | 70 | 50.00 | 0.32 | 5.57 | Yes |
| 86 | M | 13 | EU | c.11101G>T | D371_G380del | DNA-binding | 83 | 76.67 | 0.13 | 5.80 | Yes |
| 17 | F | 25 | EU | c.11391 1G>T | D371_G380del | DNA-binding | 65 | 41.66 | 0.33 | 1.43 | Yes |
| 108 | M | 14 | EU | c.11191 2insT | D371_G380del | DNA-binding | 54 | 31.67 | 0.22 | 3.51 | Yes |
| 15 | M | 29 | EU | c.1144C>T | R382W | DNA-binding | 76 | 63.33 | 0.21 | 15.76 | Yes |
| 19 | M | 15 | EU | c.1144C>T | R382W | DNA-binding | 63 | 48.33 | 0.71 | 5.88 | Yes |
| 126 | F | 11 | EU | c.1144C>T | R382W | DNA-binding | 58 | 40 | 0.12 | 7.17 | Yes |
| 128 | F | 3 | EU | c.1144C>T | R382W | DNA-binding | 29 | 38.33 | 0.53 | 1.73 | Yes |
| 60 | M | 29 | ME | c.1150T>C | F384L | DNA-binding | 71 | 55 | 0.68 | 1.33 | Yes |
| 118 | F | 32 | EU | c.1268G>A | R423Q | DNA-binding | 59 | 40 | 0.19 | 5.43 | No |
| 71 | M | 24 | EU | c.1907C>A | S636Y | SH2 | 71 | 40 | 0.31 | 2.36 | Yes |
| 16 | F | 39 | EU | c.1909G>A | V637M | SH2 | 68 | 55 | 0.18 | 8.9 | Yes |
| 115 | M | 19 | EU | c.1909G>A | V637M | SH2 | 56 | 68.33 | 0.72 | 5.72 | ND |
| 125 | M | 16 | EU | c.1909G>A | V637M | SH2 | 58 | 45 | 0.22 | 7.79 | No |
| 18 | M | 23 | EU | c.2129T>C | F710C | Transactivation | 58 | 40 | 0.46 | 11.41 | Yes |
| 116 | M | 8 | EU | c.2141C>G | T714A | Transactivation | 63 | 50 | 0.4 | 5.59 | Yes |
| 1 | M | 19 | EU | No mutation | | | 72 | 63.33 | 0.25 | 2.45 | Yes |
| 4 | M | 17 | EU | No mutation | | | 60 | 35 | 0.79 | 10.46 | Yes |
| 9 | M | 30 | EU | No mutation | | | 49 | 25 | 0.69 | 1.49 | Yes |
| 10 | M | 59 | EU | No mutation | | | 29 | 25 | 0.18 | 0.65 | No |
| 14 | F | 34 | EU | No mutation | | | 47 | 21.67 | 0.28 | 1.8 | No |
| 20 | F | 9 | EU | No mutation | | | 32 | 21.67 | 1 | 3.28 | No |
| 22 | F | 18 | EU | No mutation | | | 43 | 25 | 0.54 | 2.53 | Yes |
| 54 | M | 21 | EU | No mutation | | | 29 | 20 | 0.83 | 1.01 | Yes |
| 69 | M | 25 | EU | No mutation | | | 61 | 48.33 | 1.03 | 4.06 | Yes |
| 97 | M | 5 | EU | No mutation | | | 39 | 20 | 0.81 | 3.56 | Yes |
| 113 | F | 5 | EU | No mutation | | | 39 | 20 | 0.73 | 0.67 | No |
| 124 | M | 1 | EU | No mutation | | | 22 | 5 | 1.21 | 0.47 | ND |
| 127 | F | 7 | EU | No mutation | | | 49 | 26.67 | 0.77 | 3.8 | Yes |

*EU*, Europe; *F*, female; *M*, male; *ME*, Middle East; *ND*, not determined.
Patients with less than 0.5% $CD45RO^+$ $IL17^+$ IFN-$\gamma^-$ or 5% $CD45RO^+$ $IL17^-$ IFN-$\gamma^+$ cells are *shaded*.
* Scoring system described in Grimbacher et al.[4]
† Weighted score based on the 5 cardinal features (recurrent pneumonia, newborn rash, pathologic bone fractures, a characteristic face, and a cathedral palate) shown in Table E4. A total number of scaled points greater than 30.0 predicts a *STAT3* mutation.

to those of the entire cohort (see this article's Table E5 in the Online Repository at www.jacionline.org). Therefore, this set of diagnostic features can also be applied to children suspected of mutations in *STAT3*.

We applied the score on a control cohort of 28 patients with atopic dermatitis and IgE levels of >1000 IU/mL, and all had a weighted STAT3 score of <30 points.

In leave-1-out testing, this set of clinical features in our cohort of patients with suspected HIES had a sensitivity of 87.5% for predicting presence of a *STAT3* mutation, a specificity of 80.6%, and an error rate of 15%. On the replication set of 50 patients, this same set of features had a sensitivity of 93.9%, a specificity of 70.6%, and an error rate of 14%. Increasing the number of features did not improve the quality of our model.

None of the 11 features shown in Table E3 was significantly associated with the domain in which the *STAT3* mutation was located, hence no phenotype-genotype correlation can be reported.

## Using $T_H17$ cells in HIES as a diagnostic marker: IL-17 and IFN-γ production by $CD4^+$T cells in HIES

The induction of $T_H17$ cells after stimulation with either the superantigen SEB or the mitogen phorbol 12-myristate 13-acetate has been reported to be impaired in patients with HIES with *STAT3* mutations.[13-17] We analyzed the level of $T_H17$ cells among the PBMCs in 30 of our 100 patients with HIES and in 12 control subjects. Seventeen of these patients had mutations in *STAT3,* and 13 did not. Thirteen of the 17 patients harboring mutations in *STAT3* had less than 0.5% of IL-17–producing $CD4^+$ T cells. In contrast, all but 1 of our healthy donors had a frequency of more than 1%. The 13 patients without mutations in *STAT3* had significantly more IL-17–producing T cells than the STAT3-deficient patients ($P = .003$), but significantly fewer than the healthy controls ($P = .0001$); only 3 of these 13 patients had an extremely low frequency of IL-17–producing T cells, less than 0.5% (Fig 2). No correlation among the lack of $T_H17$ cells, the incidence of *Candida* infections, or the sex of the patients has been observed (Table III).

We measured the frequency of SEB-induced IFN-γ–producing $CD4^+$ T cells from PBMCs of the same set of subjects (Fig 2). Interestingly, the patients without a *STAT3* mutation had significantly fewer IFN-γ–producing $CD4^+$ T cells than did healthy controls ($P = .0001$). Of the 13 patients lacking *STAT3* mutations, only 1 had >5% IFN-γ–producing $CD4^+$ T cells. Patients with HIES carrying *STAT3* mutations also had significantly lower frequencies of IFN-γ–producing $CD4^+$ T cells than healthy controls ($P = .02$) but significantly higher frequencies than patients without *STAT3* mutation ($P = .007$).

We conclude that the lack of $T_H17$ cells is a useful predictive marker for heterozygous mutations in *STAT3*.

## DISCUSSION

In our cohort of 100 patients with suspected HIES, we found 18 novel mutations in *STAT3*. Because these were spread over the whole gene, it appears necessary to sequence the entire *STAT3* gene to exclude a possible mutation, an expensive process because of the size of the gene (Fig 3). However, the relative frequencies of mutations in Table II suggest that a cost-conscious strategy would be to sequence either cDNA or genomic DNA of *STAT3* in the following order: (1) the exons and intron junctions of the DNA-binding domain, (2) the exons and intron junctions of

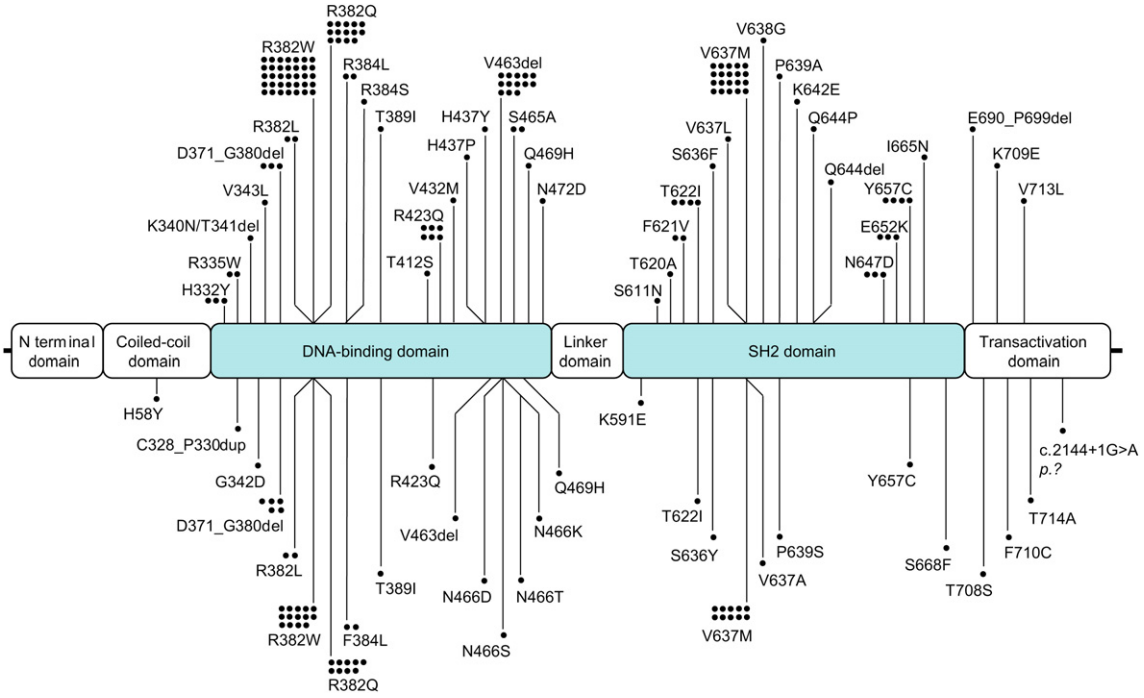


**FIG 3.** Schematic structure of STAT3. Every *black dot* represents 1 patient carrying the mutation. Patients who carry *STAT3* mutations and who were described in previous publications are shown in the *upper part* of the figure.[5,6,13-17] The 64 patients identified in this study who had *STAT3* mutations are shown in the *lower part*. The patient carrying a splice site mutation after exon 22 is shown as having DNA sequence change because the effect on the protein is not known. *p.?,* Unknown effect on protein level.

the SH2 domain, (3) the exons and intron junctions of the transactivation domain, and (4) the rest of the coding sequences. In all patients with HIES with detected *STAT3* mutations, only 1 allele was affected, consistent with the observation that complete loss of STAT3 leads to early embryonic death in STAT3 knockout mice.[26]

Mutations in the promoter region are possible, but are unlikely to cause HIES when present in heterozygosity because they will not exert a dominant-negative effect, which seems necessary for the HIES phenotype. [5]

Although more than 90% of the mutations we found are located in the exons of *STAT3*, we found 5 mutations at the intron-exon boundaries. Four of these mutations caused an in-frame deletion of the 10 amino acids encoded by exon 12. The effect of the splice site mutation on exon 22 is more complex. In healthy individuals, alternative splicing of exon 23 results in 2 major isoforms of STAT3: STAT3α and STAT3β. Whereas STAT3α contains the entire exon 23, STAT3β lacks the first 50 bp of the same exon.[27] The splice site introduced by the mutation 1 nucleotide after exon 22 causes the 43-bp exon to be skipped. In the wild-type STAT3α transcript, this deletion would result in a frame shift and possibly nonsense mediated decay. However, in conjunction with the shortened exon 23 of STAT3β, this deletion is predicted to produce a transcript coding for STAT3α with an in-frame deletion of amino acids 701 to 732, which would encompass both the tyrosine and serine phosphorylation sites of STAT3α (see this article's Figure E1 in the Online Repository at www.jacionline.org).

All 8 novel mutations that we tested deteriorated the suppression of LPS-induced production of TNF-α by the IL-10/Janus kinase/STAT3 pathway in the patients' macrophages. Compared with the most prevalent mutation affecting the DNA-binding domain of *STAT3* (R382Q), this effect was less prominent in some of the novel mutations. This result could reflect the fact that some mutations have a different impact on the retained functionality of the STAT3 dimerization, nuclear translocation, or transcriptional activation.

We confirm that patients with HIES carrying *STAT3* mutations have significantly reduced numbers of IL-17–producing $CD4^+$ T cells. On the basis of these data, we suggest that $T_H17$ cells may be used as an additional marker to distinguish between patients with HIES with or without *STAT3* mutations.

Patients without *STAT3* mutations had a striking reduction of IFN-γ–producing $CD4^+$ T cells. This cytokine imbalance has been previously described in HIES and may be explained by an intrinsic T-cell defect.[28] In our HIES cohort of *STAT3* wild-type patients, only 3 of 13 had less than 0.5% IL-17–producing T cells. These 3 patients may have defects in other proteins involved in STAT3 signaling or in the differentiation of $T_H17$ cells.

## Proposed set of diagnostic features for *STAT3*-mutant HIES

Hyper-IgE syndrome with mutations in *STAT3* is transmitted as an autosomal-dominant trait. Therefore, *STAT3*-mutated patients with HIES have a 50% risk of transmitting the disease to each of their offspring.[4,6] Thus, a positive family history contributes to the risk of having HIES. In addition clinical features have been observed to accumulate over time as affected children get older.[1] Hence, any diagnostic algorithm is likely to underdiagnose young patients. However, adding an age feature did not improve our SVM classifier.

On the basis of data from our multicenter cohort, we propose the following diagnostic guidelines for *STAT3*-mutant HIES:

- Possible: IgE ≥1000 IU/mL plus a weighted score of clinical features >30 based on recurrent pneumonia, newborn rash, pathologic bone fractures, characteristic face, and high palate (Table E4).
- Probable: These characteristics plus lack of $T_H17$ cells or a family history for definitive HIES.
- Definitive: These characteristics plus a dominant-negative heterozygous mutation in *STAT3*.

We caution, however, that none of the scores should be used to keep physicians from pursuing a molecular diagnosis in a particular patient. Patients with HIES accrue findings over time. Aggressive treatment with antibiotics can forestall infectious complications that would be diagnostic if allowed to occur. Moreover, the unique feature of HIES that abscesses or other infections are "cold," or not accompanied by a normal intensity of pain or inflammation, has not been captured by any score, but should increase clinical suspicion. HIES scoring will not replace good clinical judgment but may help the diagnostic process.

The aim of the proposed guidelines is to discern patients with HIES carrying mutations in *STAT3* to facilitate time-saving and resource-saving diagnosis of patients, thereby leading to an early and effective treatment. This screening tool will need to be evaluated and updated as new knowledge is revealed and new mutations causing HIES are identified. We acknowledge that most clinicians will continue to diagnose HIES with its key clinical features such as recurrent pneumonia, lung cysts, typical facies, retention of primary teeth, and pathological fractures. When it comes to the question whether to evaluate *STAT3,* however, this scoring should be used.

We thank the patients and their families, and we acknowledge the help of J. Milner in improving our $T_H17$-cell staining.

**Clinical implications: These diagnostic guidelines for STAT3-deficient HIES will facilitate time-saving and resource-saving diagnosis of the patients, thereby leading to an early and effective treatment.**

### REFERENCES

1. Grimbacher B, Holland SM, Gallin JI, Greenberg F, Hill SC, Malech HL, et al. Hyper-IgE syndrome with recurrent infections—an autosomal dominant multisystem disorder. N Engl J Med 1999;340:692-702.
2. Davis SD, Schaller J, Wedgwood RJ. Job's syndrome: recurrent, "cold", staphylococcal abscesses. Lancet 1966;287:1013-5.
3. Buckley RH, Wray BB, Belmaker EZ. Extreme hyperimmunoglobulinemia E and undue susceptibility to infection. Pediatrics 1972;49:59-70.
4. Grimbacher B, Schäffer AA, Holland SM, Davis J, Gallin JI, Malech HL, et al. Genetic linkage of hyper-IgE syndrome to chromosome 4. Am J Hum Genet 1999;65:735-44.
5. Minegishi Y, Saito M, Tsuchiya S, Tsuge I, Takada H, Hara T, et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. Nature 2007;448:1058-62.
6. Holland SM, DeLeo FR, Elloumi HZ, Hsu AP, Uzel G, Brodsky N, et al. *STAT3* mutations in the hyper-IgE syndrome. N Engl J Med 2007;357:1608-19.
7. Murray PJ. The JAK-STAT signaling pathway: input and output integration. J Immunol 2007;178:2623-9.
8. Reich NC, Liu L. Tracking STAT nuclear traffic. Nat Rev Immunol 2006;6:602-12.
9. Levy DE, Lee CK. What does Stat3 do? J Clin Invest 2002;109:1143-8.

10. Dong C. $T_H17$ cells in development: an updated view of their molecular identity and genetic programming. Nat Rev Immunol 2008;5:337-48.
11. Huang W, Na L, Fidel PL, Schwarzenberger P. Requirement of interleukin-17A for systemic anti-*Candida albicans* host defense in mice. J Infect Dis 2004;190:624-31.
12. Happel KI, Dubin PJ, Zheng M, Ghilardi N, Lockhart C, Quinton LJ, et al. Divergent roles of IL-23 and IL-12 in host defense against *Klebsiella pneumoniae*. J Exp Med 2005;202:761-9.
13. Renner ED, Rylaarsdam S, Aňover-Sombke S, Rack AL, Reichenbach J, Carey JC, et al. Novel signal transducer and activator of transcription 3 (STAT3) mutations, reduced $T_H17$ cell numbers, and variably defective STAT3 phosphorylation in hyper-IgE syndrome. Allergy Clin Immunol 2008;122:181-7.
14. Milner JD, Brenchley JM, Laurence A, Freeman AF, Hill BJ, Elias KM, et al. Impaired $T_H17$ cell differentiation in subjects with autosomal dominant hyper-IgE syndrome. Nature 2008;452:773-6.
15. Ma CS, Chew GYJ, Simpson N, Priyadarshi A, Wong M, Grimbacher B, et al. Deficiency of Th17 cells in hyper IgE syndrome due to mutations in *STAT3*. J Exp Med 2008;205:1551-7.
16. de Beaucoudrey L, Puel A, Filipe-Santos O, Cobat A, Ghandil P, Chrabieh M, et al. Mutations in *STAT3* and *IL12RB1* impair the development of human IL-17-producing T cells. J Exp Med 2008;205:1543-50.
17. Jiao H, Tóth B, Erdős M, Fransson I, Rákóczi E, Balogh I, et al. Novel and recurrent *STAT3* mutations in hyper-IgE syndrome patients from different ethnic groups. Mol Immunol 2008;46:202-6.
18. Renner ED, Puck JM, Holland SM, Schmitt M, Weiss M, Frosch M, et al. Autosomal recessive hyperimmunoglobulin E syndrome: a distinct disease entity. J Pediatr 2004;144:93-9.
19. Grimbacher B, Holland SM, Puck JM. Hyper-IgE syndromes. Immunol Rev 2005; 203:244-50.
20. Moin M, Farhoudi A, Movahedi M, Rezaei N, Pourpak Z, Yeganeh M, et al. The clinical and laboratory survey of Iranian patients with hyper-IgE syndrome. Scand J Infect Dis 2006;38:898-903.
21. Buettner M, Meinken C, Bastian M, Bhat R, Stössel E, Faller G, et al. Inverse correlation of maturity and antibacterial activity in human dendritic cells. J Immunol 2005;174:4203-9.
22. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2006.
23. Vapnik V. The nature of statistical learning theory. New York: Springer-Verlag; 1996.
24. Gertz EM, Wright SJ. Object-oriented software for quadratic programming. ACM TOMS 2003;29:58-81.
25. Joachims T. Making large-scale SVM learning practical. In: Schölkopf B, Burges C, Smola A, editors. Advances in kernel methods—support vector learning. Cambridge: MIT Press; 1999. p. 41-56.
26. Takeda K, Noguchi K, Shi W, Tanaka T, Matsumoto M, Yoshida N, et al. Targeted disruption of the mouse *Stat3* gene leads to early embryonic lethality. Proc Natl Acad Sci U S A 1997;94:3801-4.
27. Maritano D, Sugrue ML, Tininini S, Dewilde S, Strobl B, Fu XP, et al. The STAT3 isoforms $\alpha$ and $\beta$ have unique and specific functions. Nat Immunol 2004;5:401-9.
28. Netea MG, Kullberg BJ, van der Meer JWM. Severely impaired IL-12/IL-18/IFN$\gamma$ axis in patients with hyper IgE syndrome. Eur J Clin Invest 2005;35: 718-21.

# METHODS

## Statistical analysis

To perform leave-1-out testing, OOQP first trains an SVM by using the full set of observations and a relatively stringent termination tolerance of $10^{-5}$. If an observation is not a support vector of the SVM trained on the full set of observations, then omitting that observation and retraining the SVM produces the same linear classifier. Thus, we needed only to retrain the SVM using those observations that were support vectors—that is, those observations with an approximate Lagrange multiplier of at least $10^{-4}$. We retrained the SVM to the more relaxed tolerance of $10^{-3}$ when computing individual leave-1-out trials.

We chose as candidate feature sets those 344 sets for which OOQP reported leave-1-out accuracy of better than 80%. We repeated leave-1-out testing on these sets by using SVMlight (http://svmlight.joachims.org),[24] another package for generating SVMs, because OOQP and SVMlight use different termination criteria that are not entirely comparable. We programmed scripts to perform leave-1-out testing by calling SVMlight repeatedly.

For several feature sets, we tested the accuracy of the classifier generated by SVMlight on a replication set of 50 patients from NIH. In each of these tests, a classifier was trained by using the scores from our cohort of 100 patients and then applied to predict whether each of the 50 patients in the replication set had a *STAT3* mutation.

Both OOQP and SVMlight were run with a penalty of 1.0. Otherwise default values were used for SVMlight. By default, SVMlight uses a termination tolerance of $10^{-3}$.

## Set of clinical features best predicting the genotype

We found 12 feature sets with a leave-1-out accuracy of 85%. Of these, 1 had 5 features, 4 had 6 features each, and 7 consisted of 7 features. Table E3 shows the features that occurred more than once in the top feature sets. Pathological fractures and characteristic face occurred in all 12 sets. Moreover, at least 1 of the 2 related features, lung cysts or pneumonia, also occurred in each set. By logistic regression, pneumonia ($P = .002$), lung cysts ($P = .001$), and characteristic face ($P = .002$) were significantly positively associated with a *STAT3* mutation. The SVM weights for these 4 features were always positive in the sets in which they occurred. Midline congenital anomalies and upper respiratory tract infections were assigned negative weights by the SVM each time they occurred. Neither of these features achieved statistical significance in logistic regression.

Two of the 5 features in the best all-positive set were immunologic, and 3 were not. Importantly, criteria for these features were strict. Pneumonia was positive only if proven by radiograph; a newborn rash had to occur in the first 3 weeks of life; pathologic fractures were radiograph-proven fractures after unrecognized or minor trauma, defined as, for example, a fall from less than twice the patient's height; the characteristic facial features were assessed by an immunologist familiar with HIES; and a cathedral palate was defined as a palate higher than twice the height of a molar tooth.

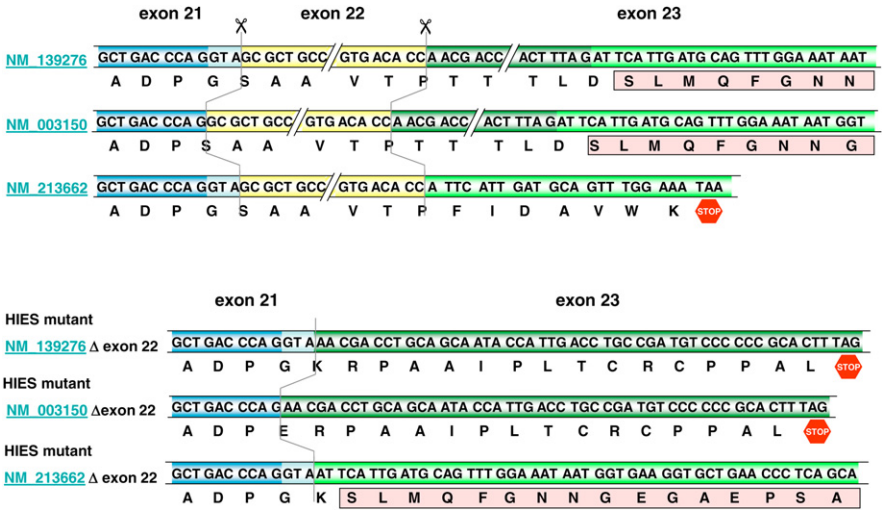


**FIG E1.** Two alternative splicing sites, 1 between exons 21 and 22 and 1 between exons 22 and 23, result in 3 alternative wild-type transcripts: 2 STAT3α transcripts (NM_139276, NM_003150) and 1 STAT3β (NM_213662) transcript. The splice site introduced by the mutation 1 nucleotide after exon 22 causes the 43-bp exon 22 to be skipped. In mutant variants of NM_139276 and NM_003150, this deletion would result in a frame shift and a premature stop-codon. In conjunction with the 50-bp shorter exon 23 of STAT3β, the deletion of exon 22 is predicted to produce a transcript that codes for a variant of STAT3α with an in-frame deletion of 31 amino acids. The amino acid sequences of the STAT3α transcripts shared with the mutated STAT3β transcript are highlighted in *pink*.

**TABLE E1.** Clinical findings and *STAT3* mutations of the 100 analyzed patients

| UPN | Sex | Age of scoring | Origin | DNA sequence change | Predicted amino acid change | Domain | NIH score* | STAT3 deficiency score† | Recurrent skin abscesses | Recurrent pneumonia | Lung cyst formation | Other unusual infections | Eosinophilia | Newborn rash | Eczema | Recurrent upper respiratory tract infections | Candidiasis | Failure to shed deciduous teeth | Scoliosis | Pathologic bone fractures | Hyperextensibility | Characteristic face for Job syndrome | Increased interalar distance | Cathedral palate | Midline anomaly |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 94 | F | 10 | EU | c.172C>T | H58Y | Coiled-coil | 42 | 8.33 | Yes | No | No | No | <700 | Yes | Severe | 3 | No | >3 | <10° | No | Yes | No | No | No | No |
| 110 | F | 28 | EU | c.1025G>A | G342D | DNA-binding | 79 | 63.33 | Yes | Yes | Yes | Yes | 700-800 | Yes | Severe | 4-6 | Nail | >3 | ND | Yes | Yes | Yes | Yes | Yes | No |
| 70 | F | 33 | EU | c.982_990dup | C328_P330dup | DNA-binding | 41 | 31.67 | Yes | Yes | No | No | <700 | No | Moderate | 3 | Oral | 2 | 10-14° | No | Yes | Yes | No | Yes | No |
| 7 | F | 12 | EU | c.1110-2A>G | D371_G380del | DNA-binding | 53 | 36.67 | Yes | Yes | Yes | No | >800 | No | Moderate | 1-2 | Nail | 2 | <10° | No | Yes | Yes | ND | ND | No |
| 114 | M | 22 | EU | c.1110-2A>G | D371_G380del | DNA-binding | 70 | 50.00 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 4-6 | Oral | ND | 10-14° | ND | ND | Yes | Yes | Yes | Yes |
| 86 | M | 12 | EU | c.1110-1G>T | D371_G380del | DNA-binding | 83 | 76.67 | Yes | Yes | Yes | Yes | <700 | Yes | Severe | 4-6 | Oral | >3 | 10-14° | Yes | Yes | Yes | ND | Yes | Yes |
| 17 | F | 24 | EU | c.1139+1G>T | D371_G380del | DNA-binding | 65 | 41.66 | Yes | Yes | Yes | Yes | 700-800 | ND | Severe | 1-2 | Nail | >3 | ND | No | No | Yes | Yes | Yes | No |
| 108 | M | 14 | EU | c.1139+2insT | D371_G380del | DNA-binding | 54 | 31.67 | No | Yes | Yes | Yes | >800 | No | Moderate | 1-2 | Systemic | >3 | <10° | No | No | Mild | No | Yes | No |
| 15 | M | 28 | EU | c.1144C>T | R382W | DNA-binding | 76 | 63.33 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 1-2 | Nail | >3 | ND | Yes | ND | Yes | Yes | Yes | No |
| 19 | M | 14 | EU | c.1144C>T | R382W | DNA-binding | 63 | 48.33 | Yes | Yes | Yes | Yes | <700 | Yes | Severe | 1-2 | Nail | >3 | <10° | Yes | Yes | Mild | Yes | Yes | No |
| 29 | F | 28 | SA | c.1144C>T | R382W | DNA-binding | 64 | 45.00 | Yes | Yes | Yes | Yes | >800 | Yes | Moderate | 4-6 | Nail | >3 | <10° | No | No | Yes | Yes | No | No |
| 31 | M | 12 | SA | c.1144C>T | R382W | DNA-binding | 57 | 41.67 | Yes | Yes | Yes | Yes | >800 | Yes | Moderate | 4-6 | Nail | 3 | ND | Yes | No | No | Yes | No | No |
| 43 | M | 20 | ME | c.1144C>T | R382W | DNA-binding | 62 | 40.00 | Yes | Yes | Yes | No | ND | Yes | Severe | 4-6 | Oral | >3 | <10° | No | Yes | Mild | Yes | Yes | No |
| 56 | M | 15 | EU | c.1144C>T | R382W | DNA-binding | 59 | 58.33 | Yes | Yes | Yes | No | >800 | Yes | Severe | ND | No | >3 | <10° | Yes | No | Yes | ND | ND | No |
| 63 | M | 27 | ME | c.1144C>T | R382W | DNA-binding | 61 | 35.00 | Yes | Yes | Yes | No | >800 | Yes | Mild | 1-2 | Nail | >3 | <10° | No | Yes | Mild | No | No | No |
| 73 | M | 17 | ME | c.1144C>T | R382W | DNA-binding | 73 | 45.00 | Yes | Yes | Yes | Yes | 700-800 | Yes | Severe | ND | ND | >3 | 15-20° | ND | Yes | Yes | Yes | ND | ND |
| 126 | F | 11 | EU | c.1144C>T | R382W | DNA-binding | 58 | 40.00 | Yes | Yes | Yes | Yes | >800 | Yes | Mild | 3 | Oral | 0 | 10-14° | No | No | Mild | Yes | Yes | ND |
| 128 | F | 2 | EU | c.1144C>T | R382W | DNA-binding | 29 | 38.33 | Yes | Yes | No | No | <700 | No | Moderate | 1-2 | Nail | 0 | <10° | Yes | No | Mild | Yes | No | No |
| 129 | M | 10 | EU | c.1144C>T | R382W | DNA-binding | 64 | 68.33 | Yes | Yes | Yes | Yes | <700 | No | Severe | 1-2 | Systemic | >3 | <10° | Yes | No | Yes | Yes | Yes | No |
| 133 | F | 40 | EU | c.1144C>T | R382W | DNA-binding | 74 | 71.67 | Yes | Yes | No | Yes | >800 | Yes | Mild | 4-6 | Nail | 0 | 10-14° | Yes | Yes | Yes | Yes | No | Yes |
| 134 | M | 5 | EU | c.1144C>T | R382W | DNA-binding | 48 | 40.00 | Yes | Yes | No | No | >800 | Yes | Severe | 1-2 | Oral | 0 | <10° | No | Yes | Yes | Yes | Yes | Yes |
| 135 | M | 7 | EU | c.1144C>T | R382W | DNA-binding | 60 | 50.00 | Yes | Yes | Yes | Yes | 700-800 | Yes | Severe | 3 | Oral | ND | | No | Yes | Yes | ND | Yes | No |
| 79 | F | 7 | EU | c.1145G>T | R382L | DNA-binding | 56 | 58.33 | Yes | Yes | No | No | >800 | Yes | Severe | 3 | Nail | 3 | <10° | Yes | ND | Yes | ND | ND | No |
| 98 | M | 25 | EU | c.1145G>T | R382L | DNA-binding | 45 | 36.67 | Yes | Yes | Yes | Yes | >800 | No | Moderate | 3 | Oral | ND | <10° | No | No | Yes | ND | No | No |
| 25 | F | 11 | EU | c.1145G>A | R382Q | DNA-binding | 72 | 48.33 | Yes | Yes | Yes | Yes | <700 | Yes | Severe | >6 | Systemic | >3 | <10° | Yes | Yes | Mild | No | No | No |
| 37 | F | 6 | EU | c.1145G>A | R382Q | DNA-binding | 46 | 41.66 | No | Yes | No | No | <700 | ND | Severe | >6 | No | >3 | <10° | No | Yes | Yes | Yes | Yes | No |
| 44 | M | 9 | ME | c.1145G>A | R382Q | DNA-binding | 62 | 35.00 | Yes | Yes | Yes | No | 700-800 | Yes | Moderate | >6 | ND | >3 | <10° | No | Yes | Mild | Yes | No | No |
| 47 | F | 24 | ME | c.1145G>A | R382Q | DNA-binding | 62 | 53.33 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 1-2 | No | ND | ND | Yes | No | Yes | Yes | No | ND |
| 61 | M | 14 | ME | c.1145G>A | R382Q | DNA-binding | 57 | 41.66 | Yes | Yes | Yes | ND | ND | No | Severe | >6 | Nail | >3 | ND | No | No | Yes | ND | Yes | No |
| 72 | M | 19 | ME | c.1145G>A | R382Q | DNA-binding | 60 | 28.33 | Yes | Yes | Yes | ND | 700-800 | Yes | Severe | 4-6 | Systemic | >3 | ND | ND | ND | No | Yes | ND | ND |
| 77 | M | 19 | ME | c.1145G>A | R382Q | DNA-binding | 61 | 50.00 | Yes | Yes | Yes | Yes | 700-800 | ND | Moderate | >6 | Nail | 0 | ND | Yes | ND | Yes | Yes | ND | No |
| 92 | M | 12 | ME | c.1145G>A | R382Q | DNA-binding | 50 | 45.00 | Yes | Yes | Yes | ND | <700 | Yes | Severe | ND | ND | ND | ND | ND | Yes | Yes | Yes | No | ND |
| 95 | F | 15 | EU | c.1145G>A | R382Q | DNA-binding | 48 | 40.00 | Yes | Yes | Yes | Yes | 700-800 | ND | Moderate | 1-2 | ND | ND | ND | Yes | Yes | Mild | Yes | ND | ND |
| 6 | F | 50 | EU | c.1150T>C | F384L | DNA-binding | 54 | 58.33 | Yes | Yes | Yes | No | <700 | Yes | Severe | ND | No | 0 | 10-14° | Yes | No | Yes | Yes | No | No |
| 60 | M | 29 | ME | c.1150T>C | F384L | DNA-binding | 71 | 55.00 | Yes | Yes | Yes | Yes | >800 | ND | Mild | >6 | Nail | >3 | <10° | Yes | Yes | Yes | Yes | Yes | ND |
| 65 | M | 23 | ME | c.1166C>T | T389I | DNA-binding | 52 | 45.00 | Yes | Yes | No | No | >800 | Yes | Severe | 1-2 | Nail | 2 | <10° | No | No | Yes | Yes | No | No |
| 118 | F | 31 | EU | c.1268G>A | R423Q | DNA-binding | 59 | 40.00 | Yes | Yes | Yes | No | <700 | No | Severe | 4-6 | No | 0 | 10-14° | Yes | Yes | Yes | Yes | No | Yes |
| 130 | F | 14 | EU | c.1387_1389del | V463del | DNA-binding | 56 | 41.66 | No | Yes | Yes | Yes | >800 | No | Moderate | 1-2 | No | >3 | <10° | No | No | Yes | Yes | Yes | No |
| 26 | F | 10 | EU | c. 1396 A>G | N466D | DNA-binding | 55 | 40.00 | Yes | Yes | No | Yes | <700 | Yes | Severe | 1-2 | Oral | >3 | <10° | No | Yes | Yes | Yes | Yes | No |
| 2 | F | 30 | EU | c.1397A>G | N466S | DNA-binding | 41 | 25.00 | No | Yes | Yes | No | >800 | No | Mild | 4-6 | No | >3 | ND | Yes | ND | Mild | ND | ND | ND |
| 46 | F | 15 | ME | c.1397A>C | N466T | DNA-binding | 66 | 36.67 | Yes | Yes | Yes | No | ND | No | Moderate | >6 | ND | >3 | >20° | No | Yes | Yes | Yes | No | No |
| 102 | M | 11 | EU | c.1398C>G | N466K | DNA-binding | 45 | 35.00 | Yes | Yes | No | No | <700 | No | Moderate | >6 | Oral | >3 | <10° | Yes | Yes | Yes | Yes | No | No |
| 36 | M | 16 | EU | c.1407G>T | Q469H | DNA-binding | 57 | 56.67 | Yes | No | No | Yes | ND | Yes | Severe | 1-2 | Oral | >3 | <10° | Yes | No | Yes | Yes | Yes | No |
| 57 | F | 18 | EU | c.1771A>G | K591E | SH2 | 75 | 61.67 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 1-2 | Nail | >3 | <10° | Yes | Yes | Yes | ND | ND | ND |
| 5 | M | 10 | EU | c.1865C>T | T622I | SH2 | 47 | 41.66 | Yes | Yes | Yes | No | >800 | No | Moderate | 1-2 | No | 2 | 10-14° | No | No | Yes | ND | Yes | No |
| 71 | M | 23 | EU | c.1907C>A | S636Y | SH2 | 71 | 40.00 | Yes | Yes | Yes | Yes | >800 | Yes | Mild | 1-2 | Oral | 2 | >20° | No | Yes | Mild | Yes | Yes | ND |
| 16 | F | 38 | EU | c.1909G>A | V637M | SH2 | 68 | 55.00 | Yes | Yes | Yes | No | >800 | No | Severe | 4-6 | Nail | >3 | ND | Yes | Yes | Yes | Yes | Yes | No |
| 32 | M | 14 | SA | c.1909G>A | V637M | SH2 | 37 | 25.00 | Yes | Yes | No | Yes | <700 | Yes | Moderate | 3 | No | 2 | <10° | No | No | Mild | Yes | No | Yes |
| 33 | M | 24 | SA | c.1909G>A | V637M | SH2 | 45 | 31.67 | Yes | Yes | Yes | No | <700 | No | Moderate | 4-6 | Oral | 0 | <10° | No | No | Yes | Yes | No | No |
| 42 | F | 11 | ME | c.1909G>A | V637M | SH2 | 58 | 41.66 | Yes | Yes | Yes | Yes | >800 | No | Moderate | 1-2 | No | >3 | ND | No | No | Yes | Yes | Yes | No |
| 45 | M | 9 | ME | c.1909G>A | V637M | SH2 | 44 | 18.33 | Yes | Yes | Yes | No | ND | Yes | Moderate | 4-6 | Oral | >3 | <10° | No | Yes | No | Yes | No | No |
| 66 | F | 20 | EU | c.1909G>A | V637M | SH2 | 56 | 40.00 | Yes | Yes | No | Yes | <700 | No | Severe | 1-2 | Nail | >3 | <10° | Yes | Yes | Yes | Yes | Yes | No |
| 67 | M | 12 | EU | c.1909G>A | V637M | SH2 | 59 | 53.33 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 1-2 | No | 1 | <10° | Yes | ND | Mild | No | Yes | No |
| 106 | M | 3 | EU | c.1909G>A | V637M | SH2 | 30 | 11.67 | Yes | Yes | No | Yes | >800 | No | Moderate | 3 | No | 0 | <10° | No | No | Mild | Yes | No | No |
| 115 | M | 19 | EU | c.1909G>A | V637M | SH2 | 56 | 68.33 | Yes | Yes | ND | ND | ND | ND | Severe | ND | ND | >3 | <10° | Yes | No | Yes | Yes | Yes | No |
| 125 | M | 15 | EU | c.1909G>A | V637M | SH2 | 58 | 45.00 | Yes | Yes | Yes | Yes | >800 | Yes | Severe | 4-6 | No | ND | 10-14° | No | No | Yes | Yes | No | ND |
| 104 | M | 5 | EU | c.1910T>C | V637A | SH2 | 72 | 50.00 | Yes | Yes | Yes | No | >800 | No | Severe | >6 | Nail | >3 | <10° | Yes | Yes | Yes | Yes | No | No |
| 49 | M | 7 | NA | c.1915T>C | P639S | SH2 | 55 | 35.00 | Yes | Yes | Yes | No | <700 | Yes | Severe | 4-6 | Nail | >3 | <10° | No | No | Mild | No | Yes | Yes |
| 109 | F | 21 | EU | c.1970A>G | Y657C | SH2 | 33 | 0.00 | Yes | No | No | Yes | >800 | No | Moderate | 3 | Nail | 0 | <10° | No | No | No | ND | No | No |
| 27 | M | 14 | SA | c.2003C>T | S668F | SH2 | 56 | 41.67 | Yes | Yes | Yes | Yes | >800 | Yes | Moderate | 4-6 | Oral | 2 | <10° | Yes | No | No | Yes | No | No |
| 103 | F | 22 | EU | c.2124C>G | T708S | Transactivation | 53 | 40.00 | No | Yes | Yes | No | >800 | Yes | Severe | 1-2 | Oral | >3 | <10° | ND | ND | Yes | Yes | ND | No |
| 18 | M | 22 | EU | c.2129T>C | F710C | Transactivation | 58 | 40.00 | Yes | Yes | Yes | No | >800 | Yes | Severe | 1-2 | Oral | 2 | <10° | No | No | Mild | Yes | Yes | No |
| 116 | M | 8 | EU | c.2141C>G | T714A | Transactivation | 63 | 50.00 | Yes | Yes | No | Yes | >800 | Yes | Severe | >6 | Oral | ND | 10-14° | No | Yes | Yes | Yes | Yes | ND |
| 96 | M | 9 | EU | c.2144+1G>A | *p.?* | Transactivation | 32 | 16.67 | No | Yes | Yes | No | >800 | ND | Moderate | 1-2 | No | ND | <10° | No | ND | Mild | ND | ND | No |

**TABLE E1.** (*Continued*)

| | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | M | 18 | EU | No mutation | 72 | 63.33 | Yes | Yes | No | Yes | >800 | Yes | Severe | >6 | Nail | >3 | 10-14° | Yes | No | Yes | Yes | Yes | No |
| 4 | M | 16 | EU | No mutation | 60 | 35.00 | Yes | No | No | No | >800 | No | Severe | 3 | Nail | >3 | 10-14° | Yes | Yes | Yes | Yes | Yes | Yes |
| 9 | M | 30 | EU | No mutation | 49 | 25.00 | Yes | No | No | No | 700-800 | Yes | Severe | 4-6 | Nail | >3 | <10° | No | No | Yes | Yes | No | No |
| 10 | M | 58 | EU | No mutation | 29 | 25.00 | Yes | No | No | No | <700 | Yes | Moderate | 3 | No | 2 | <10° | No | No | Yes | Yes | No | No |
| 11 | F | 54 | EU | No mutation | 55 | 53.33 | Yes | Yes | Yes | No | 700-800 | No | Moderate | >6 | Nail | 0 | <10° | Yes | No | Mild | No | No | No |
| 12 | F | 14 | EU | No mutation | 38 | 15.00 | Yes | No | No | ND | 700-800 | Yes | Severe | 4-6 | No | 3 | <10° | No | No | Mild | Yes | No | No |
| 14 | F | 33 | EU | No mutation | 47 | 21.67 | Yes | Yes | No | Yes | <700 | No | Severe | >6 | No | 0 | 10-14° | No | No | Yes | Yes | No | Yes |
| 20 | F | 8 | EU | No mutation | 32 | 21.67 | No | Yes | No | No | >800 | No | Mild | 1-2 | No | 3 | <10° | No | No | Mild | Yes | Yes | No |
| 21 | M | 8 | EU | No mutation | 49 | 26.67 | Yes | Yes | Yes | No | >800 | No | Severe | 3 | Oral | 2 | <10° | No | No | Mild | Yes | ND | No |
| 22 | F | 18 | EU | No mutation | 43 | 25.00 | Yes | No | No | No | >800 | Yes | Severe | 4-6 | Nail | >3 | <10° | No | No | Yes | No | No | No |
| 28 | M | 15 | SA | No mutation | 44 | 28.33 | Yes | Yes | Yes | No | <700 | Yes | Moderate | 3 | Oral | 2 | <10° | No | Yes | No | Yes | No | No |
| 30 | F | 14 | SA | No mutation | 48 | 18.33 | Yes | Yes | Yes | Yes | >800 | Yes | Moderate | 4-6 | Oral | ND | <10° | No | No | No | Yes | No | No |
| 35 | M | 7 | EU | No mutation | 39 | 20.00 | Yes | Yes | Yes | No | 700-800 | ND | Severe | ND | ND | 0 | <10° | No | Yes | No | No | Yes | No |
| 38 | M | 20 | EU | No mutation | 57 | 35.00 | Yes | Yes | Yes | ND | >800 | Yes | Severe | 1-2 | No | 2 | <10° | No | Yes | Yes | No | ND | No |
| 52 | M | 13 | EU | No mutation | 35 | 26.67 | Yes | Yes | No | No | >800 | No | Mild | 1-2 | No | 0 | <10° | No | No | Mild | No | No | No |
| 54 | M | 22 | EU | No mutation | 29 | 20.00 | No | Yes | No | No | >800 | No | No | 1-2 | Oral | 0 | <10° | No | No | No | No | No | No |
| 55 | M | 18 | EU | No mutation | 26 | 0.00 | Yes | No | No | No | >800 | No | Severe | 1-2 | No | 0 | 15-20° | No | No | No | No | No | No |
| 58 | F | 9 | EU | No mutation | 38 | 25.00 | No | Yes | Yes | No | <700 | Yes | Moderate | 1-2 | No | >3 | <10° | No | No | Mild | ND | No | No |
| 59 | M | 21 | EU | No mutation | 45 | 33.33 | Yes | Yes | Yes | Yes | <700 | Yes | Moderate | ND | No | 0 | 10-14° | No | No | ND | No | Yes | Yes |
| 69 | M | 25 | EU | No mutation | 61 | 48.33 | Yes | Yes | No | No | >800 | Yes | Severe | 4-6 | Nail | >3 | 10-14° | Yes | No | Yes | Yes | Yes | Yes |
| 76 | F | 29 | ME | No mutation | 42 | 20.00 | Yes | Yes | Yes | Yes | <700 | ND | Moderate | 4-6 | Nail | 0 | <10° | No | No | No | No | No | No |
| 78 | M | 6 | EU | No mutation | 46 | 45.00 | No | Yes | ND | No | >800 | Yes | Severe | 4-6 | Systemic | ND | <10° | No | No | Yes | Yes | Yes | No |
| 80 | M | 38 | ME | No mutation | 48 | 20.00 | Yes | Yes | Yes | No | >800 | ND | Severe | 4-6 | Systemic | ND | <10° | No | No | No | No | No | No |
| 81 | F | 8 | ME | No mutation | 44 | 26.67 | Yes | Yes | No | Yes | >800 | ND | Moderate | 3 | Nail | 0 | <10° | No | No | Mild | Yes | No | No |
| 82 | M | 10 | ME | No mutation | 63 | 28.33 | Yes | Yes | Yes | Yes | >800 | Yes | Moderate | >6 | Systemic | 3 | <10° | No | Yes | No | Yes | No | No |
| 83 | M | 9 | ME | No mutation | 40 | 21.67 | No | Yes | Yes | Yes | >800 | ND | Moderate | 4-6 | Nail | 0 | <10° | No | ND | Mild | ND | No | No |
| 88 | M | 15 | ME | No mutation | 53 | 26.67 | Yes | Yes | Yes | Yes | >800 | ND | Severe | 4-6 | No | ND | ND | ND | Yes | Mild | Yes | ND | ND |
| 90 | M | 15 | EU | No mutation | 29 | 11.67 | Yes | No | No | No | 700-800 | No | Severe | 3 | No | 2 | 10-14° | No | No | Mild | Yes | Yes | No |
| 97 | M | 4 | EU | No mutation | 39 | 20.00 | Yes | Yes | No | Yes | >800 | Yes | Severe | 1-2 | Oral | 0 | ND | No | ND | Mild | ND | ND | No |
| 105 | F | 7 | EU | No mutation | 19 | 13.33 | No | No | No | No | <700 | No | Mild | 1-2 | No | 0 | <10° | Yes | Yes | No | No | No | No |
| 112 | M | 4 | EU | No mutation | 21 | 6.67 | Yes | No | ND | No | 700-800 | ND | Severe | ND | ND | ND | ND | No | ND | Mild | ND | ND | No |
| 113 | F | 5 | EU | No mutation | 39 | 20.00 | No | No | No | Yes | >800 | Yes | Severe | 1-2 | No | 0 | <10° | No | Yes | Mild | Yes | Yes | ND |
| 122 | M | 11 | EU | No mutation | 29 | 10.00 | Yes | Yes | No | Yes | <700 | No | Moderate | 3 | No | 0 | <10° | No | No | No | No | Yes | No |
| 123 | F | 24 | EU | No mutation | 30 | 6.67 | Yes | No | No | Yes | <700 | No | Severe | 1-2 | No | 2 | <10° | No | No | Mild | No | No | No |
| 124 | M | 1‡ | EU | No mutation | 22 | 5.00 | No | No | No | No | >800 | No | Severe | 1-2 | ND | 0 | <10° | No | ND | No | ND | Yes | No |
| 127 | F | 6 | EU | No mutation | 49 | 26.67 | Yes | Yes | Yes | No | >800 | No | Severe | 1-2 | Oral | 2 | <10° | No | No | Mild | Yes | Yes | Yes |

*EU*, Europe; *F*, female; *M*, male; *ME*, Middle East; *NA*, North America; ND, not determined; *SA*, South America; *p.?,* unknown effect on protein level.
Patients positive for the 5 cardinal clinical features (recurrent pneumonia, newborn rash, pathologic bone fractures, a characteristic face, and a cathedral palate) are shaded in *blue*.
Patients with a mild facial phenotype are shaded in *light blue*.
*Scoring system described in Grimbacher et al.[4]
†Weighted score based on the 5 features (recurrent pneumonia, newborn rash, pathologic bone fractures, a characteristic face, and a cathedral palate) shown in Table E4. A total number of scaled points greater than 30 predicts the presence of a *STAT3* mutation. Patients with a score greater than 30 are shaded in *gray*.
‡UPN 124 was scored 8 days before his first birthday.

**TABLE E2.** Association of NIH-HIES score with mutation status

| Kind of mutation | Patients positive | Patients negative | Rank-sum *P* value |
|---|---|---|---|
| Any mutation in *STAT3* | 64 | 36 | 3.9e-7 |
| DNA binding domain | 42 | 22 | .094 |
| SH2 domain | 17 | 47 | NS |

The first column indicates the kind of mutations considered, the second column shows the number of patients positive for that mutation, and the third column shows the number of patients negative for that mutation. The fourth column is the *P* value of a Wilcoxon rank-sum test against the hypothesis that both groups have the same distribution of score. The rank-sum test for the first row includes all 100 patients in our cohort; the other 2 rows include only the 64 mutation-positive patients.

**TABLE E3.** The 11 features that occur more than once in the 12 feature sets with best leave-1-out accuracy

| Feature | Times included | Weight signs | *P* value |
|---|---|---|---|
| Recurrent pneumonia | 11 | Positive | .002 |
| Lung cyst formation | 3 | Positive | .001 |
| Other unusual infections | 3 | Mixed | NS |
| Newborn rash | 10 | Positive | NS |
| Upper respiratory tract infections | 4 | Negative | NS |
| Scoliosis | 3 | Positive | NS |
| Pathologic bone fractures | 12 | Positive | NS |
| Characteristic face for Job syndrome | 12 | Positive | .002 |
| Increased interalar distance | 5 | Mixed | NS |
| Cathedral palate | 10 | Positive | NS |
| Midline anomaly | 2 | Negative | NS |

The second column shows the number of sets that include the feature. The third column indicates whether the weight assigned the feature is positive or negative in each of 12 SVM classifiers in which that feature occurs. The features "increased interalar distance" and "other unusual infections" were sometimes assigned positive and sometimes assigned negative weights. The fourth column is the *P* value of the null hypothesis that the slope of the logistic regression function for the indicated feature is 0. The sign of the logistic regression function was positive for the features recurrent pneumonia, lung cyst formation, and characteristic face for Job syndrome.

**TABLE E4.** HIES STAT3 score

**STAT3-Score: Patient_____ DOB: _____ Scoring date: _____ Gender: _____**

| | | Points | | | | | |
|---|---|---|---|---|---|---|---|
| | **Clinical findings** | **0** | **2** | **4** | **5** | **6** | **8** |
| 1. | Pneumonias (X-ray proven, total #) | none | 1 | 2 | - | 3 | >3 |
| 2. | Newborn rash | absent | - | present | - | - | - |
| 3. | Pathologic bone fractures | none | - | 1-2 | - | - | >2 |
| 4. | Characteristic face for Job syndrome | absent | mild | - | present | - | - |
| 5. | Cathedral palate | absent | present | - | - | - | - |

| | | Points × | Scale | Scaled Points |
|---|---|---|---|---|
| 1. | Pneumonias | | 2.5 | |
| 2. | Newborn rash | | 2.08 | |
| 3. | Pathologic bone fractures | | 3.33 | |
| 4. | Characteristic face for Job syndrome | | 3.33 | |
| 5. | Cathedral palate | | 2.5 | |
| | Total (Sum 1-5) Scaled Points | | | |

This score sheet is intended to help predict whether a patient already known to have serum IgE above 1000 IU/mL is likely to have a mutation in *STAT3*. **A total number of scaled points greater than 30 predicts a *STAT3* mutation is likely.**

An example scoring a hypothetical patient with 3 pneumonias, newborn rash, no fractures, a characteristic face and normal palate:

**STAT3-Score: Patient_B.G. DOB: 23.01.1967 Scoring date: 22.08.2008 Gender: __M_**

| | | Points | | | | | |
|---|---|---|---|---|---|---|---|
| | **Clinical findings** | **0** | **2** | **4** | **5** | **6** | **8** |
| 1. | Pneumonias (X-ray proven, total #) | none | 1 | 2 | - | (3) | >3 |
| 2. | Newborn rash | absent | - | (present) | - | - | - |
| 3. | Pathologic bone fractures | (none) | - | 1-2 | - | - | >2 |
| 4. | Characteristic face for Job syndrome | absent | mild | - | (present) | - | - |
| 5. | Cathedral palate | (absent) | present | - | - | - | - |

| | | Points × | Scale | Scaled Points |
|---|---|---|---|---|
| 1. | Pneumonias | 6 | 2.5 | 15 |
| 2. | Newborn rash | 4 | 2.08 | 8.32 |
| 3. | Pathologic bone fractures | 0 | 3.33 | 0 |
| 4. | Characteristic face for Job syndrome | 5 | 3.33 | 16.65 |
| 5. | Cathedral palate | 0 | 2.5 | 0 |
| | Total (Sum 1-5) Scaled Points | | | 39.97 |

The score 39.97 is above the threshold of 30, and so the patient is predicted to have a *STAT3* mutation. Because the hypothetical patient does not have fractures or a cathedral palate, this shows that not all 5 of the features need to have a high score for the total weighted score to be above the threshold.

**TABLE E5.** Effectiveness of the best SVM classifier by age range in our cohort of 100 patients

| | Error | Sensitivity | Specificity | Positive | Negative |
|---|---|---|---|---|---|
| Age 0-8 y | 10 | 88.9 | 90.9 | 9 | 11 |
| Age 9-16 y | 13.2 | 85.2 | 90.3 | 27 | 11 |
| Age 17-30 y | 21.2 | 87.0 | 60.0 | 23 | 10 |
| Age 31+ y | 11.1 | 100.0 | 75.0 | 5 | 4 |

The first column shows an age range for each row, where age is determined at the date of scoring. The second, third, and fourth columns show the error rate, the sensitivity, and the specificity of the SVM classifier when applied to the subset of 100 patients in the given age range. The fifth and sixth columns list the number of individuals in the given age range that are either positive or negative for a *STAT3* mutation.