Diagnosis of Kawasaki Disease
Using a Minimal Whole-Blood Gene Expression Signature

Victoria J. Wright, PhD; Jethro A. Herberg, PhD; Myrsini Kaforou, PhD; Chisato Shimizu, MD; Hariklia Eleftherohorinou, PhD; Hannah Shailès, PhD; Anouk M. Barendregt, BSc; Stephanie Menikou, PhD; Stuart Gormley, MRes; Maurice Berk, PhD; Long Truong Hoang, PhD; Adriana H. Tremoulet, MD; John T. Kanegaye, MD; Lachlan J. M. Coin, PhD; Mary P. Glodé, MD; for the Immunopathology of Respiratory, Inflammatory and Infectious Disease Study (IRIS) Consortium and the Pediatric Emergency Medicine Kawasaki Disease Research Group (PEMKDRG)

IMPORTANCE To date, there is no diagnostic test for Kawasaki disease (KD). Diagnosis is based on clinical features shared with other febrile conditions, frequently resulting in delayed or missed treatment and an increased risk of coronary artery aneurysms.

OBJECTIVE To identify a whole-blood gene expression signature that distinguishes children with KD in the first week of illness from other febrile conditions.

DESIGN, SETTING, AND PARTICIPANTS The case-control study comprised a discovery group that included a training and test set and a validation group of children with KD or comparator febrile illness. The setting was pediatric centers in the United Kingdom, Spain, the Netherlands, and the United States. The training and test discovery group comprised 404 children with infectious and inflammatory conditions (78 KD, 84 other inflammatory diseases, and 242 bacterial or viral infections) and 55 healthy controls. The independent validation group comprised 102 patients with KD, including 72 in the first 7 days of illness, and 130 febrile controls. The study dates were March 1, 2009, to November 14, 2013, and data analysis took place from January 1, 2015, to December 31, 2017.

MAIN OUTCOMES AND MEASURES Whole-blood gene expression was evaluated using microarrays, and minimal transcript sets distinguishing KD were identified using a novel variable selection method (parallel regularized regression model search). The ability of transcript signatures (implemented as disease risk scores) to discriminate KD cases from controls was assessed by area under the curve (AUC), sensitivity, and specificity at the optimal cut point according to the Youden index.

RESULTS Among 404 patients in the discovery set, there were 78 with KD (median age, 27 months; 55.1% male) and 326 febrile controls (median age, 37 months; 56.4% male). Among 202 patients in the validation set, there were 72 with KD (median age, 34 months; 62.5% male) and 130 febrile controls (median age, 17 months; 56.9% male). A 13-transcript signature identified in the discovery training set distinguished KD from other infectious and inflammatory conditions in the discovery test set, with AUC of 96.2% (95% CI, 92.5%-99.9%), sensitivity of 81.7% (95% CI, 60.0%-94.8%), and specificity of 92.1% (95% CI, 84.0%-97.0%). In the validation set, the signature distinguished KD from febrile controls, with AUC of 94.6% (95% CI, 91.3%-98.0%), sensitivity of 85.9% (95% CI, 76.8%-92.6%), and specificity of 89.1% (95% CI, 83.0%-93.7%). The signature was applied to clinically defined categories of definite, highly probable, and possible KD, resulting in AUCs of 98.1% (95% CI, 94.5%-100%), 96.3% (95% CI, 93.3%-99.4%), and 70.0% (95% CI, 53.4%-86.6%), respectively, mirroring certainty of clinical diagnosis.

CONCLUSIONS AND RELEVANCE In this study, a 13-transcript blood gene expression signature distinguished KD from other febrile conditions. Diagnostic accuracy increased with certainty of clinical diagnosis. A test incorporating the 13-transcript disease risk score may enable earlier diagnosis and treatment of KD and reduce inappropriate treatment in those with other diagnoses.
Kawasaki disease (KD) is an acute inflammatory disorder predominantly seen in young children. Since it was first described in Japan,1 KD has emerged as the most common cause of acquired heart disease, with an incidence in children younger than 5 years ranging from 265 cases per 100,000 in Japan,2 to 51 to 194 cases per 100,000 in other Asian countries,3,5 to 8 to 20 cases per 100,000 in Europe6 and the United States,7 respectively. What makes KD of such concern is its association with vasculitis, affecting predominantly the coronary arteries, which results in coronary artery aneurysms (CAAs) in up to 25% of untreated children.8 Death from myocardial infarction may occur due to thrombotic occlusion of the aneurysms or from the later development of stenotic lesions due to vascular remodeling in the damaged artery. Long-term outcome studies9,10 of children with giant CAAs indicate a worrisome prognosis, with more than 50% needing revascularization or experiencing myocardial infarction within a 30-year period.

Treatment with intravenous immunoglobulin (IVIG) and, for those who do not respond, additional IVIG11 or other anti-inflammatory agents, such as corticosteroids and infliximab, is effective in abrogating the inflammatory process and reduces the risk of CAAs to 5% to 10%.12 Because KD is difficult to distinguish from other common febrile conditions, many children with KD are not diagnosed and treated early enough to prevent development of CAAs.13 Furthermore, patients who do not fulfill the clinical criteria for diagnosing KD (so-called incomplete KD) may experience CAAs. Delayed diagnosis is a consistent risk factor for development of CAAs, and treatment is often commenced only when coronary dilatation is already demonstrated on echocardiography.

The symptoms of KD are similar to those of several other childhood febrile illnesses, including staphylococcal and streptococcal toxic shock syndromes, measles and other viral illnesses (eg, adenovirus infection, Rocky Mountain spotted fever), and childhood inflammatory diseases, leading to diagnostic difficulty and thus delay in diagnosis and treatment. Guidelines have been developed to facilitate diagnosis based on clinical signs and symptoms, echocardiography, and laboratory variables,14 but there remains an urgent need for an accurate test to distinguish KD from other conditions causing prolonged fever in children.

In the era of precision medicine, diagnosis of many conditions previously based on clinical features alone is being replaced by diagnosis based on molecular pathology. Host blood gene expression signatures have been shown to identify several specific infectious and inflammatory diseases, including tuberculosis,15 bacterial and viral infections,16,17 and systemic lupus erythematosus.18 Support for a diagnostic approach to KD based on gene expression signatures comes from identification of microRNA biomarkers in KD,19,20 although existing studies are limited by the range of comparator patients or a need to extract RNA from exosomes. We explored use of whole-blood gene expression patterns to distinguish KD from other childhood infectious and inflammatory conditions. We present a gene expression signature, discovered and validated in independent patient groups, that distinguishes KD from a range of bacterial, viral, and inflammatory illnesses.

Methods

Patient Study Groups
The differential diagnosis for KD includes multiple infectious and inflammatory conditions. Therefore, in this case-control study, we established a discovery group of children with KD and a range of other infectious and inflammatory diseases with clinical signs, inflammatory markers, and duration of fever overlapping KD. Patients were prospectively recruited at pediatric centers in the United Kingdom, Spain, the Netherlands, and the United States if they had febrile illness and required blood testing for clinical investigation as part of the UK-based Immunopathology of Respiratory, Inflammatory and Infectious Disease (IRIS) study; the Spanish GENDRES (Genetic, Vitamin D, and Respiratory Infections Research Network) study (http://www.gendres.org); the Dutch Kawasaki Study; or the US-based Kawasaki Disease Research Center Program (https://medschool.ucsd.edu/som/pediatrics/research/centers/kawasaki-disease). The training and test discovery group comprised 404 children with infectious and inflammatory conditions (78 KD, 84 other inflammatory diseases, and 242 bacterial or viral infections) and 55 healthy controls. The independent validation group comprised 102 patients with KD, including 72 in the first 7 days of illness, and 130 febrile controls. The study dates were March 1, 2009, to November 14, 2013, and data analysis took place from January 1, 2015, to December 31, 2017.

Children with KD represented a combination of those seen directly in emergency departments and patients referred from regional centers. Our study included only patients recruited before initiation of IVIG for treatment. For discovery of a diagnostic signature, we included patients with KD in the first 7 days of illness because we aimed to develop a test for use early in the illness before coronary artery damage occurs. However, to explore...
the performance of the signature at all stages of illness, we included patients up to day 10 of their illness in the validation study.

Febrile controls whose duration of illness before hospital presentation varied were recruited, with blood samples collected as soon as possible after presentation and before clinical diagnosis was confirmed. Febrile controls were assigned to diagnostic groups using predefined criteria once the results of all investigations were available (Figure 1 and eMethods in the Supplement). Children with comorbidities likely to influence gene expression, such as immunosuppressive treatments, were excluded. We included comparator groups of children seen with inflammatory illness, including juvenile idiopathic arthritis and Henoch-Schönlein purpura. Comparison of the duration of illness, inflammatory markers, and demographics between patients with KD and febrile controls is summarized in Table 1.

Patients in the validation study group were similarly recruited as part of biomarker studies of febrile children seen in the hospital and requiring blood tests, as described previously.21,22 Healthy control children with no recent history of fever or immunization were recruited alongside patients with KD and febrile control patients in the discovery and validation studies. Data from healthy controls were used to standardize data obtained in different microarray experiments but were not used to evaluate the performance of the signature.

KD Case Definition
Kawasaki disease was diagnosed on the basis of the American Heart Association criteria,14 with 2-dimensional echocardiography performed soon after presentation (2 and 6 weeks after onset). Patients with fewer than 4 of the 5 classic criteria (bilateral nonpurulent conjunctivitis, oral mucosal changes, peripheral extremity changes, rash, and cervical lymphadenopathy >1.5 cm) were included as having incomplete KD if the maximum coronary artery z score (Zmax) (standard deviation units from the mean internal diameter normalized for body surface area) at any time during the illness for the left anterior descending or right coronary arteries was 2.5 or higher or if the patients satisfied the algorithm for incomplete KD in the American Heart Association guidelines. Patients were classified as having normal (Zmax <2.5), small (Zmax 2.5 to <5.0), or large (Zmax ≥5.0) CAAs. Because of interoperator variability in coronary artery dimensions, we set a high (Zmax ≥5.0) threshold to define patients with confirmed aneurysms and thus definite diagnosis of KD.
Further Classification of KD by Diagnostic Certainty

Because there is no gold standard for diagnosis, some patients may meet the criteria for KD but have other conditions. Therefore, we further categorized patients with KD in the validation study group based on certainty of clinical diagnosis. All clinical records, laboratory results, echocardiogram reports, response to treatment, and follow-up were reviewed by an independent pediatric infectious disease specialist and expert on KD (M.P.G.) masked to the analysis. Patients with documented CAAs (Zmax ≥5.0) persisting 6 weeks after onset were considered to have definite KD because there is no other known self-resolving inflammatory illness in childhood that leads to CAAs. The remaining patients (all of whom were treated with IVIG for suspected KD) were classified as having highly probable, possible, or unlikely KD by the expert reviewer. This review identified no unlikely KD cases.

Febrile Control Children

With Infection or Other Inflammatory Syndromes

Children seen with febrile illnesses were recorded as having definite bacterial infection, definite viral infection, suspected bacterial or viral infection, juvenile idiopathic arthritis, or Henoch-Schönlein purpura. The criteria shown in Figure 1 and described in eMethods in the Supplement were used to make this determination.

Oversight and Conduct of the Study

Patients were categorized into disease groups (Figure 1) after evaluation of all results by 2 independent clinicians not involved in the patients’ care (J.A.H., A.M.B., J.T.K., M.P.G., and J.C.B.). All blood samples were anonymized, and transcriptomic data sets were analyzed only after clinical assignments were finalized and dispatched for independent verification (eMethods in the Supplement).

Discovery and Validation of the Gene Expression Signature

The overall study design and signature discovery pipeline are shown in Figure 2. Whole blood was collected at the time of recruitment into blood RNA tubes (PAXgene; PreAnalytiX), frozen, extracted, and analyzed on arrays (HumanHT-12 version 4.0 BeadChip; Illumina). An earlier array (HumanHT-12 version 3.0 BeadChip; Illumina) with

Table 1. Clinical Characteristics and Initial Laboratory Values for Patients With Kawasaki Disease and Febrile Controls in Discovery and Validation Study Groups*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Discovery Set Kawasaki Disease</th>
<th>Febrile Controlsa</th>
<th>Validation Set Kawasaki Diseaseb</th>
<th>Febrile Controlsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>78</td>
<td>326</td>
<td>72</td>
<td>130</td>
</tr>
<tr>
<td>Age, median (IQR), mo</td>
<td>27 (16 to 45)</td>
<td>37 (9 to 116)</td>
<td>34 (17 to 51)</td>
<td>17 (5 to 47)</td>
</tr>
<tr>
<td>Male sex, No. (%)</td>
<td>43 (55.1)</td>
<td>184 (56.4)</td>
<td>45 (62.5)</td>
<td>74 (56.9)</td>
</tr>
<tr>
<td>Illness day at sample collection, median (IQR)</td>
<td>5 (4 to 6)</td>
<td>6 (4 to 9)</td>
<td>5 (5 to 6)</td>
<td>5 (3 to 7)</td>
</tr>
<tr>
<td>Laboratory values, median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin z score</td>
<td>−1.3 (−2.0 to −0.3)</td>
<td>NA</td>
<td>−1.2 (−2.0 to −0.4)</td>
<td>NA</td>
</tr>
<tr>
<td>C-reactive protein, mg/L</td>
<td>119 (48 to 192)</td>
<td>66 (23 to 174)</td>
<td>87 (59 to 173)</td>
<td>62 (16 to 162)</td>
</tr>
<tr>
<td>Platelet count, ×10³/μL</td>
<td>352 (303 to 448)</td>
<td>254 (167 to 351)</td>
<td>408 (324 to 474)</td>
<td>277 (176 to 352)</td>
</tr>
<tr>
<td>White blood cell count, /μL</td>
<td>14 200 (10 400 to 18 300)</td>
<td>8000 (6000 to 12 900)</td>
<td>13 900 (11 000 to 19 000)</td>
<td>11 000 (7700 to 16 000)</td>
</tr>
<tr>
<td>Neutrophil count, /μL</td>
<td>9000 (6600 to 12 400)</td>
<td>5000 (3100 to 9400)</td>
<td>10000 (7300 to 12600)</td>
<td>7000 (3600 to 13400)</td>
</tr>
<tr>
<td>Ethnicity, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. not statedf</td>
<td>0</td>
<td>23 (7.1)</td>
<td>0</td>
<td>10 (8.3)</td>
</tr>
<tr>
<td>African</td>
<td>3 (3.8)</td>
<td>28 (8.6)</td>
<td>2 (2.8)</td>
<td>23 (19.2)</td>
</tr>
<tr>
<td>Asian, including Indian subcontinent and Far East</td>
<td>12 (15.4)</td>
<td>29 (8.9)</td>
<td>12 (16.7)</td>
<td>12 (10.0)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>20 (25.6)</td>
<td>186 (57.1)</td>
<td>20 (27.8)</td>
<td>68 (56.7)</td>
</tr>
<tr>
<td>Mixed</td>
<td>15 (19.2)</td>
<td>28 (8.6)</td>
<td>23 (31.9)</td>
<td>7 (5.8)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (3.8)</td>
<td>12 (3.7)</td>
<td>1 (1.4)</td>
<td>10 (8.3)</td>
</tr>
<tr>
<td>Coronary artery status, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>45 (57.7)</td>
<td>NA</td>
<td>52 (72.2)</td>
<td>NA</td>
</tr>
<tr>
<td>Dilated</td>
<td>25 (32.1)</td>
<td>NA</td>
<td>15 (20.8)</td>
<td>NA</td>
</tr>
<tr>
<td>Aneurysm</td>
<td>8 (10.3)</td>
<td>NA</td>
<td>5 (6.9)</td>
<td>NA</td>
</tr>
<tr>
<td>IVIG resistant, No. (%)</td>
<td>18 (23.1)</td>
<td>NA</td>
<td>15 (20.8)</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; IVIG, intravenous immunoglobulin; NA, not applicable.

SI conversion factors: To convert C-reactive protein level to nanomoles per liter, multiply by 9.524; neutrophil count to ×10⁹/L, multiply by 0.001; platelet count to ×10⁹/L, multiply by 1.0; and white blood cell count to ×10⁹/L, multiply by 0.001.

* There were no significant differences between patients with Kawasaki disease in the discovery and validation sets.

a Healthy controls were not included.

b Data refer to the 72 patients in the first week of Kawasaki disease.

c Illness day 1 is the first day of fever (in Kawasaki disease) or symptoms (in febrile controls).

d Hemoglobin was normalized by age (data unavailable for febrile controls).

e Ethnicity percentages were calculated for the denominator with recorded data.
largely overlapping probes was used in a subset of the validation study group. Details of laboratory methods are provided in eMethods in the Supplement.

**Statistical Analysis**

**Transcript Signature Discovery**

Analysis of the transcriptomic data was conducted with statistical software (R, version 3.2.2; R Foundation for Statistical Computing). As shown in Figure 2, the discovery study group was randomly divided into an 80% training set and a 20% test set. The signature was identified in the training set and validated in the test set and in the validation study group, established using previously reported acute and convalescent patients with KD21 and acute bacterial and viral patients22 (eMethods in the Supplement). After quality control and filtering (eMethods in the Supplement), significantly differentially expressed transcripts in patients with KD compared with all other diseases were identified in the training set.

**Small Signature Discovery Using Parallel Regularized Regression Model Search**

A range of statistical methods are available to identify signatures from significantly differentially expressed transcripts, including least absolute shrinkage and selection operator (LASSO)24 and elastic net.25 However, these approaches produce large signatures that may not be easy to translate into a bedside diagnostic test. Therefore, we developed a novel variable selection method, parallel regularized regression model search, that identifies and ranks transcript signatures on the basis of their least number of transcripts and highest accuracy in discrimination26 (eMethods in the Supplement). The method first evaluates all possible 1- and 2-transcript models distinguishing KD from comparator diseases based on all significantly differentially expressed transcripts and takes the 100 best-fitting 2-transcript models to the next round, when a further transcript is added to the model and all combinations are again evaluated. The process continues with the incremental

---

**Figure 2. Study Design**

The overall study pipeline shows sample handling, derivation of test and training data sets, data processing, and analysis pipeline. Version 3 arrays indicate HumanHT-12, version 3.0 BeadChip (Illumina); version 4 arrays indicate HumanHT-12, version 4.0 BeadChip (Illumina); and ComBat indicates the ComBat algorithm.23 DB indicates definite bacterial; DV, definite viral; FC, fold change; HC, healthy controls; HSP, Henoch-Schönlein purpura; JIA, juvenile idiopathic arthritis; KD, Kawasaki disease; PReMS, parallel regularized regression model search; SDE, significantly differentially expressed; and U, infections of uncertain bacterial or viral etiology.

- See Supplemental Methods (RNA sample extraction and processing), as well as Statistical Methods in eMethods in the Supplement.
- Healthy controls were used in model building but were excluded from estimates of model accuracy.
- See Statistical Methods in eMethods in the Supplement; 146 acute KD samples (HumanHT-12, version 4.0) were used in Combat, of which 101 were taken forward.
- Diagnostic performance was assessed on 72 patients (within the first 7 days of illness).
- Includes convalescent KD and healthy controls.
regulated transcripts and might facilitate development of diagnostic signatures. The DRS combines a method that assigns individual disease risk based on the disease risk score (DRS) was determined by cross-validation. We applied the previously reported DRS implemented as a DRS, had a diagnostic performance in the discovery test set distinguishing KD from other infectious diseases and healthy controls (defined as log2 fold change >1 in KD vs at least 1 of the comparator groups). To identify small signatures suitable for developing as a diagnostic test, we next undertook variable selection using parallel regularized regression model search. Genes with positive values show increased expression in Kawasaki disease relative to other diseases, and genes with negative values show decreased expression in Kawasaki disease.

Results

The numbers of patients in each diagnostic category are shown in Figure 2. Clinical and demographic features of patients with KD and febrile controls are summarized in Table 1, with further details of control patients listed in eTable 1 in the Supplement. Principal component analysis of the normalized transcript expression profiles was performed separately on the discovery (training and test) and validation groups (eFigure 1 and eFigure 2 in the Supplement). Study groups clustered together in the discovery group and in the validation group after combining KD and case-control data using the ComBat algorithm (Statistical Methods in eMethods in the Supplement).

Identification of Minimal Transcript Signatures

In total, 1600 transcripts passed quality control and were significantly differentially expressed between KD and all other diseases and healthy controls (defined as log2 fold change >1 in KD vs at least 1 of the comparator groups). To identify small signatures suitable for developing as a diagnostic test, we next undertook variable selection using parallel regularized regression model search. This approach identified a 13-transcript signature (Table 2) that, when implemented as a DRS, had a diagnostic performance in the discovery test set distinguishing KD from other infectious and inflammatory conditions, with an AUC of 96.2% (95% CI, 92.5%-99.9%), sensitivity of 81.7% (95% CI, 60.0%-94.8%), and specificity of 92.1% (95% CI, 84.0%-97.0%) (Figure 3A and B).

Signature Performance in Validation Set

When the signature was applied to all of the 72 KD cases in the validation set, who were in the first 7 days of illness, the AUC was 94.6% (95% CI, 91.3%-98.0%), with sensitivity of 85.9% (95% CI, 76.8%-92.6%) and specificity of 93.7%. The performance was slightly reduced in the 30 patients diagnosed later (days 8-10) (eTable 2 and eFigure 3 in the Supplement).

Because clinical features of KD overlap those of other conditions and because any KD study group is likely to include patients misclassified as KD, we assessed whether certainty of clinical diagnosis corresponded to the predictive performance of the KD DRS. The performance of the 13-transcript signature in the patients with definite, highly probable, or possible KD in the validation set mirrored certainty of clinical diagnosis, with AUCs of 98.1% (95% CI, 94.5%-100%), 96.3% (95% CI, 93.3%-99.4%), and 70.0% (95% CI, 53.4%-86.6%), respectively (Figure 3C and D and eTable 2 in the Supplement).
Diagnosis of Kawasaki Disease Using a Minimal Whole-Blood Gene Expression Signature

Original Investigation Research

enesses and because the consequences of delayed treatment in the differential diagnosis of many childhood febrile illnesses. Conversely, because the diagnosis of KD is considered may be delayed while awaiting exclusion of other conditions, treatment with IVIG for presumed KD is often initiated without confirmation of diagnosis.

The diagnosis of KD now relies on the presence of 4 of the 5 characteristic clinical criteria. Fewer criteria are accepted if coronary artery abnormalities (dilatation or aneurysms) are detected on echocardiography. Children with incomplete KD who do not fulfill the classic diagnostic criteria but have prolonged fever and inflammation are at an increased risk of developing CAAs.33 One reason for the greater risk of CAAs in incomplete KD is the delayed diagnosis that often occurs in patients lacking all clinical features. Because the clinical features of KD overlap those of many other common childhood conditions, treatment with IVIG may be delayed while awaiting exclusion of other conditions. Conversely, because the diagnosis of KD is considered in the differential diagnosis of many childhood febrile illnesses and because the consequences of delayed treatment may be severe, overtreatment with IVIG or immunosuppressive second-line treatments may occur. A diagnostic test that accurately distinguishes KD from other infectious and inflammatory processes would be a significant advance in management of the disorder, reduce unnecessary investigations and inappropriate treatments, and enable earlier treatment with IVIG and other anti-inflammatory agents.

In establishing our discovery and validation study groups, we aimed to include a wide range of disorders with similar features overlapping those of KD, including both infectious and inflammatory diseases. The signature that we have identified distinguished KD from a wide range of other conditions with similar duration of fever and overlapping levels of inflammation. Because KD is diagnosed based on a constellation of clinical features and because there is no gold standard for diagnosis, evaluation of biomarker test results is difficult. Any cohort of children treated with IVIG for presumed KD is likely to include some patients with non-KD illness but with similar features. To evaluate the correspondence of the KD DRS with levels of diagnostic certainty, we categorized patients in the validation set as having definite, highly probable KD; KD-P, possible KD; and U, infections of uncertain bacterial or viral etiology.

Discussion

We identified a 13-transcript signature that distinguishes patients with KD from patients with bacterial, viral, and inflammatory illnesses. The high sensitivity and specificity of this signature for early diagnosis of KD suggests that it might form the basis of a diagnostic test. Our findings herein extend previous gene expression studies21,28-32 in KD that focused on immunopathogenesis. The diagnosis of KD now relies on the presence of 4 of the 5 characteristic clinical criteria. Fewer criteria are accepted if coronary artery abnormalities (dilatation or aneurysms) are detected on echocardiography. Children with incomplete KD who do not fulfill the classic diagnostic criteria but have prolonged fever and inflammation are at an increased risk of developing CAAs.33 One reason for the greater risk of CAAs in incomplete KD is the delayed diagnosis that often occurs in patients lacking all clinical features. Because the clinical features of KD overlap those of many other common childhood conditions, treatment with IVIG may be delayed while awaiting exclusion of other conditions. Conversely, because the diagnosis of KD is considered in the differential diagnosis of many childhood febrile illnesses and because the consequences of delayed treatment may be severe, overtreatment with IVIG or immunosuppressive second-line treatments may occur. A diagnostic test that accurately distinguishes KD from other infectious and inflammatory processes would be a significant advance in management of the disorder, reduce unnecessary investigations and inappropriate treatments, and enable earlier treatment with IVIG and other anti-inflammatory agents.

In establishing our discovery and validation study groups, we aimed to include a wide range of disorders with features overlapping those of KD, including both infectious and inflammatory diseases. The signature that we have identified distinguished KD from a wide range of other conditions with similar duration of fever and overlapping levels of inflammation. Because KD is diagnosed based on a constellation of clinical features and because there is no gold standard for diagnosis, evaluation of biomarker test results is difficult. Any cohort of children treated with IVIG for presumed KD is likely to include some patients with non-KD illness but with similar features. To evaluate the correspondence of the KD DRS with levels of diagnostic certainty, we categorized patients in the validation set as having definite, highly probable KD; KD-P, possible KD; and U, infections of uncertain bacterial or viral etiology.

Discussion

We identified a 13-transcript signature that distinguishes patients with KD from patients with bacterial, viral, and inflammatory illnesses. The high sensitivity and specificity of this signature for early diagnosis of KD suggests that it might form the basis of a diagnostic test. Our findings herein extend previous gene expression studies21,28-32 in KD that focused on immunopathogenesis. The diagnosis of KD now relies on the presence of 4 of the 5 characteristic clinical criteria. Fewer criteria are accepted if coronary artery abnormalities (dilatation or aneurysms) are detected on echocardiography. Children with incomplete KD who do not fulfill the classic diagnostic criteria but have prolonged fever and inflammation are at an increased risk of developing CAAs.33 One reason for the greater risk of CAAs in incomplete KD is the delayed diagnosis that often occurs in patients lacking all clinical features. Because the clinical features of KD overlap those of many other common childhood conditions, treatment with IVIG may be delayed while awaiting exclusion of other conditions. Conversely, because the diagnosis of KD is considered in the differential diagnosis of many childhood febrile illnesses and because the consequences of delayed treatment may be severe, overtreatment with IVIG or immunosuppressive second-line treatments may occur. A diagnostic test that accurately distinguishes KD from other infectious and inflammatory processes would be a significant advance in management of the disorder, reduce unnecessary investigations and inappropriate treatments, and enable earlier treatment with IVIG and other anti-inflammatory agents.

In establishing our discovery and validation study groups, we aimed to include a wide range of disorders with features overlapping those of KD, including both infectious and inflammatory diseases. The signature that we have identified distinguished KD from a wide range of other conditions with similar duration of fever and overlapping levels of inflammation. Because KD is diagnosed based on a constellation of clinical features and because there is no gold standard for diagnosis, evaluation of biomarker test results is difficult. Any cohort of children treated with IVIG for presumed KD is likely to include some patients with non-KD illness but with similar features. To evaluate the correspondence of the KD DRS with levels of diagnostic certainty, we categorized patients in the validation set as having definite, highly probable KD; KD-P, possible KD; and U, infections of uncertain bacterial or viral etiology.
Regarding the transcripts in the signature (Table 2), expression was lower in patients with KD compared with the non-KD group for 5 of the 13 transcripts. Of these, SIOOP, previously reported to have increased expression in acute KD relative to convalescence \( ^{35} \) or viral infections \( ^{32,35} \) was most abundant in patients with bacterial infection. The IFI27 gene has been reported to be upregulated in children with viral compared with bacterial infections \( ^{36} \) and autoimmune diseases \( ^{37,38} \) consistent with reduced expression of genes induced by type 1 interferons reported in acute KD vs adenovirus infection. \(^{32}\)

CD163 is a transmembrane receptor expressed in macrophages and monocytes involved in bacterial clearance during acute infection. \(^{39}\) A network analysis of the signature using pathway analysis (Ingenuity Pathways Analysis; Ingenuity Systems) revealed that 7 of the 13 transcripts in the signature were connected in a network around a central hub of tumor necrosis factor and interleukin 6 (Figure 4 in the Supplement).

**Strengths and Limitations**

We recognize both strengths and limitations in our study. First, the epidemiology of KD varies globally by ethnicity. Although we included patients with a mix of ethnicities in both discovery and validation cohorts, further studies are required to establish the performance of the signature in other geographic populations. Second, in the validation experiment, data from different Illumina microarray versions and studies were combined using the ComBat algorithm to achieve normalization. This normalization may reduce both experimental and biological sources of variability between data sets; consequently, the accuracy (AUC) of the signature in the validation set may be an underestimate. Conversely, although we showed that the ComBat algorithm successfully normalized the data sets, residual batch associations may have falsely increased the performance of the signature. Third, to develop a signature applicable in a wide range of febrile conditions, we discovered the signature through comparison of KD cases with a wide range of febrile controls with a spectrum of KD diagnosis likelihoods. This potentially biased the test toward better discriminatory value than might be applicable in the clinical setting. Fourth, we discovered the signature using patients with KD in the first 7 days of illness, with the aim of identifying a test for use early in the disease to enable treatment before coronary injury has occurred. Because the performance of the signature was lower in patients with KD seen after the seventh day, further work is required to establish the optimal signature for diagnosis in patients with KD with late presentation.

**Conclusions**

The results of our study suggest that KD can be distinguished from the range of infectious and inflammatory conditions with which it is often clinically confused using 13 transcripts in blood. Development of a test based on this gene expression signature is made more achievable because of the small number of transcripts in our signature and the rapidly evolving technologies for detecting nucleic acids. A diagnostic test would be a major advance allowing earlier treatment and thus prevention of cardiac complications of this serious childhood disease. Our findings represent a step toward better diagnosis of diseases based on molecular signatures rather than the clinical criteria and are relevant to many other clinical syndromes.
Diagnosis of Kawasaki Disease Using a Minimal Whole-Blood Gene Expression Signature

Original Investigation Research

approval of the manuscript; and decision to submit the manuscript for publication.

Group Information: Immunopathology of Respiratory, Inflammatory and Infectious Disease Study (IRIS) Consortium members: Imperial College London (United Kingdom): Michael J. Carter, PhD, Lachlan J. M. Cox, PhD, Harikila Elefterohorinou, PhD, Erin Fitzgerald, MSC, Stuart Gormley, MRes, Jethro A. Herberg, PhD, Clive J. Hoggart, PhD, Victoria A. Janes, MD, Kelsey D. Jones, MD, Mirai Kawasaki, MD, Stephanie Menikou, PhD, Sofia Mustafa, MSc, Stéphane Paulus, FRCPCH, Joanna Reid, MSC, Hannah Shales, PhD, Victoria J. Wright, PhD, and Michael Levin, FRCPCH; Southampton (United Kingdom): Saul Faust, PhD, Jenni McInerney, RN, and Sanjay Patel, FRCPCH; Oxford (United Kingdom): Andrew J. Pollard, PhD, Louise Willis, RNC, and Zoe Young, RNC. Micropathology Ltd (United Kingdom): Colin Fink, FRCPPath, and Ed Sumner, PhD; Emma Children’s Hospital, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands: Anouk M. Barendregt, BSc, D. Schoenberg, MD, J. Merijn van den Berg, PhD, and Taco W. Kuipers, PhD; and Hospital Clinico Universitario de Santiago (Spain): Miriam Ceby Lopez, PhD, Antonio Salas, PhD, Antonio Justicia Grande, MD, Irene Rivero, MD, Alberto Gómez Carballa, PhD, Jacobo Pardo Seco, PhD, José María Martíndez Sánchez, MD, Lorenzo Redondo Collazo, MD, Carmen Rodríguez-Tenreiro, PhD, Lucia Vilanova Martinón Sánchez, PhD, Lorenzo Redondo Collazo, MD, Carmen Rodríguez-Tenreiro, PhD, Lucia Vilanova Trillo, LPN, and Federico Martín Torr Reese. Pediatric Emergency Medicine Kawasaki Disease Research Group (PEMKDORG): Members Rody Children’s Hospital–San Diego, San Diego, California: John T. Kenagaye, MD, Lindsay T. Grubensky, PhD, Jim R. Hanley, MD, Paul Ishimine, MD, Jamie Lien, MD, Simon J. Lucio, MD, Seema Shah, MD, and Federico Martín-Torres, PhD. Additional Contributions: David Inwald, PhD (Imperial College Healthcare National Health Service Trust), critically appraised the manuscript (no compensation was received). We thank the patients and their families who participated in the study, as well as the clinical teams for their assistance in recruiting patients to the study.

Data Repository: The data discussed in this publication have been deposited in the National Center for Biotechnology Information’s Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE73464 (http://www.ncbi.nlm.nih.gov/geo/).

REFERENCES

1. Kawasaki T, Kosaki F, Okawa S, Shigematsu I, Du ZD, Zhao D, Du J, et al; Beijing Kawasaki Disease Study (Beads) Group Information: Immunopathology of Respiratory, Inflammatory and Infectious Disease Study (IRIS) Consortium members: Imperial College London (United Kingdom): Michael J. Carter, PhD, Lachlan J. M. Cox, PhD, Harikila Elefterohorinou, PhD, Erin Fitzgerald, MSC, Stuart Gormley, MRes, Jethro A. Herberg, PhD, Clive J. Hoggart, PhD, Victoria A. Janes, MD, Kelsey D. Jones, MD, Mirai Kawasaki, MD, Stephanie Menikou, PhD, Sofia Mustafa, MSc, Stéphane Paulus, FRCPCH, Joanna Reid, MSC, Hannah Shales, PhD, Victoria J. Wright, PhD, and Michael Levin, FRCPCH; Southampton (United Kingdom): Saul Faust, PhD, Jenni McInerney, RN, and Sanjay Patel, FRCPCH; Oxford (United Kingdom): Andrew J. Pollard, PhD, Louise Willis, RNC, and Zoe Young, RNC. Micropathology Ltd (United Kingdom): Colin Fink, FRCPPath, and Ed Sumner, PhD; Emma Children’s Hospital, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands: Anouk M. Barendregt, BSc, D. Schoenberg, MD, J. Merijn van den Berg, PhD, and Taco W. Kuipers, PhD; and Hospital Clinico Universitario de Santiago (Spain): Miriam Ceby Lopez, PhD, Antonio Salas, PhD, Antonio Justicia Grande, MD, Irene Rivero, MD, Alberto Gómez Carballa, PhD, Jacobo Pardo Seco, PhD, José María Martíndez Sánchez, MD, Lorenzo Redondo Collazo, MD, Carmen Rodríguez-Tenreiro, PhD, Lucia Vilanova Martinón Sánchez, PhD, Lorenzo Redondo Collazo, MD, Carmen Rodríguez-Tenreiro, PhD, Lucia Vilanova Trillo, LPN, and Federico Martín-Torres, PhD. Pediatric Emergency Medicine Kawasaki Disease Research Group (PEMKDORG): Members Rody Children’s Hospital–San Diego, San Diego, California: John T. Kenagaye, MD, Lindsay T. Grubensky, PhD, Jim R. Hanley, MD, Paul Ishimine, MD, Jamie Lien, MD, Simon J. Lucio, MD, Seema Shah, MD, and Federico Martín-Torres, PhD. Additional Contributions: David Inwald, PhD (Imperial College Healthcare National Health Service Trust), critically appraised the manuscript (no compensation was received). We thank the patients and their families who participated in the study, as well as the clinical teams for their assistance in recruiting patients to the study.

Data Repository: The data discussed in this publication have been deposited in the National Center for Biotechnology Information’s Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE73464 (http://www.ncbi.nlm.nih.gov/geo/).

REFERENCES

1. Kawasaki T, Kosaki F, Okawa S, Shigematsu I, Du ZD, Zhao D, Du J, et al; Beijing Kawasaki Disease Study (Beads) Group Information: Immunopathology of Respiratory, Inflammatory and Infectious Disease Study (IRIS) Consortium members: Imperial College London (United Kingdom): Michael J. Carter, PhD, Lachlan J. M. Cox, PhD, Harikila Elefterohorinou, PhD, Erin Fitzgerald, MSC, Stuart Gormley, MRes, Jethro A. Herberg, PhD, Clive J. Hoggart, PhD, Victoria A. Janes, MD, Kelsey D. Jones, MD, Mirai Kawasaki, MD, Stephanie Menikou, PhD, Sofia Mustafa, MSc, Stéphane Paulus, FRCPCH, Joanna Reid, MSC, Hannah Shales, PhD, Victoria J. Wright, PhD, and Michael Levin, FRCPCH; Southampton (United Kingdom): Saul Faust, PhD, Jenni McInerney, RN, and Sanjay Patel, FRCPCH; Oxford (United Kingdom): Andrew J. Pollard, PhD, Louise Willis, RNC, and Zoe Young, RNC. Micropathology Ltd (United Kingdom): Colin Fink, FRCPPath, and Ed Sumner, PhD; Emma Children’s Hospital, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands: Anouk M. Barendregt, BSc, D. Schoenberg, MD, J. Merijn van den Berg, PhD, and Taco W. Kuipers, PhD; and Hospital Clinico Universitario de Santiago (Spain): Miriam Ceby Lopez, PhD, Antonio Salas, PhD, Antonio Justicia Grande, MD, Irene Rivero, MD, Alberto Gómez Carballa, PhD, Jacobo Pardo Seco, PhD, José María Martíndez Sánchez, MD, Lorenzo Redondo Collazo, MD, Carmen Rodríguez-Tenreiro, PhD, Lucia Vilanova Trillo, LPN, and Federico Martín-Torres, PhD. Pediatric Emergency Medicine Kawasaki Disease Research Group (PEMKDORG): Members Rody Children’s Hospital–San Diego, San Diego, California: John T. Kenagaye, MD, Lindsay T. Grubensky, PhD, Jim R. Hanley, MD, Paul Ishimine, MD, Jamie Lien, MD, Simon J. Lucio, MD, Seema Shah, MD, and Federico Martín-Torres, PhD. Additional Contributions: David Inwald, PhD (Imperial College Healthcare National Health Service Trust), critically appraised the manuscript (no compensation was received). We thank the patients and their families who participated in the study, as well as the clinical teams for their assistance in recruiting patients to the study.

Data Repository: The data discussed in this publication have been deposited in the National Center for Biotechnology Information’s Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE73464 (http://www.ncbi.nlm.nih.gov/geo/).


34. Newburger JW, Takahashi M, Gerber MA, et al; Committee on Rheumatic Fever, Endocarditis and Kawasaki Disease; Council on Cardiovascular Disease in the Young; American Heart Association; American Academy of Pediatrics. Diagnosis, treatment, and long-term management of Kawasaki disease: a statement for health professionals from the Committee on Rheumatic Fever, Endocarditis and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association. Circulation. 2004;110(17):2747-2771. doi:10.1161/01.CIR.0000145143.19711.78


