Rheumatology 2009;48:899–905
Advance Access publication 28 May 2009
doi:10.1093/rheumatology/kep125

# Differential cytokine profiles in juvenile idiopathic arthritis subtypes revealed by cluster analysis

**Henk-Jan van den Ham[1], Wilco de Jager[2], Johannes W. J. Bijlsma[3], Berent J. Prakken[2] and Rob J. de Boer[1]**

**Objectives.** With the introduction of high-throughput biomarker measurements, traditional analysis of these markers is increasingly difficult. Using samples from a diverse group of patients, we tested the applicability of cluster analysis to these data. Using this method, we aim to visualize some of the patterns specific to certain disease groups. In particular, we focus on juvenile idiopathic arthritis (JIA), a multifactorial autoimmune disorder that ultimately leads to chronic inflammation of the joints.
**Methods.** Cytokine measurements were performed using multiplex immunoassays. Using heuristic clustering methods, we set out to compare the pattern of 30 cytokines in plasma and SF of JIA, RA, OA, or diabetes type II patients and healthy controls.
**Results.** Analysis shows that oligo- and polyarticular JIA have similar biomarker profiles, both in plasma and SF. Systemic onset JIA (SoJIA) has a profile distinct from other JIA subtypes, suggesting that they involve different inflammatory processes. SoJIA samples do, however, cluster together with RA in SF, suggesting that these two conditions have similar cytokine profiles. Furthermore, we identify several clusters of ILs and chemokines that are co-expressed, suggesting that they are co-regulated.
**Conclusions.** We show that previously undetected clusters of cytokines and patients can be identified by applying cluster analysis to multiplex data. Cytokine clusters identified in plasma and SF samples were quite different, which underscore the differential cytokine signalling in these two compartments, and suggest that plasma samples may not be suitable for estimating joint biomarker profiles and inflammation.

**Key words:** Juvenile idiopathic arthritis, Cluster analysis, Inflammation, Cytokines, Biomarkers.

## Introduction

Juvenile idiopathic arthritis (JIA) is an autoimmune disorder that is thought to be caused by a combination of environmental and genetic factors [1]. All subtypes of JIA are characterized by chronic inflammation of the joints [2]. The local inflammatory reaction is controlled by infiltrating immune cells in combination with different cytokines [3, 4]. The multifactorial pathogenesis of JIA is reflected in the many cytokines that have been reported to play a role in either disease pathology or in recovery from disease [4–9]. These regulating cytokines may serve as biomarkers for diagnosis or treatment fine-tuning, and a more complete knowledge of the cytokine signalling network in the local inflammatory process is fundamental to finding more effective clinical intervention strategies [10].

Technological advances in the past few years have led to the availability of multiplex immunoassays that can now measure many cytokines simultaneously in a 50-μl sample [11]. While providing a wealth of information about the bio-markers related to certain disorders, these data are known to be noisy, and analysis can be complex due to the large number of cytokines. These high-throughput techniques therefore call for new methods of analysis to identify patterns invisible to normal statistics.

We aim to visualize some of the disease-related patterns by applying heuristic cluster analysis methods to multiplex cytokine measurements. The multidimensional character of multiplex data (i.e. many cytokines concurrently measured in many samples) allows the application of cluster analysis as an alternative to normal statistical analysis [12], and a diverse panel of samples should enable the cluster algorithm to clearly differentiate between different sample and disease types. Here, we show that cluster analysis can improve the analysis substantially, and is therefore a useful tool for the analysis of multiplex data.

## Methods

### *Participants*

All children in this cross-sectional study were classified according to the ILAR criteria [13], and were divided into three major subgroups: (persistent) oligoarticular JIA (oJIA), polyarticular JIA (pJIA) (including five extended oJIA) and systemic onset JIA (SoJIA). RA patients were diagnosed according the 1987 ACR criteria [14]. The oJIA, pJIA, SoJIA and diabetes mellitus (DM) data were analysed and published before [12]; the healthy control (HC), RA and OA data are novel. Patient characteristics are outlined in Table 1.

Blood samples were obtained at either the outpatient clinics of the University Medical Centre Utrecht, Utrecht, The Netherlands, or at Great Ormond Street Hospital for Children, London, UK. The study had full ethical approval at both centres. Furthermore, SF samples were obtained at the time of a therapeutic joint aspiration from both RA and JIA. Informed consent was obtained either from parents or from the individuals directly if they were >12 years of age. Blood and SF samples were collected and were centrifuged to remove cells, and were stored frozen at −80°C until analysis.

### *Multiplex immunoassay*

All 30 antibody pairs and recombinant proteins used were purchased from commercial sources as described previously [15]. Aspecific heterophilic immunoglobulins such as RF were pre-absorbed from all samples with protein-L (Pierce, Rockford, IL, USA) as described earlier [15]. Measurements and data analysis of all assays were performed using the Bio-Plex system in combination with the Bio-Plex Manager software V.4.0 using five parametric curve fitting (Bio-Rad Laboratories, Hercules, CA, USA).

[1]Theoretical Biology and Bioinformatics, Utrecht University, [2]Department of Paediatric Immunology and [3]Department of Rheumatology and Clinical Immunology, University Medical Centre Utrecht, Utrecht, The Netherlands.

Submitted 6 October 2008; revised version accepted 20 April 2009.

Correspondence to: Henk-Jan van den Ham, Theoretical Biology and Bioinformatics, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands. E-mail: h.j.vandenham@uu.nl



© The Author 2009. Published by Oxford University Press on behalf of the British Society for Rheumatology. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

TABLE 1. Clinical and sample characteristics

| | oJIA | pJIA | SoJIA | RA | OA | DM | HC |
|---|---|---|---|---|---|---|---|
| Sex, female/male | 21/9 | 13/7 | 8/7 | 3/6 | 4/1 | 5/4 | 11/9 |
| Age, mean ± S.D., years | 9.3 ± 4.0 | 9.2 ± 3.3 | 9.7 ± 4.1 | 58.4 ± 20.3 | 56.7 ± 38.1 | 12.8 | 9.1 ± 5.3 |
| Disease duration, mean ± S.D., years | 5.9 ± 3.3 | 5.7 ± 3.4 | 4.4 ± 3.1 | 9.3 ± 13 | 9.1 ± 12 | NA | – |
| Active/Remission | 20/10 | 13/7 | 7/8 | 9/0 | – | – | – |
| ESR, mean ± S.D., mm/h | 21 ± 17 | 18 ± 15 | 46 ± 58 | 34 ± 23 | ND | ND | ND |
| CRP, mean ± S.D., mg/l | 29 ± 26 | 33 ± 34 | 45 ± 43 | 36 ± 28 | ND | ND | ND |
| RF positive[a], *n* (%) | 7 (23) | 3 (15) | 0 (0) | 6 | 0 | ND | ND |
| NSAID, *n* (%) | 24 (80) | 19 (95) | 14 (93) | 9 (100) | 3 | NA | 0 (0) |
| MTX, *n* (%) | 12 (40) | 11 (55) | 8 (53) | 7 (78) | 0 (0) | NA | 0 (0) |
| Corticosteroids, *n* (%) | 1 (3) | 2 (10) | 4 (27) | 0 (0) | 0 (0) | NA | 0 (0) |
| Anti-TNF-a, *n* (%) | 0 (0) | 3 (15) | 2 (13) | 4 (44) | 0 (0) | NA | 0 (0) |
| No. of samples | | | | | | | |
| Plasma | 30 | 20 | 15 | 9 | 0 | 9 | 20 |
| SF | 19 | 9 | 5 | 10 | 5 | 0 | 0 |

ND: not determined; NA: not available. [a]Determined by IgM ELISA.

### Cluster analysis

Cluster analysis was performed both for patients and cytokines. The measured cytokine concentrations were log-transformed after imposing a detection limit of 0.1 pg/ml. Approximately one-quarter of these cytokine measurements was below detection limit (1218 out of 4530 measurements), and most of these were in plasma samples (965 out of 3090 measurements). The distance measure used to quantify the distance between two data points was $(1-\rho)$, where $\rho$ is the pair-wise complete Pearson's correlation between two cytokines or patients. Hierarchical clustering using Ward's (minimum variance) method was subsequently used to cluster the data. For cytokine clustering, measurement values below detection limit were set to detection limit, and were in this way included in the clustering. Incorporating below detection limit data in the patient clustering resulted in clusters of patients having the same missing cytokines (data not shown). CCL5 was excluded from clustering the plasma samples, since it is known that platelet release of CCL5 may occur *in vitro* during processing, which masks its true plasma concentration [16]. By combining cytokine and sample clustering, heat maps were created with colours corresponding to the relative expression of the cytokine.

To assess the significance of disease type distribution over different clusters, a hypergeometric test was used. Every patient group within a tree branch or cluster was tested for significant presence in or absence from this cluster. In total, 24 tests and 16 tests were performed for patient clusters in plasma and SF, respectively. To account for multiple testing, we applied a Bonferroni correction, and considered a *P*-value of 0.002 (plasma) or 0.003 (SF) to be significant. To assess the robustness of the cytokine trees, bootstrapping ($n = 100\,000$) was applied to both the plasma and SF trees. All statistical analyses were performed using R, a free software environment for statistical computing and graphics (http://www.r-project.org/). Bootstrapping was performed using the Pvclust package [17].

## Results

Cytokine levels in clinical samples obtained from a diverse group of patients were measured using multiplex immunoassays. Patient groups include (persistent) oJIA, pJIA (including five extended oJIA cases), SoJIA, RA, DM, OA as well as HCs. Both plasma and SF cytokines were measured where available (Table 1). Summaries of the data are shown in Supplementary Tables 1 (plasma) and 2 (SF) (supplementary data are available at *Rheumatology* Online).

To test the applicability of cluster analysis to cytokine profiles, we first performed an analysis to see if this method could differentiate between SF and plasma samples. Using correlation clustering and Ward's linkage method, a cluster analysis was performed using all available samples. No label information or previous data selection method was used to stratify the data. Cluster analysis clearly classifies plasma and SF samples into two separate clusters, with only five SF samples and one plasma sample being misclassified (Fig. 1). This shows that cluster analysis is able to pick up significant patterns in cytokine profiles.

To observe possible different cytokine profiles in patients with different disorders, we performed cluster analysis on samples and cytokines of plasma origin. This analysis applies clustering in two ways: it (i) clusters patients based on cytokine expression profiles, and (ii) clusters cytokines based on patient profiles. The result is represented as a heat map, where red indicates high, and cyan indicates low relative expression, respectively. In general, the chequered pattern of colours indicates that cytokine measurements are noisy (as expected), and that IL measurements are frequently below detection limit (indicated by grey). Nonetheless, the heat map shows areas of relatively high (hot spots) and low (cold spots) cytokine expression in certain clusters of patients. The patient clustering shows a clear, but incomplete, separation of disease types into clusters. The patient tree was divided into six clusters (Fig. 2), and hypergeometric tests were used to assess the significance of the disease type distribution over the clusters (Table 2). In total, 24 hypergeometric tests were performed for the plasma sample clustering. Due to the small number of samples and the large number of tests, the absence of a patient group from a cluster was often not significant. However, the presence of a patient group within a cluster was generally significant. Using the obtained *P*-values, each cluster was annotated using the dominant disease type in the cluster: cluster 1: SoJIA; cluster 2: combined oJIA and pJIA (oJIA/pJIA); cluster 3: all RA and DM samples (RA/DM); cluster 4: an oJIA cluster; clusters 5 and 6: HC (Fig. 2). Although several patients are not classified correctly, this shows that cluster analysis is able to identify clinically recognized disease groups, and can classify patients into these disease groups based on their plasma cytokine profiles.

Cytokine clustering of plasma samples yields a tree that relates cytokines to each other based on their expression in the samples. ILs and chemokines appear largely segregated on opposite sides of the tree, with a few exceptions (CXCL8, IL-4 and IL-17, Fig. 2). The heat map shows a hot spot of CXCL10, CXCL9, CCL22, CCL17, OPG and CCL2 in cluster 4 (oJIA), and part of cluster 5 (HC). Conversely, there are two cold spots for these cytokines in cluster 1 (SoJIA), and part of cluster 6 (HC) (Fig. 2). This patient subgroup of oJIA and pJIA appears to have a different cytokine expression profile compared with other oJIA and pJIA patients (primarily in clusters 2 and 3) and SoJIA (cluster 1). The known heterogeneity within this disorder is therefore evident from the cytokine expression profiles.

Next, we performed two-way cluster analysis on the SF samples, which represent more directly the local site of inflammation.

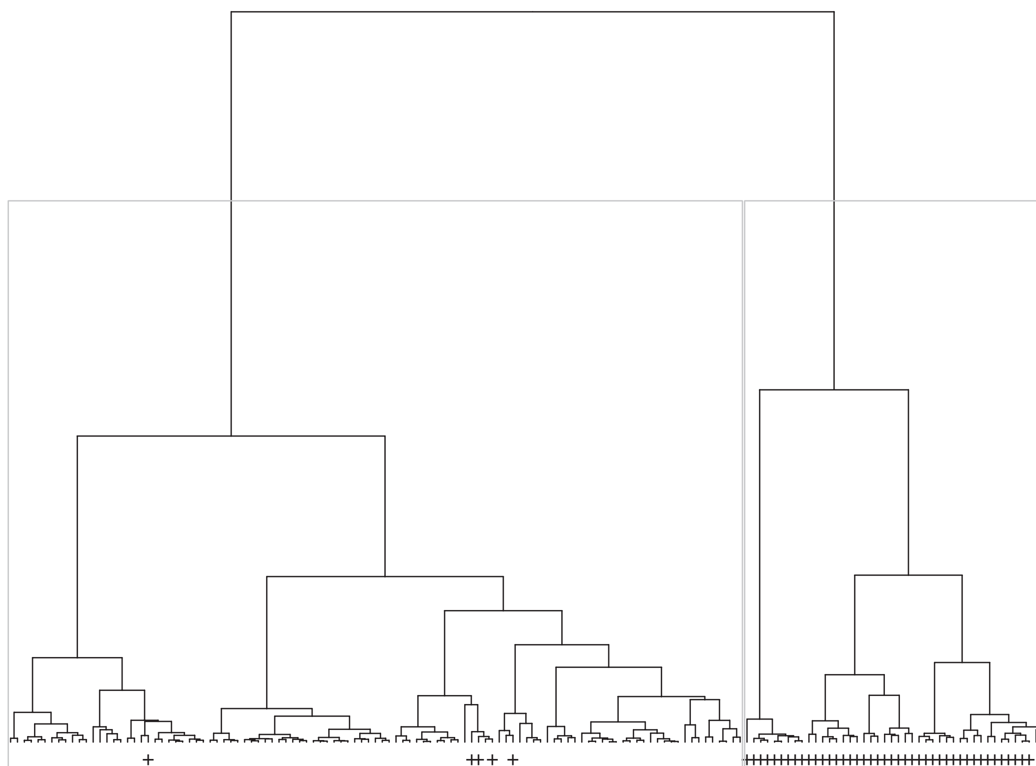

FIG. 1. Cluster analysis classifies plasma and SF samples. SF samples marked (+), grey boxes indicate the two clusters that were identified. The left cluster contains 107 samples, and the right cluster contains 44 samples. CCL5 measurements have been excluded from this analysis.

The patient tree was divided into four clusters (Fig. 3). As in the plasma clustering, disease groups are asymmetrically divided, and separation is not complete with some clusters containing two or three disease types. In total, 16 hypergeometric tests were carried out to assess the significance of the disease group present within or absence from a cluster (Table 3). As before, each cluster was annotated by its significantly present disease group. Cluster 4 consists predominantly of RA and SoJIA samples, which cluster together in SF (Fig. 3), but not in plasma (Fig. 2). They are characterized by a hot spot of the IL cluster TNF-$\alpha$/IL-2. The RA and SoJIA samples intermingle, showing that the cytokine profile of these two different disorders is similar. Cluster 3 is an OA cluster, which can be characterized as having an overall low expression of all cytokines. Clusters 1 and 2 are oJIA/pJIA clusters that have two different profiles (Fig. 3). Each of these clusters appears to have their own specific cytokine hot spot, with cluster 2 (oJIA/pJIA) having high CXCL10, CCL3, IL-17 and CXCL9, and cluster 1 (oJIA/pJIA) having high CCL2, CCL11, sCD54, sCD106, sRANKL and OPG. This shows that cluster analysis classifies patients into distinct clinical groups based on cytokine profiles measured in SF.

When we investigated the difference between these two groups of oJIA and pJIA patients, we unexpectedly found that these samples cluster geographically. The SF samples obtained in UK nearly all cluster together in cluster 2 (oJIA/pJIA) without any other samples in this cluster [Fig. 3; note the UK sample in cluster 4 (RA/SoJIA) is from a SoJIA patient]. Dutch samples are primarily in cluster 1 (oJIA/pJIA) and some are in cluster 4 (RA/SoJIA) (Fig. 3). In plasma, UK samples also cluster together, although to a lesser degree than in SF (data not shown). Further analysis of the ESR and CRP values in the oJIA and pJIA patients showed that the British patients had higher levels than Dutch subjects [ESR 68 (46–171) and 28 (7–80), $P = 0.0317$; CRP 63.1 (18–110) and 29.5 (3–110), respectively; $P = 0.051$], which suggests more disease activity in the UK cohort at the time of sample collection. However, we cannot rule out the possibility that sample collection, storage, transport or processing may have contributed to the clustering of UK samples.

Multiplex data have so far been largely used to characterize patient groups and phenotypes, but the data in themselves have not been used for studying the regulation of cytokines themselves. Excluding CCL5, bootstrapped cytokine trees were generated for plasma and SF samples separately (Fig. 4). From these trees, one can identify cytokines that are co-expressed, and could therefore have a regulatory link between them. The bootstrap value (bp) indicates the percentage of times that this cluster was identified by cluster analysis performed on re-sampled data. The bootstrap values for the two trees are generally low (on average 44 and 47% for plasma and SF, respectively); in particular the nodes high up in the trees (Fig. 4, nodes 24–27). There are, however, small clusters of cytokines that seem robust, such as an inflammatory cluster including IFN-$\gamma$, IL-1$\beta$, -6, -13 and TNF-$\alpha$ (bp = 82%) in the plasma tree. Other well-supported cytokine clusters in the plasma samples include IL-1$\alpha$ and -2 (bp = 87%), OPG, CCL2 and CCL11 (bp = 96%), a CXCL9 and CXCL10 cluster (bp = 100%) and a CCL18, sCD54 and sCD106 cluster (bp = 84%, Fig. 4a). Robust clusters in SF include the chemokine clusters OSM and CCL22 (bp = 97%), a CCL3 and CXCL10 cluster (bp = 96%) and a somewhat larger cluster of sCD54, CCL2, CCL11, sCD106, OPG and sRANKL (bp = 56%, Fig. 4b). This last cluster seems to have its equivalent in plasma samples, but more cytokines appear to be involved. The inflammatory cytokine cluster found in plasma appears in altered form in SF, but its bootstrap values are low, and many other cytokines appear in the cluster. Using several other clustering methods (complete linkage, UPGMA), similar clusters were identified, demonstrating that these clusters are reasonably robust with respect to the clustering technique applied. We furthermore tried to construct cytokine trees for particular disease groups, but since this requires the application of cluster analysis on even smaller subsets of the data, we found that the bootstrap values for these trees were too low to be considered significant. In summary,

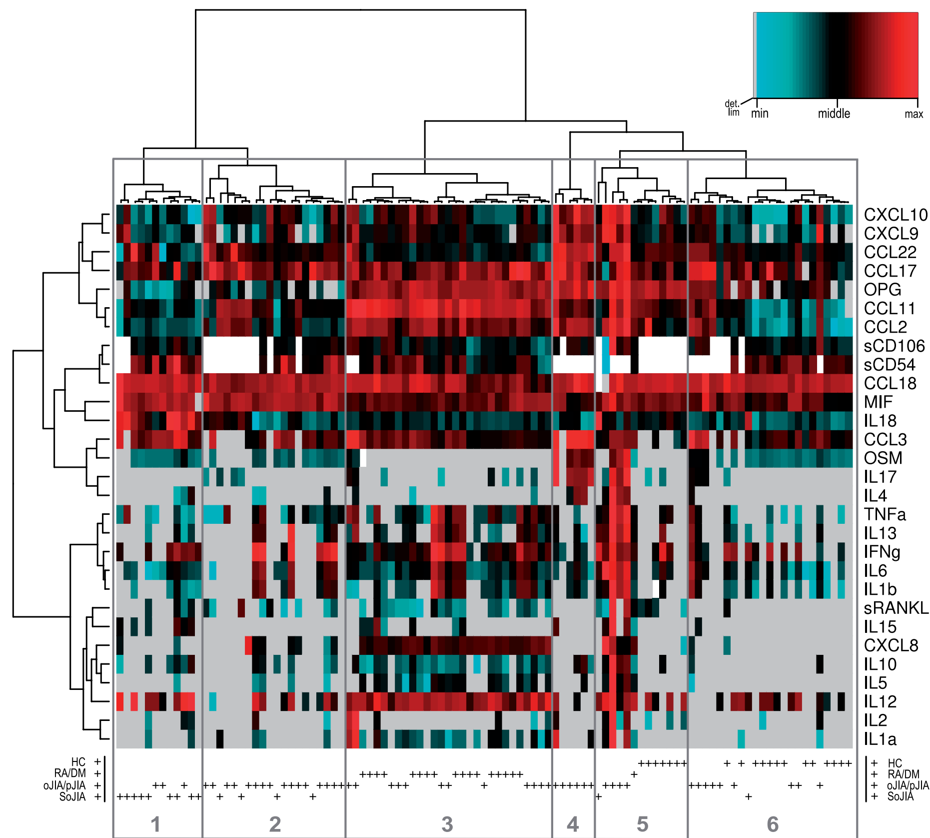


Fig. 2. Two-way cluster analysis of plasma samples. Patients were clustered (tree shown at the top) and divided into six clusters (grey boxes, clusters are numbered 1–6). The patient category HC, RA/DM, oJIA and pJIA or SoJIA is indicated at the bottom for each sample. Cytokines were clustered, with the tree and cytokine names shown on the left and right of the heat map, respectively. The relative abundance of a disease category within each of the clusters was tested for significance by hypergeometric test (Table 1). The heat map colours represent concentration of cytokines relative to the minimum and maximum of all values in this analysis; grey and white represent detection limited values and missing values, respectively.

Table 2. Patient and disease type distribution for plasma sample clustering

| | Plasma cluster (cluster size) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 (12) | | | 2 (20) | | | 3 (29) | | |
| Patient group | *n* | *d* | *P* | *n* | *d* | *P* | *n* | *d* | *P* |
| HC (20) | 0 | 0 | 0.063 | 0 | 0 | 0.0079 | 0 | 0 | 0.00057 |
| RA/DM (9/9) | 0/0 | 0 | 0.086 | 0/0 | 0/0 | 0.014 | 9/8[a] | 0.59 | $<10^{-10}$ |
| oJIA/pJIA (30/20) | 1/2 | 0.25 | 0.057 | 8/8[a] | 0.8 | 0.0014 | 8/4 | 0.41 | 0.12 |
| SoJIA (15) | 9[a] | 0.75 | $<10^{-6}$ | 4 | 0.2 | 0.19 | 0 | 0 | 0.0045 |

| | 4 (6) | | | 5(13) | | | 6 (23) | | |
|---|---|---|---|---|---|---|---|---|---|
| Patient group | *n* | *d* | *P* | *n* | *d* | *P* | *n* | *d* | *P* |
| HC (20) | 0 | 0 | 0.26 | 7[a] | 0.54 | $<10^{-8}$ | 13[a] | 0.56 | $10^{-5}$ |
| RA/DM (9/9) | 0/0 | 0 | 0.31 | 0/1 | 0.08 | 0.18 | 0 | 0 | 0.006 |
| oJIA/pJIA (30/20) | 6/0 | 1 | 0.011 | 3/1 | 0.31 | 0.10 | 4/5 | 0.39 | 0.11 |
| SoJIA (15) | 0 | 0 | 0.37 | 1 | 0.08 | 0.29 | 1 | 0.04 | 0.08 |

Significance of patient and disease type distribution for plasma sample clustering calculated by a hypergeometric test. Cluster size and patient group size within the analysis are shown in brackets. The number of a certain disease group within a cluster (*n*), the density of this group within the cluster (*d*) and the *P*-value (*P*) are shown. [a]Significant presence ($P < 0.002$ after Bonferroni correction) of a certain disease category within a cluster.

we find reasonably robust clusters of cytokines, but these are not the same in plasma and SF. This shows that cytokine expression in these two compartments is regulated differently.

## Discussion

The multiplex immunoassay is a novel tool that allows multiple biomarkers to be measured concurrently in a single sample; without this technique, measuring 30 different markers in inherently small paediatric samples would hardly be feasible. The increasing dimensions of the data make the analysis of multiplex data increasingly complex, and calls for a more integrated approach. Using heuristic clustering methods, we were able to segregate samples into sample type (plasma/SF) and disease background with high accuracy and significance. oJIA/pJIA and SoJIA are classified into separate clusters, and clearly have

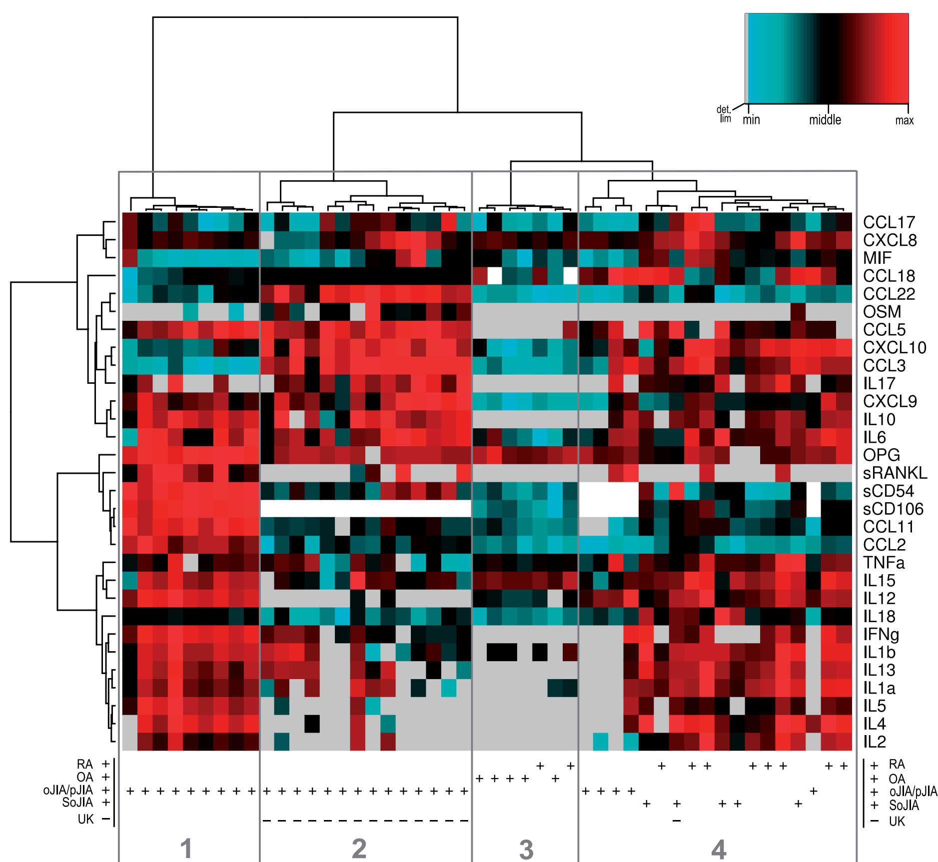


FIG. 3. Two-way cluster analysis of SF samples. Patients were clustered (tree shown at the top) and divided into four clusters (grey boxes, clusters are numbered 1–4). The patient categories RA, OA, oJIA and pJIA or SoJIA are indicated at the bottom for each sample. Cytokines were clustered, and the tree and cytokine names shown on the left and right side of the heat map, respectively. The relative abundance of a disease category within each of the clusters was tested for significance by a hypergeometric test (Table 2). The heat map colours represent concentration of cytokines relative to the minimum and maximum of all values in this analysis; grey and white represent detection limited values and missing values, respectively.

TABLE 3. Patient and disease type distribution for SF sample clustering

| | SF cluster (cluster size) | | | | | |
|---|---|---|---|---|---|---|
| | 1 (9) | | | 2 (14) | | |
| Patient group | *n* | *d* | *P* | *n* | *d* | *P* |
| RA (10) | 0 | 0 | 0.14 | 0 | 0 | 0.046 |
| OA (5) | 0 | 0 | 0.63 | 0 | 0 | 0.47 |
| oJIA/pJIA (19/9) | 6/3[a] | 1 | $<10^{-5}$ | 10/4[a] | 1 | $<10^{-9}$ |
| SoJIA (5) | 0 | 0 | 0.63 | 0 | 0 | 0.47 |
| | 4 (7) | | | 5(18) | | |
| Patient group | *n* | *d* | *P* | *n* | *d* | *P* |
| RA (10) | 2 | 0.28 | 0.27 | 8[a] | 0.44 | 0.038 |
| OA (5) | 5[a] | 0.71 | $<10^{-6}$ | 0 | 0 | 0.37 |
| oJIA/pJIA (19/9) | 0/0 | 0 | 0 | 3/2 | 0.27 | 0.23 |
| SoJIA (5) | 0 | 0 | 0 | 5[a] | 0.27 | $10^{-4}$ |

Significance of patient and disease type distribution for SF clustering by hypergeometric test. The cluster size and patient group size within the analysis are shown in brackets. The number of a certain disease group within a cluster (*n*), the density of this group within the cluster (*d*) and the *P*-value (*P*) are shown. [a]Significant presence ($P < 0.003$ after Bonferroni correction) of a certain disease category within a cluster.

different cytokine profiles. In SF, SoJIA cluster together with RA samples. Combined, these observations suggest that the ongoing joint inflammation in SoJIA is more similar to RA than to other JIA subtypes. There are, however, always some oJIA and pJIA samples in the SoJIA cluster, both for plasma and SF. The cytokine profile of these misclassified patients is apparently more similar to SoJIA than to other oJIA/pJIA patients. Furthermore, RA and DM plasma samples associate in a single cluster (cluster 3, Fig. 2). Additional analysis of the RA/DM cluster shows, however, that RA and DM samples largely segregate into disease-specific sub-clusters (data not shown). This implies that, even though RA and DM samples cluster together, these diseases do not have identical biomarker profiles. RA and DM biomarker profiles are evidently more similar to each other than to other patients groups or HCs.

The lack of separation between oJIA and pJIA patients indicates that these JIA subtypes have similar biomarker expression, suggesting that the inflammatory process is similar. There do,

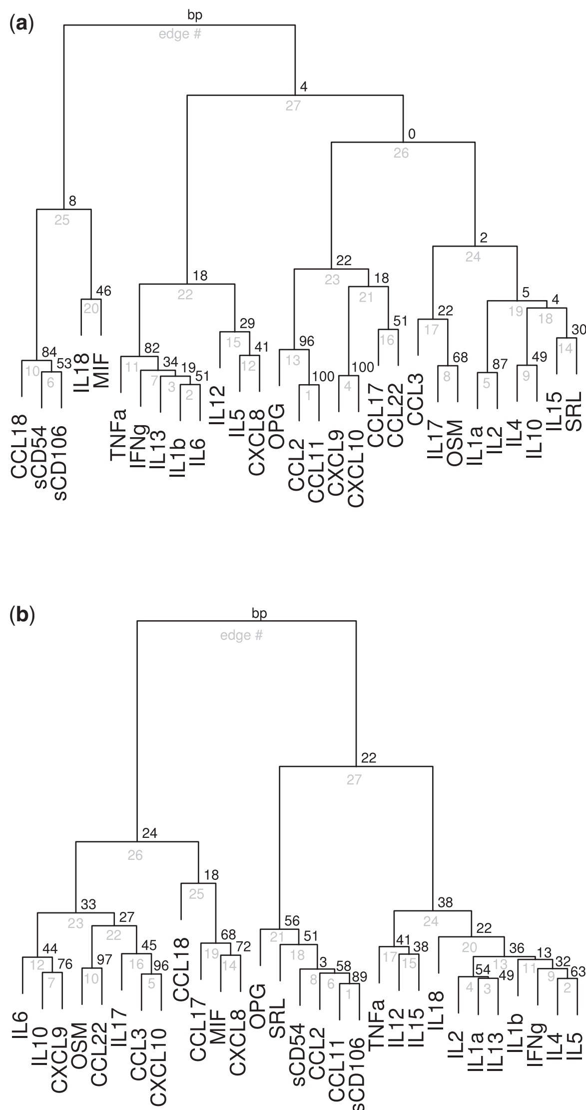


FIG. 4. Co-expression of cytokines quantified by cluster analysis. Two separate trees were constructed for plasma (**a**) and SF (**b**) samples. Bootstrapping was performed ($n = 100\,000$) to affirm stability of the observed clustering, and are shown as a percentage by black numbers at the tree nodes; the nodes are labelled with grey numbers.

however, appear to be large differences within this oJIA/pJIA subset, because these patients appear in almost every cluster, both in plasma and SF. Similar studies with well-defined populations such as early onset RA, revealed specific, but transient, up-regulated cytokine clusters that could be used for disease classification [18]. Repeating this analysis on such data before starting therapy in any form, combined with follow-up at frequent intervals, may help to refine classification protocols.

In addition to its ability to handle high-dimensional data, an additional advantage of cluster analysis is that groups emerge from the analysis, and do not have to be pre-defined. In this way, we identified, to our surprise, differential cytokine expression in patients from the UK or The Netherlands. The likelihood of a spurious clustering due to differential treatment of either of the sample groups (e.g. transport or storage) was low given the similar logistics in both centres, but cannot be ruled out. Furthermore, not all British samples were in the same cluster (one British SoJIA is in a different cluster, Fig. 4), and additional analysis shows that patients from the UK, though also diagnosed with oJIA and pJIA, have increased CRP and ESR levels. This suggests that the patients in the British cohort have more disease activity, perhaps reflecting differential patient referral or selection criteria in the UK and The Netherlands. More SF and plasma samples from different patient groups, in particular HCs from the UK, are needed to clarify this point.

Considering that cluster analysis identifies groups, it is interesting to note that we were unable to find significant patterns related to the disease status and treatment regimen of the patients. Previous analysis of these data suggests that there are clear differences between active and remissive disease, and between different treatment strategies. This study does not reproduce these differences, perhaps because cluster analysis divides groups along the main axis of variation, i.e. the factor that causes the largest differences between groups. This suggests that disease phenotype has a stronger influence on the cytokine profiles in plasma and SF than treatment or disease status has. Redoing this analysis with more patients with the same clinical disease, but with different treatment regimens could potentially identify treatment subgroups within disease phenotypes.

The multiplex measurements can be used not only for subsetting patients as discussed above, but also for clustering the expression profiles of the cytokines themselves. This approach should provide insights into which cytokines are co-expressed, and could therefore be co-regulated. At the site of inflammation, two clusters of cytokines can be identified (IL-6, -10 and CXCL9; CCL3, IL-17 and CXCL10) related to T-cell function, T-cell proliferation and recruitment of inflammatory T cells (Th1) [19–21]. Cytokine receptor CXCR3, which binds CXCL9 and CXCL10, and CCR4/CCR5, which both bind CCL3, is known to be up-regulated on several cell populations in the inflamed joint, and are critical for lymphocyte homing to the site of inflammation in both JIA [22] and RA [23]. In plasma, several ligands for both CXCR3 (CXCL9 and CXCL10) and CCR5 (CCL2 and CCL11) cluster together in combination with IL-10 and -17, which are both known to have a regulatory function in arthritis [5, 24, 25]. Additionally, the inclusion of the classical Th2 cytokine IL-13 in the plasma cluster of pro-inflammatory cytokines hints at a different role for IL-13 in a complex inflammatory environment in arthritis.

While the patterns found in this study commend the use of cluster analysis as a prognostic tool, the results of this study should be treated with caution. Cytokine measurements are known to be noisy, perhaps due to the various polymorphisms that have been described for several cytokines such as IL-6, -10, TNF-$\alpha$ and macrophage migration inhibitory factor (MIF), and which have been linked to arthritic diseases [5, 26–28]. Secondly, the amount of available data is limited; a dataset of 151 samples may seem quite large from a clinical point of view, but from the analytical point of view it is quite small. Thirdly, our analysis was complicated by the fact that many cytokine concentrations were below detection limit. This forced us to treat data below the detection limit (low expression) as missing values. Other normalization and clustering techniques were successful to a degree of identifying small homogeneous disease sub-clusters, but did not yield clusters as presented here. We expect that redoing this analysis with more samples will probably increase the robustness considerably.

Collectively, the findings of this exploratory study illustrate the power of cluster analysis concurrently with the heterogeneity between, and within, the JIA disease groups. Further steps in applying cluster analysis to multiplex data include generating larger datasets with more cytokines to improve accuracy and reliability. We believe that this novel analysis of plasma and SF cytokine patterns shows that combining a multiplex immunoassay with cluster analysis promises to be an essential tool that could greatly enhance our understanding on the role of cytokines in pathogenesis, and may contribute to fine-tuning therapeutic interventions in the future.

## Rheumatology key messages

- Multiplex immunoassays are a useful tool to generate biomarker profiles of individual patients.
- Cluster analysis is a useful tool for analysing biomarker profiles.
- Cluster analysis visualizes previously identified as well as unexpected groups in both cytokines and patients.


## Acknowledgements

We thank Dr Nico Wulffraat (UMC Utrecht, Utrecht, The Netherlands) and Dr Lucy Wedderburn (University College London, London, UK) for assembling patient material that was essential for these studies.

H.J.v.d.H. and R.J.d.B. were financially supported by the VIRGO consortium, an Innovative Cluster approved by the Netherlands Genomics Initiative and partially funded by the Dutch Government (BSIK 03012), and by the Netherlands Organization for Scientific Research (NWO) VICI grant 016.048.603. W.d.J. and B.J.P. were supported by the Dutch Arthritis Foundation (Reumafonds). Furthermore B.J.P. is supported by a VIDI innovative grant from Netherlands Organization for Scientific Research (NWO).

*Disclosure statement*: The authors have declared no conflicts of interest.


## Supplementary data

Supplementary data are available at *Rheumatology* Online.

## References


1 Prahalad S. Genetics of juvenile idiopathic arthritis: an update. Curr Opin Rheumatol 2004;16:588–94.
2 Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369:767–78.
3 Choy EH, Panayi GS. Cytokine pathways and joint inflammation in rheumatoid arthritis. N Engl J Med 2001;344:907–16.
4 Woo P. Cytokines and juvenile idiopathic arthritis. Curr Rheumatol Rep 2002;4: 452–57.
5 Crawley E, Kon S, Woo P. Hereditary predisposition to low interleukin-10 production in children with extended oligoarticular juvenile idiopathic arthritis. Rheumatology (Oxford) 2001;40:574–78.
6 De Benedetti F, Ravelli A, Martini A. Cytokines in juvenile rheumatoid arthritis. Curr Opin Rheumatol 1997;9:428–33.
7 De Benedetti F, Pignatti P, Bernasconi S *et al.* Interleukin 8 and monocyte chemoattractant protein-1 in patients with juvenile rheumatoid arthritis. Relation to onset types, disease activity, and synovial fluid leukocytes. J Rheumatol 1999;26:425–31.
8 Ozen S, Saatci U, Bakkaloglu A *et al.* Interleukin-1, -6, and -8 levels in juvenile chronic arthritis. Clin Rheumatol 1997;16:173–8.
9 Yilmaz M, Kendirli SG, Altintas D, Bingol G, Antmen B. Cytokine levels in serum of patients with juvenile rheumatoid arthritis. Clin Rheumatol 2001;20:30–5.
10 Asquith DL, McInnes IB. Emerging cytokine targets in rheumatoid arthritis. Curr Opin Rheumatol 2007;19:246–51.
11 Vignali DA. Multiplexed particle-based flow cytometric assays. J Immunol Methods 2000;243:243–55.
12 de Jager W, Hoppenreijs EP, Wulffraat NM, Wedderburn LR, Kuis W, Prakken BJ. Blood and synovial fluid cytokine signatures in patients with juvenile idiopathic arthritis: a cross-sectional study. Ann Rheum Dis 2007;66:589–98.
13 Petty RE, Southwood TR, Manners P *et al.* International League of Associations for Rheumatology Classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
14 Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
15 de Jager W, Prakken BJ, Bijlsma JW, Kuis W, Rijkers GT. Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. J Immunol Methods 2005;300:124–35.
16 Kameyoshi Y, Dorschner A, Mallet AI, Christophers E, Schroder JM. Cytokine RANTES released by thrombin-stimulated platelets is a potent attractant for human eosinophils. J Exp Med 1992;176:587–92.
17 Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 2006;22:1540–2.
18 Raza K, Falciani F, Curnow SJ *et al.* Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin. Arthritis Res Ther 2005;7:R784–95.
19 Charo IF, Ransohoff RM. The many roles of chemokines and chemokine receptors in inflammation. N Engl J Med 2006;354:610–21.
20 Iellem A, Mariani M, Lang R *et al.* Unique chemotactic response profile and specific expression of chemokine receptors CCR4 and CCR8 by CD4(+)CD25(+) regulatory T cells. J Exp Med 2001;194:847–53.
21 Pharoah DS, Varsani H, Tatham RW *et al.* Expression of the inflammatory chemokines CCL5, CCL3 and CXCL10 in juvenile idiopathic arthritis, and demonstration of CCL5 production by an atypical subset of CD8+ T cells. Arthritis Res Ther 2006;8:R50.
22 Wedderburn LR, Robinson N, Patel A, Varsani H, Woo P. Selective recruitment of polarized T cells expressing CCR5 and CXCR3 to the inflamed joints of children with juvenile idiopathic arthritis. Arthritis Rheum 2000;43:765–74.
23 Ruth JH, Rottman JB, Katschke KJ Jr *et al.* Selective lymphocyte chemokine receptor expression in the rheumatoid joint. Arthritis Rheum 2001;44:2750–60.
24 Harrington LE, Hatton RD, Mangan PR *et al.* Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. Nat Immunol 2005;6:1123–32.
25 Hirota K, Yoshitomi H, Hashimoto M *et al.* Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. J Exp Med 2007;204:2803–12.
26 De Benedetti F, Meazza C, Vivarelli M *et al.* Functional and prognostic relevance of the -173 polymorphism of the macrophage migration inhibitory factor gene in systemic-onset juvenile idiopathic arthritis. Arthritis Rheum 2003; 48:1398–407.
27 Fishman D, Faulds G, Jeffery R *et al.* The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. J Clin Invest 1998;102:1369–76.
28 Ozen S, Alikasifoglu M, Bakkaloglu A *et al.* Tumour necrosis factor alpha G—>A -238 and G—>A -308 polymorphisms in juvenile idiopathic arthritis. Rheumatology (Oxford) 2002;41:223–7.