RHEUMATOLOGY

Rheumatology 2021;60:4765–4777
doi:10.1093/rheumatology/keab098
Advance Access publication 29 January 2021

Original article

A distinguishing profile of chemokines, cytokines and biomarkers in the saliva of children with Sjögren's syndrome

M. Paula Gomez Hernandez[1], Emily E. Starman[2], Andrew B. Davis[3], Miyuraj Harishchandra Hikkaduwa Withanage[4], Erliang Zeng[4], Scott M. Lieberman[5], Kim A. Brogden [2] and Emily A. Lanzel[6]

Abstract

Objective. SS is an autoimmune disease most commonly diagnosed in adults but can occur in children. Our objective was to assess the presence of chemokines, cytokines and biomarkers (CCBMs) in saliva from these children that were associated with lymphocyte and mononuclear cell functions.

Methods. Saliva was collected from 11 children diagnosed with SS prior to age 18 years and 16 normal healthy children. A total of 105 CCBMs were detected in multiplex microparticle-based immunoassays. ANOVA and *t* test (0.05 level) were used to detect differences. Ingenuity Pathway Analysis (IPA) was used to assess whether elevated CCBMs were in annotations associated with immune system diseases and select leukocyte activities and functions. Machine learning methods were used to evaluate the predictive power of these CCBMs for SS and were measured by receiver operating characteristic (ROC) curve and area under curve (AUC).

Results. Of the 105 CCBMs detected, 43 (40.9%) differed in children with SS from those in healthy study controls ($P < 0.05$) and could differentiate the two groups ($P < 0.05$). Elevated CCBMs in IPA annotations were associated with autoimmune diseases and with leukocyte chemotaxis, migration, proliferation, and regulation of T cell activation. The best AUC value in ROC analysis was 0.93, indicating that there are small numbers of CCBMs that may be useful for diagnosis of SS.

Conclusion. While 35 of these 43 CCBMs have been previously reported in SS, 8 CCBMs had not. Additional studies focusing on these CCBMs may provide further insight into disease pathogenesis and may contribute to diagnosis of SS in children.

Key words: chemokines, cytokines, biomarkers, saliva, children, Sjögren's syndrome

CLINICAL SCIENCE

Rheumatology key messages

- Of the 105 chemokines, cytokines and biomarkers (CCBMs) detected, 43 differed in children with Sjögren's syndrome from those in healthy controls ($P < 0.05$).
- Of the 43 CCBMs detected, 35 have previously been detected in adults with Sjögren's syndrome and 8 are new.
- Elevated CCBMs were associated with leukocyte chemotaxis, migration, proliferation, and regulation of T-cell activation.

[1]Pediatric Dentistry, [2]Iowa Institute for Oral Health Research, College of Dentistry, [3]Department of Otolaryngology, College of Medicine, [4]Division of Biostatistics and Computational Biology, College of Dentistry, [5]Division of Rheumatology, Allergy and Immunology, Stead Family Department of Pediatrics, Carver College of Medicine and [6]Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, IA, USA

Submitted 29 October 2020; accepted 17 January 2021

Correspondence to: Emily A. Lanzel, Department of Oral Pathology, Radiology and Medicine, College of Dentistry, S386 DSB, 801 Newton Road, University of Iowa, Iowa City, IA 52242, USA.
E-mail: emily-lanzel@uiowa.edu

Introduction

SS is a chronic, systemic autoimmune disease that primarily involves the salivary and lacrimal glands, resulting in xerostomia and xeropthalmia. It typically occurs in adults around the fourth or fifth decade of life [1]. Diagnosis depends upon the results of clinical and biopsy tests [2, 3]. Characteristic symptoms include dry eyes, dry mouth, fatigue, musculoskeletal pain, and swelling of the major salivary glands. There can be gradual and progressive damage and dysfunction of exocrine glands, which result in whole-body dryness [3]. SS is confirmed when minor salivary gland or parotid gland biopsies show periepithelial mononuclear cell infiltrates [4]. Autoantibodies, acute phase proteins, and inflammatory cell–derived chemokines, cytokines and biomarkers (CCBMs) can be produced and serve as diagnostic biomarkers [5, 6].

The pooled prevalence rate of primary SS is 0.0608% and the pooled incidence rate is 0.0069% [7]; secondary SS is often associated with RA or SLE [8]. In children, the prevalence of SS is not well known, at least in part due to a lack of well-established and highly sensitive criteria for diagnosis [9]. Only 81 cases were reported in 2000–2010, and the age of diagnosis ranged from 4 to 16 years (mean age 9.84 years) [1, 10]. The ratio of female to male in adults is 9:1 and in children is 5:1 [1, 10].

Early events leading to the development of SS are not well known. Triggers of the onset of chronic inflammation may occur decades prior to significant tissue damage. For example, a case–control study assessing the association between infections and SS using inpatient and outpatient data from the Sweden National Health Care registers found that infections of the lung, skin and urogenital tract increase the risk of developing SS as adults, and infections were more prominently associated with developing primary SS with Ro/SSA and La/SSB antibodies [11].

The objective of our study was to assess the presence of CCBMs in the saliva of children with SS, seeking to identify those that may be associated with the presence of immune cell infiltrates in their salivary tissues. We hypothesized that the saliva of these children would have CCBMs similar to those of adults with SS. We also hypothesized that some CCBMs would be novel and associated with lymphocyte and mononuclear cell activities. Overall, a unique profile of CCBMs would have the potential to aid in diagnosing SS in children and may help to identify a window of opportunity to intervene therapeutically to alter disease progression.

Methods

Human samples

Twenty-seven children consented to participate, and 1.0–7.0 ml of unstimulated whole saliva was collected as approved by the Human Institutional Review Board of the University of Iowa (IRB ID No.: 200907702). Between 30 August 2016 and 23 May 2017, we obtained consent and collected data and saliva from 11 children or young adults who had been formally diagnosed with SS by a paediatric rheumatologist prior to age 18 years and were being followed by a paediatric rheumatologist for the management of SS at the time of saliva collection. Diagnosis of SS was based on expert opinion and not restricted to any specific set of criteria. Diagnoses were supported by histopathologic evidence of salivary gland inflammation (Table 1). No child was diagnosed with another systemic rheumatologic disease. Saliva was collected from 16 normal healthy children, matched for gender and age, who served as study controls. These 16 children did not have a contributory health history, were not taking any medications, and did not have significant allergies. They were patients in the Pediatric Dentistry and Orthodontic Clinics at the College of Dentistry and the Department of Pediatrics at the Carver College of Medicine, University of Iowa. The features of all children are shown in Table 1.

Determination of CCBMs

Saliva samples were thawed on ice and centrifuged at a relative centrifugal force of 16 100 (13 200 RPM, Eppendorf, 5415 D centrifuge, Brinkmann Instruments, Inc., Westbury, NY) for 5 min at 24°C to pellet particulates and debris. Supernatants were removed and held on ice.

The concentrations (pg/ml) of 105 CCBMs were determined in each sample, in triplicate, using multiplex fluorescent microparticle-based immunoassays (Luminex Human Magnetic Assay, R&D Systems, Minneapolis, MN). Immunoassays were run according to the manufacturer's instructions on all samples at the same time as so not to create batch effects. CCBM concentrations were interpolated from their median fluorescence intensity (MFI) values using five parameter logistic curves created from the standard concentrations and their respective MFI readings on the Luminex 100 IS using xPonent v3.1 software (Luminex, Austin, TX) or on the readout files using Milliplex Analyst v5.1 software (EMD Millipore, Billerica, MA). CCBM concentrations below the curve were interpolated from their MFI values using curves created from zero concentration to the lowest standard concentration and their respective MFI readings.

Statistical analysis

We calculated a mean value for each of the three replications and then applied a log10-transformation to the concentrations of all CCBMs before analysis and comparisons. The log10-transformation attenuated the positive skew in the distributions of the CCBM concentrations and made their normality assumption more defensible. We used the statistical package SAS® System version 9.4 (SAS Institute Inc., Cary, NC, USA). We used a one-way ANOVA, followed by the *t* test, to detect the differences between CCBMs concentrations in saliva samples. The Benjamini–Hochberg procedure was used

Table 1 Features of children formally diagnosed with SS

ID	Gender	Age (time since diagnosis) years	Focal sialadenitis[a]	Antibodies			Dry eye (subj)	Dry mouth (subj)	Recurrent acute parotitis	Medications (at time of saliva collection)[b]			
				SS-A	SS-B	RF				NSAID	HCQ	PRED	Other
Children, 10–21 years of age, diagnosed with SS													
UIBB.1837	F	17.9 (9.8)	c	+	+	+	+	−	−	−	−	−	−
UIBB.1845	F	10.7 (1.3)	++	+	+	+	−	−	−	−	+	−	−
UIBB.1846	F	15.7 (0.0)	++	−	+	−	−	−	−	+	+	−	−
UIBB.1856	F	19.2 (1.4)	+	−	−	−	+	+	−	+	+	−	−
UIBB.1860	F	16.1 (1.0)	++	−	−	−	+	+	−	−	+	+	MMF, e
UIBB.1886	F	17.9 (1.7)	++	+	−	−	+	+	−	−	−	−	−
UIBB.1904	F	16.6 (1.2)	++	−	−	−	+	−	−	−	+/−	–	−
UIBB.1907	F	18.9 (1.5)	++	−	−	−	+	+	+	+/−	+/−	−	−
UIBB.1914	F	14.7 (1.4)	+	−	−	−	−	+	+	−	+/−	−	−
UIBB.1950	F	20.9 (10.0)	d	+	+	+	+	+	+	−	−	−	e
UIBB.1965	M	11.5 (1.4)	++	−	+	−	+	−	−	−	−	−	−
Children, matched for gender and age, who served as study controls													
UIBB.2195	F	18											
UIBB.2189	M	11											
UIBB.2193	F	17											
UIBB.2194	F	19											
UIBB.2190	F	17											
UIBB.2192	F	15											
UIBB.2120	F	10											
UIBB.2185	F	20											
UIBB.2191	F	10											
UIBB.2179	F	15											
UIBB.2187	F	16											
UIBB.2073	F	14											
UIBB.2118	F	18											
UIBB.2181	M	11											
UIBB.2188	F	17											
UIBB.2072	F	16											

F: female; M: male; subj: subjective; PRED: prednisone. [a]++, focus score $\geq$ 1 focus/4 mm^2; +, foci present but focus score not reported or $<$1 focus/4 mm^2. [b]+, taking; –, not prescribed; ±, prescribed but not taking regularly. [c]lymphoplasmacytic infiltrate and benign lymphoepithelial lesions. [d]extensive lymphocytic infiltrates with germinal centres and acinar atrophy (noted on histopathology of resected parotid glands). [e]rituximab was given $\geq$8 months prior to saliva sampling.

to control the false discovery rate. All statistical tests utilized a 0.05 level of significance.

Hierarchical clustering

Hierarchical clustering was used to illustrate differences in the mean values of CCBMs in saliva samples identified as being different in statistical analysis. We used Euclidean distance function as a method to calculate the dissimilarity of two CCBM profiles, and used average linkage to calculate the distance between two clusters.

Principal component analysis

Principal component analysis (PCA) analysis was performed on the mean values of CCBMs in saliva samples using R package *rgl* (https://cran.r-project.org/web/packages/rgl/).

Ingenuity pathway analysis

Ingenuity pathway analysis (IPA) (Qiagen, Redwood City, CA) was used to assess whether the significant CCBM responses were in Canonical Pathway annotations associated with immune system diseases or underlying causes and Function and Disease annotations associated with leukocyte activities and functions. Statistical significance was calculated using Fisher's Exact Test.

Feature selection and classification

Features here refer to CCBMs. The concentrations of all CCBMs produced in the statistical analysis were standardized and scaled between 0 and 1. Five feature selection methods (i.e. Correlation [12], Information Gain [13], Information Gain Ratio [14], Symmetrical Uncertainty [15], and RELIEF [16]) were applied to rank CCBMs. CCBMs were also ranked using an ensemble method that ranks them by aggregating ranks of all five aforementioned feature selection methods. Furthermore, discrete sets of CCBMs (i.e. Union, AtLeast2, AtLeast3, AtLeast4, SelectedByFive) were obtained using a Venn diagram representing overlapped features of the five feature selection methods. The Union feature set at a given rank k is the union of all top k features selected by five methods. The SelectedByFive feature set at a given rank k are the common features of all top k feature sets each selected by five methods. This definition is analogous for feature sets of AtLeast2 , AtLeast3 and AtLeast4 . For example, the AtLeast2 feature set at a given rank k includes the top k features that are selected by at least two feature selection methods.

Each top k feature set was evaluated using classification methods (classifiers). The classifiers used were K-Nearest Neighbour (k = 3) [17], AdaBoost (trees = 100) [18], Support Vector Machine (Linear Kernel) [19], Support Vector Machine (rbf Kernel) [20], Naïve Bayes [21], Random Forest (trees = 100) [22], Logistic Regression, and Gaussian Process [23]. The classifier used selected features to train a model that then could be used to predict SS. The performance of a classifier on specific feature set was evaluated using leave-one-out cross-validation and Receiver Operating Characteristic (ROC) curve. The average Area Under Curve (AUC) was used to measure the performance.

Results

CCBMs in saliva

Forty-six of 105 CCBMs ranged from undetected to 969.3 pg/ml; 48 of the 105 CCBMs ranged from 1.0 to 799.3 ng/ml , and 11 of the 105 CCBMs ranged from 1.0 to 239.6 μg/ml. In the latter group, C9, B2M, MMP9, TIMP1, and AMBP concentrations were among the highest detected [Fig. 1, Supplementary Fig. S1, available at *Rheumatology* online) [24] and Mendeley Data repository (http://dx.doi.org/10.17632/yphm77tg24.1].

Fifty-five of the 105 CCBMs were significantly higher in concentration ($P < 0.05$) in the children diagnosed with SS *vs* the healthy children study controls. Of these 55, 43 CCBMs were significantly different in concentration ($P < 0.05$) after correction using the Benjamini–Hochberg procedure to control the false discovery rate and avoid type 1 false-positive errors. Thus, 40.9% (43 of 105) of the CCBMs we selected varied between the children diagnosed with SS *vs* the healthy children study controls (Fig. 1, Table 2).

Thirty-five of these 43 CCBMs have been previously detected in other studies assessing CCBMs in tears, blood, and saliva of individuals with SS by proteomics, mass spectroscopy, and immunoassay analyses (Supplementary Table S1, available at *Rheumatology* online). CALCA, CCL1, CCL8, CCL26, GDF2, IL2RA, MIA and ULBP2 were found to be unique.

Hierarchical clustering

There were broad ranges in concentrations of CCBMs in 11 children with and 16 children without SS. Seven of the 11 older children with SS (aged 16–20) formed a cluster separate from the 16 healthy study control children (Fig. 2). Four of the 11 younger children with SS (aged 10–16) were found within the clusters of the 16 healthy study control children. Two of these latter children with SS (e.g. UIBB 1845 and 1860) clustered together.

PCA

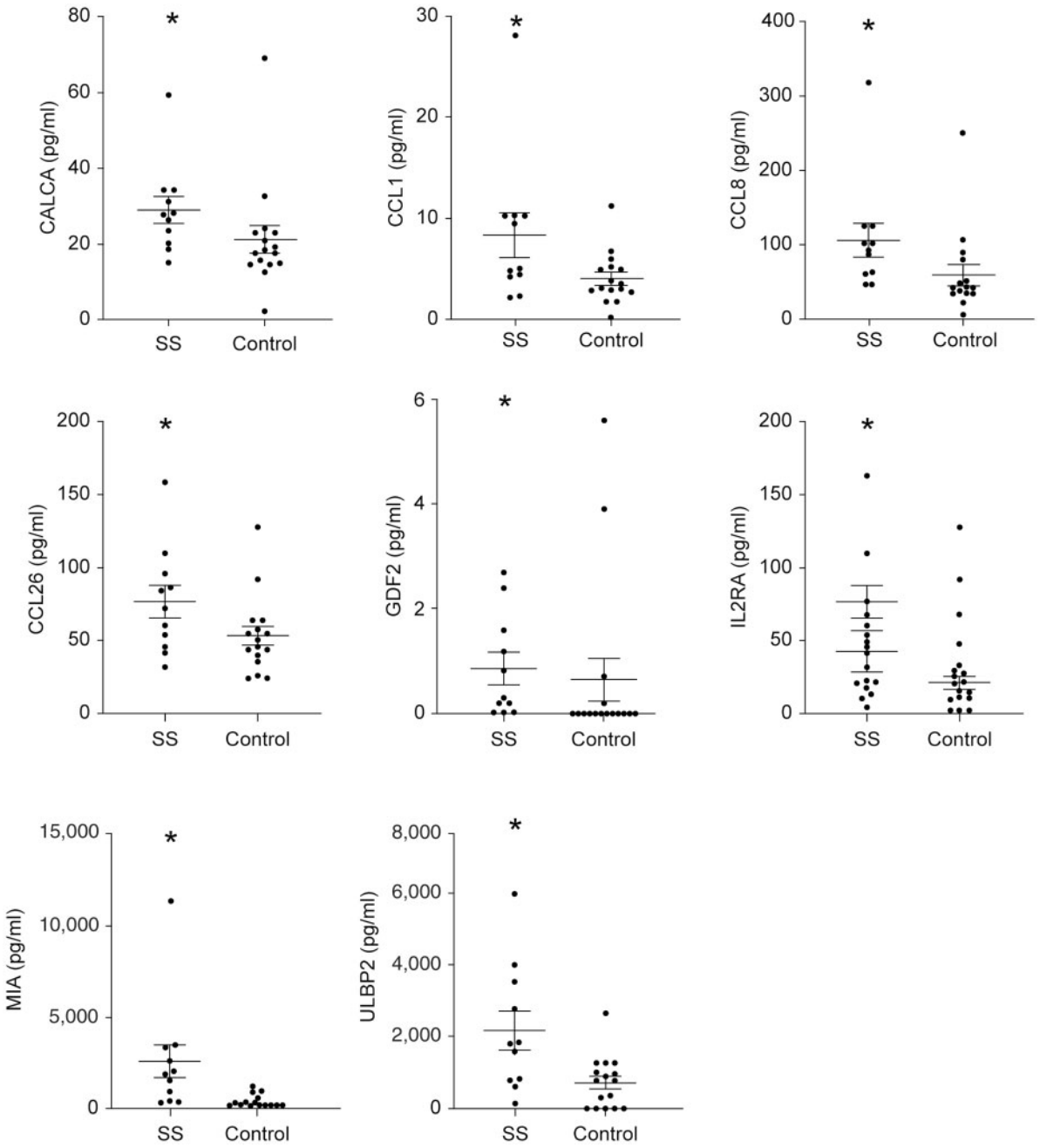
If the biomarkers were SS-specific, we would anticipate that children with SS and normal controls would be well separated in the space formed by the principal components calculated from syndrome-specific biomarker profiles. PCA analysis of the mean values of the 43 significant CCBMs was able to differentiate between the 11 children diagnosed with SS and the 16 healthy study control children (Fig. 3).

IPA

Forty-three CCBMs in children diagnosed with SS were in Canonical Pathway annotations associated with CTDs, immunological diseases, inflammatory diseases, and infectious diseases (Supplementary Table S2, available at *Rheumatology* online). CTDs included inflammation of joint, rheumatic disease, polyarthritis, RA, CIA and LE. Immunological diseases included hypersensitive reactions, systemic autoimmune syndromes, and delayed hypersensitive reactions. Finally, infectious diseases included CCBMs involved in microbial infections and sepsis, parasitic infections, and viral infections, including replication of RNA viruses and replication of viruses.

Many of these CCBMs were annotated to functions associated with (i) cellular movement, (ii) immune cell trafficking and (iii) cell-to-cell signalling and interaction (Supplementary Table S3, available at *Rheumatology* online). The most significant functions were associated

Fig. 1 Plots of eight newly identified CCBMs in saliva of childrenformally diagnosed with SS



The mean and s.e.m. bars are included. The asterisk shows significance at the $P < 0.05$ level. CCBMs: chemokines, cytokines, and biomarkers.

with cell movement of mononuclear leukocytes, cell movement of lymphocytes, lymphocyte migration, cell movement of leukocytes, chemotaxis of mononuclear leukocytes, T cell migration, cell movement of myeloid cells, or leukocyte migration. Significant functions were also associated with cell-to-cell signalling that included recruitment of cells, recruitment of mononuclear leukocytes, recruitment of leukocytes, activation of cells, cell movement of T cells, activation of leukocytes, activation of lymphocytes, response of mononuclear leukocytes, and activation of T cells.

Table 2 Descriptive statistics of CCBMs in children with SS and healthy study controls

CCBM Symbol	Children with SS (n = 11) pg/ml	Children matched for age and gender (n = 16) pg/ml
AMBP[a]	4 156 363.6[b] (763 043.3)[c] 1 790 000.0–8 590 000.0[d]	24 558 50.1 (587 895.1) 0.0–8 540 000.0
B2M[a]	122 671.6 (14 485.0) 98 850.0–266 527.2	98 060.6 (6038.4) 9110.0–111 270.0
C9[a]	897 454.5 (125 229.6) 378 000.0–1 540 000.0	1 129 250.0 (326 143.0) 323 000.0–5 720 000.0
CA9[a]	11.1 (2.7) 0.0–24.4	5.1 (0.0) 0.0–13.3
CALCA	29.2 (3.6) 15.2–59.5	21.5 (3.6) 2.4–69.3
CCL1	8.3 (2.2) 2.2–28.1	4.0 (0.6) 0.1–11.2
CCL4[a]	175.3 (34.5) 12.1–311.8	27.2 (12.7) 0.0–164.2
CCL8	106.7 (22.8) 46.7–318.0	59.8 (14.2) 6.5–250.4
CCL15	479.7 (125.9) 0.0–1284.1	285.6 (139.4) 0.0–1932.1
CCL17[a]	73.5 (8.3) 45.6–139.4	56.3 (5.4) 1.9–89.4
CCL26	76.8 (11.0) 32.1–158.9	53.4 (6.6) 24.3–128.0
CCL27[a]	4.6 (0.8) 1.7–9.7	2.6 (0.3) 0.6–5.8
CCL28[a]	4826.9 (856.2) 184.8–8866.1	5807.6 (1649.6) 604.6–25993.3
CXCL10[a]	694.0 (230.9) 70.2–2567.6	191.8 (70.1) 1.6–790.9
CXCL11[a]	33.8 (6.7) 4.2–64.4	10.4 (3.7) 0.0–58.2
FSTL1[a]	1441.6 (874.0) 0.0–9626.5	499.4 (466.3) 0.0–7482.9
GAS6[a]	4065.2 (1002.7) 479.6–10 733.3	2363.3 (319.5) 733.5–5406.1
GDF2	0.9 (0.3) 0.0–2.7	0.7 (0.4) 0.0–5.6
IFNA1[a]	30.6 (6.6) 13.4–90.1	28.6 (10.5) 2.5–181.1
IFNB1[a]	14.6 (2.4) 6.2–30.6	8.2 (1.4) 1.5–20.0
IFNG[a]	197.1 (32.6) 76.5–462.2	125.5 (27.6) 26.7–485.8
IFNGR1[a]	15.2 (2.5) 5.9–31.5	9.6 (1.5) 0.3–24.1
IL1B[a]	1141.9 (255.5) 217.0–2695.5	419.6 (132.1) 4.0–2198.9
IL6[a]	26.8 (4.4) 7.4–51.7	15.0 (3.3) 1.4–58.6
IL10[a]	5.2 (1.6) 0.5–13.9	2.4 (0.9) 0.0–13.0
IL12A[a]	988.2 (68.8) 701.0–1405.2	842.3 (69.6) 86.6–1198.0
IL12B[a]	304.6 (84.6) 7.5–716.1	57.2 (16.2) 0.0–226.9
IL21[a]	67.9 (9.4) 39.6–147.0	43.6 (7.9) 4.3–141.7
IL23A[a]	818.9 (208.8) 44.2–2088.6	253.0 (102.1) 1.4–1208.3

(continued)

TABLE 2 Continued

CCBM Symbol	Children with SS ($n = 11$) pg/ml	Children matched for age and gender ($n = 16$) pg/ml
IL27[a]	251.4 (52.6) 0.0–507.9	34.2 (12.7) 0.0–170.3
IL2RA	42.8 (14.1) 4.4–163.6	21.6 (4.4) 1.9–68.0
IRX1[a]	7.5 (1.7) 0.0–16.2	6.4 (3.3) 0.0–52.3
MIA	2600.4 (941.2) 346.9–11 360.0	398.3 (86.7) 120.8–1217.2
MMP9[a]	319 415.4 (78 440.8) 44 963.3–773 330.0	123 722.3 (27 752.9) 19 960.0–399 156.7
PECAM1[a]	1330.4 (506.2) 14.0–5521.4	412.4 (190.5) 0.0–2723.2
S100A8[a]	21 952.3 (3905.9) 7453.4–47 880.0	9254.6 (2206.6) 1928.6–32 826.7
TIMP1[a]	111 442.1 (1752.2) 98 626.7–118 480.0	102 741.4 (6263.5) 9962.5–113 810.0
TNFRSF1B[a]	246.7 (73.1) 68.8–865.3	115.4 (18.4) 20.7–348.3
TNFRSF8[a]	15.2 (2.4) 6.6–30.7	6.9 (1.1) 0.1–18.4
TNFRSF13B[a]	263.6 (13.4) 202.2–342.8	234.0 (17.8) 19.2–308.0
TNFRSF18[a]	178.0 (15.1) 101.9–246.1	101.6 (9.7) 57.1–171.8
TSLP[a]	16.8 (2.4) 7.1–33.1	8.6 (1.3) 3.6–23.7
ULBP2	2180.0 (534.1) 149.9–6014.8	725.8 (179.9) 0.0–2667.5

[a]CCBMs identified by proteomics, mass spectroscopy, and immunoassays in tears, blood and saliva of individuals with SS (see Supplemental Table 1, available at *Rheumatology* online). [b]mean. [c]S.E.M. [d]minimum–maximum values.

Feature selection and classification

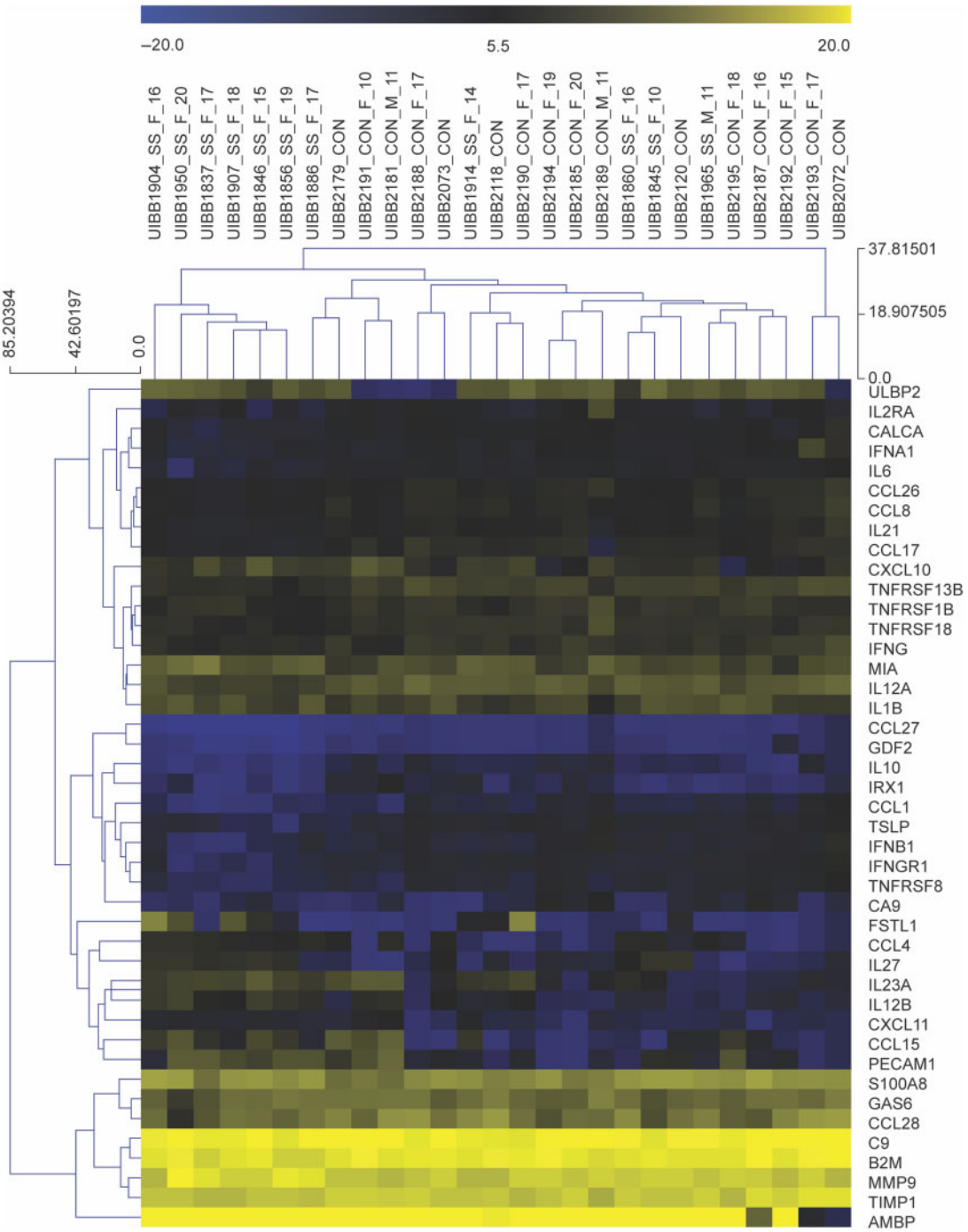
We tested eight classifiers of the top k different sets of CCBMs. We measured the performance of each classifier of feature sets of various k values to identify the best- performing models. The ROC curves and the AUC values indicated that a small number of CCBMs could be identified and used as predictor markers for SS diagnosis in children (Fig. 4, Supplementary Table S4, available at *Rheumatology* online). Among the best-performing models, the k-Nearest Neighbour classifier had the highest AUC value (AUC = 0.93, Fig. 4) for a feature set consisting of only two CCBMs: IL27 and CCL4. It is worth mentioning that two different feature sets could have the same prediction power, that is, the same AUC values (Supplementary Table S4, available at *Rheumatology* online).

Discussion

We selected 105 CCBMs associated with lymphocyte and mononuclear cell functions and determined their concentrations in the saliva of 11 children diagnosed with SS and 16 normal healthy children who served as study controls. We found that 43/105 CCBMs were significantly different in the children with SS. Upon closer examination, 35 CCBMs in this profile have been reported to be present in individuals with SS [25] (Supplementary Table S1, available at *Rheumatology* online). Eight CCBMs in this profile (e.g. CALCA, CCL1, CCL8, CCL26, GDF2, IL2RA, MIA and ULBP2) were unique and have not been reported to be associated with SS (Supplementary Table S1, available at *Rheumatology* online).

The disease in children is not well defined, and the criteria for diagnosing SS are different [9, 26]. For example, xerostomia and xerophthalmia are not often the primary manifestations. Approximately half of the children present with parotitis and the other half present with less-specific clinical features such as joint pain [9, 27–29]. Our study identified CCBMs in saliva of children with SS that may provide additional measures for use in developing child-specific diagnostic criteria. Prospective studies are needed to assess the utility of these salivary CCBMs in the diagnosis of SS in children and to determine whether these CCBMs may contribute to a better understanding of the pathogenesis of SS. Notably, the ability to more easily and

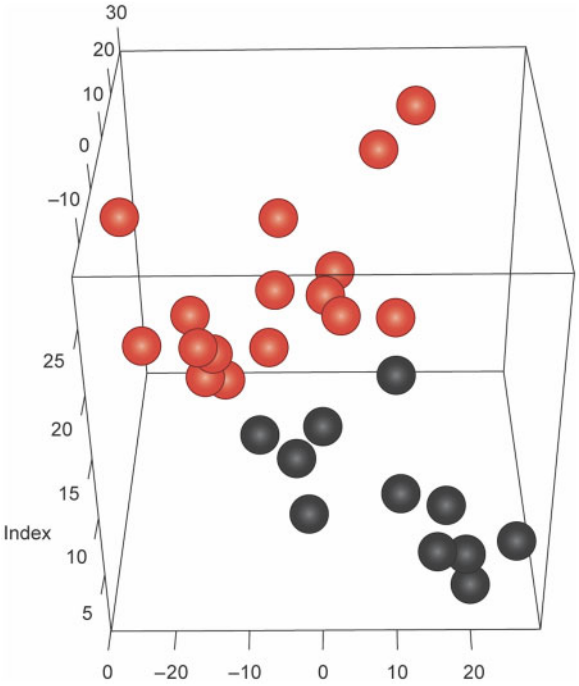
Fig. 2 Children and CCBMs were grouped together using a two-wayhierarchical clustering approach



There were broad ranges in concentrations with consistent patterns. Seven of 11 older children with SS (aged 16–20) formed a cluster separate from the 16 healthy study control children. Four of 11 younger children with SS (aged 10–16) were found within the clusters of the 16 healthy study control children. CCBMs: chemokines, cytokines, and biomarkers.

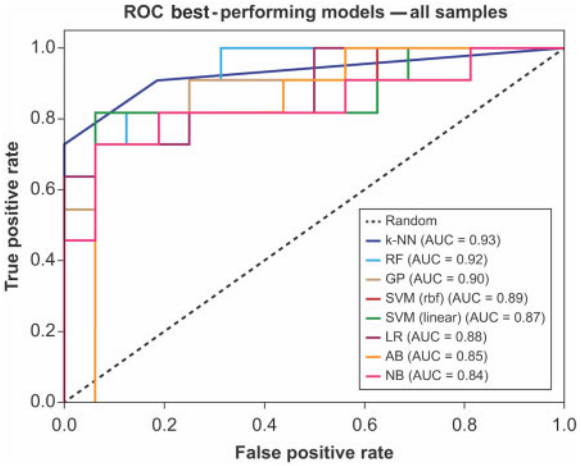
objectively diagnose SS in children may aid in identification of a window of opportunity for therapeutic interve ntion, which may enable prevention of progression to the classic profound sicca symptoms that develop over time. This is a reasonable concept, as anti-SSA/Ro and anti-SSB/La antibodies can also be detected up to 18–20 years before the appearance of symptoms and diagnosis of SS [30, 31].

Fig. 3 PCA of 43 CCBMs could differentiate between children with SS and healthy study controls



The PCA illustration shows the mean values of CCBMs in the children's saliva samples for each group. The 43 significant CCBMs could differentiate these two groups, indicating that the effect of all significant CCBMs collectively contributes to SS. CCBMs: chemokines, cytokines, and biomarkers.

Fig. 4 ROC curves of classifiers on different feature sets that predict CCBMs for SS in children



The area under curve (AUC) values indicate CCBMs can be served as predictor biomarkers for SS diagnosis in children. Eight classifiers were used, including k-NN: k-Nearest Neighbour, RF: Random Forest, GP: Gaussian Process, SVM (rbf): Support Vector Machine with rbf Kernel, SVM (Linear): Support Vector Machine with Linear Kernel, LR: Logistic Regression, AB: AdaBoost, and NB: Naïve Bayes. CCBMs: chemokines, cytokines, and biomarkers.

The profile of 43 CCBMs were related to CTDs, immunological diseases, inflammatory diseases, and infectious diseases (Supplementary Table S2, available at *Rheumatology* online). These included elevated concentrations of GAS6, FSTL1, MMP9 and IFNA1. Elevated GAS6 levels occur in individuals with SLE and liver fibrosis [32, 33]. Elevated FSTL1 levels occur in individuals with RA, ulcerative colitis, SLE, SS, SSc, and PM/ DM [34]. Elevated MMP9 levels occur in individuals with multiple sclerosis [35] and in individuals with SS [36, 37]. IFNA1 is involved in the pathogenesis of SLE [38] and SS [39, 40].

There were significant associations of CCBM responses with CTDs (Supplementary Table S2, available at *Rheumatology* online). Functional annotations were associated with inflammation of joints (27 CCBMs), rheumatic disease (28 CCBMs), polyarthritis (13 CCBMs), RA (20 CCBMs), CIA (10 CCBMs) and LE (12 CCBMs). There were significant associations of CCBM responses with immunological diseases (Supplementary Table S2, available at *Rheumatology* online). These included hypersensitive reactions (19 CCBMs), systemic autoimmune syndromes (25 CCBMs) and delayed hypersensitive reactions (9 CCBMs).

Interestingly, some of these 43 CCBMs were found in annotations associated with microbial infections, often hypothesized to be triggers for SS. There were functional annotations of CCBMs associated with microbial infections (14 CCBMs) and sepsis (12 CCBMs), parasitic infections (11 CCBMs), and viral infections (24 CCBMs). There were CCBMs in annotations associated with antimicrobial innate immune functions, and these included antimicrobial (18 CCBMs) and antiviral responses (14 CCBM). This is supported by the reports that B2M, CALCA, CCL1, CCL17, CXCL10 and FSTL1 have antimicrobial activities; IFNA1 has antiviral activity, and IFNG can trigger a cellular response to both viral and microbial infections. To what extent viral and or bacterial infections lead to auto- inflammatory events and SS remains to be determined in separate studies.

SS is characterized by lymphocytic infiltrates in the exocrine glands [4]. Infiltrates contain mononuclear cells in SS lesions that vary according to lesion severity and correlate with disease manifestations [41, 42]. T cell infiltrates can lead to secondary B cell activation and autoantibody production [43]. We found that many of these 43 CCBMs were in annotations associated with cellular movement, immune cell trafficking, and cell-to-cell signalling interactions (Supplementary Table S3, available at *Rheumatology* online). Ten elevated chemokines attract T cells (CCL8, CCL15, CCL17, CCL27, CCL28, CXCL10 and CXCL11); NK cells (CCL4 and CXCL10); monocytes (CCL1, CCL4, CCL8, CCL15 and CXCL10), eosinophils (CCL8) and basophils (CCL8 and CCL26). This profile would create an environment rich in monocytes, T cells, eosinophils and basophils. This likely sets

up an inflammatory environment in glandular tissue in which additional chemokines are produced by activated T cells (CCL1) and $CD8^+$ T cells (CCL4), further recruiting and activating arriving T cells (CCL17).

We observed higher ($P < 0.05$) concentrations of four CCBMs in the IL12 family of heterodimeric cytokines that included IL12A, IL12B, IL23 and IL27 [44]. This family mediates a number of diverse immunoregulatory activities. IL12 and IL23 are considered pro-inflammatory and pro-stimulatory cytokines active on Th1 and Th17 T cells, whereas IL27 is considered to be an immunoregulatory cytokine [44]. Furthermore, IL27 can induce the production of IFNG and IL10 (both elevated in Fig. 1) and can regulate other T cell subsets and also support chemokine responses. IL27 can act as a critical initiator of adaptive immune responses by promoting the rapid clonal expansion of naïve $CD4^+$ T cells, IFNG production, and TH1 polarization. The antibody in this immunoassay is detecting the heterodimer and does not detect a recombinant EBI-3 monomer. We assume the antibody pair used in the immunoassay can detect the IL27p28 monomer, but we have not tested that specifically.

We observed higher concentrations of IL21, a cytokine secreted from T cells that can induce the differentiation, proliferation and activity of macrophages, NK cells, B-cells and cytotoxic T cells [45] and induce the production of IFNG.

IFNs are hypothesized to play an important role in the pathogenesis of SS [39, 40, 46], and we observed significant increases in Type I IFNs (IFNA1 and IFNB1), Type II IFN (IFNG) and IFNGR1 responses. IFNA1 and IFNB1 activate cellular antiviral effects and enhance the cytotoxic activity of NK cells and macrophages ; IFNA1 primes $CD8^+$ T cells and IFNA1 activates cytotoxic T cells. IFNs can also upregulate other CCBMs like B2M [47].

CALCA, CCL1, CCL8, CCL26, GDF2, IL2RA, MIA and ULBP2 were unique CCBM responses in that they have not been previously reported to be associated with SS (Supplementary Table S1, available at *Rheumatology* online). The functions of chemokines CCL1, CCL8 and CCL26 were mentioned above. CALCA is a pro-peptide of calcitonin that has vasodilator activity, antimicrobial activity, and a role in calcium metabolism. It can occur in response to microbial and viral infections and can be associated with inflammatory responses [48, 49]. A recent example is the presence of CALCA in patients with COVID-19 [50]. In annotations, CALCA was associated with infections, inflammatory responses, and monocyte chemotaxis. GDF2 is a secreted ligand of the TGF-β superfamily of proteins and can have cytokine activities and CCBM regulatory functions. For example, GDF2 can recruit human peripheral blood monocytes [51] and decrease the activity of MMP-9 [52]. IL2RA is required for mediating IL2- induced effects, including T cell proliferation [53], T cell activation [54], and activation-induced cell death of T cells [55]. In annotations, it had a role in apoptotic processes and activation-induced cell death of T cells. MIA originally isolated from melanoma cell cultures has been used as an indicator for tumour load [56, 57]. It is expressed in salivary gland tissue [58], and in annotations, it is thought to have growth factor activity. ULBP2 is one of a family of cell membrane proteins expressed on both transformed and stressed cells [59]. It is a stress-induced molecule and a ligand for NKG2D that activates NK cells and provides co-stimulation for T cells [60]. In context here, ULBP2 is regulated by IL12 (family), and ULBP2 is thought to regulate the production and expression of IFNG, CCL1 and CCL4 [61, 62]. ULBP2 also plays a role in cell activation, proliferation, and killing by NK cells and cytotoxic T cells [62, 63].

Forty-three of 105 CCBMs is an unreasonably large number of markers to be of practical use for differentiating children with SS from healthy study controls. Smaller combinations of CCBMs have been reported as identifying adults with SS. For example, a 4-plex profile (containing FGF-4, clusterin, IL4 and IL5) and a 6-plex profile was able to differentiate adults with SS from healthy study controls [64]. A 3-plex profile containing CPD, α-enolase and B2M was able to differentiate adults with SS and had an AUC value of 0.99 [65]. The accuracy of these marker sets for children is not known. Therefore, we used ROC analysis on the 105 CCBMs to identify feature sets of CCBMs that may also have accurate diagnostic ability. There were strong performing feature sets containing 2–5 CCBMs with AUC values ranging from 0.89 for IL27, MIA, CCL4, TNFRSF18 and TNFA to 0.92 for IL27 and CCL4 (Supplementary Table S4, available at *Rheumatology* online). Among the best- performing models, the k-Nearest Neighbour classifier had the highest AUC value (AUC = 0.93, Fig. 4) on a feature set consisting of only two CCBMs: IL27 and CCL4. These CCBMs were also in the group of 43/105 CCBMs that were found to be different ($P < 0.05$) in children with SS compared with the healthy study controls.

In summary, additional studies are needed to determine whether the new CCBMs identified here in the saliva of children with SS represent reliable early markers of disease or, rather, a paediatric-specific disease process. Further understanding of the pathogenic roles of these CCBMs may provide new targets for therapeutic intervention. Ultimately, objective measures for diagnosing SS in children may provide a unique window of opportunity in which to initiate immunomodulating therapies to alter the course of disease progression that is currently not possible.

Acknowledgements

We acknowledge the support and help of C. Michael Knudson MD, PhD, Kristen Coleman MS, Rita D. Sigmund MA and Joe Galbraith, Tissue Procurement Core, University of Iowa Carver College of Medicine, The University of Iowa, Iowa City, IA 52242. We also acknowledge the support and help of Kevin Knudtson, PhD, Director, Genomics Division, Iowa Institute of Human Genetics, University of Iowa Carver College of Medicine, The University of Iowa, Iowa City, IA USA.

Funding: This work was supported by a Pilot Research Grant from the Sjögren's Syndrome Foundation.

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

Data availability statement

The CCBM responses in the saliva from 11 children diagnosed with SS and 16 normal healthy children will be published as a dataset in Data-in-Brief [24] and is published as a dataset in Mendeley Data repository (http://dx.doi.org/10.17632/yphm77tg24.1).

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Patel R, Shahane A. The epidemiology of Sjögren's syndrome. Clin Epidemiol 2014;6:247–55.

2 Jonsson R, Brokstad KA, Jonsson MV, Delaleu N, Skarstein K. Current concepts on Sjögren's syndrome – classification criteria and biomarkers. Eur J Oral Sci 2018;126:37–48.

3 Vivino FB, Bunya VY, Massaro-Giordano G *et al.* Sjögren's syndrome: an update on disease pathogenesis, clinical manifestations and treatment. Clin Immunol 2019;203:81–121.

4 Barone F, Campos J, Bowman S, Fisher BA. The value of histopathological examination of salivary gland biopsies in diagnosis, prognosis and treatment of Sjögren's syndrome. Swiss Med Wkly 2015;145:w14168.

5 Tzec-Pérez A, Barbosa-Cobos RE, Lugo-Zamudio G *et al.* Cytokine and chemokine levels in serum and saliva as disease activity biomarkers in patients with primary Sjögren's syndrome [abstract 663]. Arthritis Rheumatol 2016;68(suppl 10). https://acrabstracts.org/abstract/cytokine-and-chemokine-levels-in-serum-and-tears-as-disease-activity-biomarkers-in-patients-with-primary-sjogrens-syndrome/ (2 March 2021, date last accessed).

6 Katsiougiannis S, Wong DT. The proteomics of saliva in Sjögren's syndrome. Rheum Dis Clin North Am 2016;42:449–56.

7 Qin B, Wang J, Yang Z *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. Ann Rheum Dis 2015;74:1983–9.

8 Moutsopoulos HM. Sjögren's syndrome: autoimmune epithelitis. Clin Immunol Immunopathol 1994;72:162–5.

9 Yokogawa N, Lieberman SM, Sherry DD, Vivino FB. Features of childhood Sjögren's syndrome in comparison to adult Sjögren's syndrome: considerations in establishing child-specific diagnostic criteria. Clin Exp Rheumatol 2016;34:343–51.

10 de Souza TR, Silva IH, Carvalho AT *et al.* Juvenile Sjögren syndrome: distinctive age, unique findings. Pediatr Dent 2012;34:427–30.

11 Mofors J, Arkema EV, Bjork A *et al.* Infections increase the risk of developing Sjögren's syndrome. J Intern Med 2019;285:670–80.

12 Hall M. Correlation-based feature selection for discrete and numeric class machine learning. In: Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000);Stanford, CA,USA: Stanford University, 2000; 359–66.

13 Jiang X, Jao J, Neapolitan R. Learning predictive interactions using information gain and Bayesian network scoring. PLoS One 2015;10:e0143247.

14 Praveena Priyadarsini R, Valarmathi ML, Sivakumari S. Gain ratio based feature selection method for privacy preservation. ICTACT J Soft Comput 2011;1:201–5.

15 Singh B, Kushwaha N, Vyas OP. A feature subset selection technique for high dimensional data using symmetric uncertainty. J Data Anal Inform Process 2014;2:95–105.

16 Kononenko I. Estimating attributes: analysis and extensions of RELIEF. In: Bergadano F, De Raedt L, eds. Machine learning: ECML-94: European conference on machine learning, 6 April 1994; Catania, Italy: Springer Berlin Heidelberg, 1994; 171–82.

17 Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE T Inform Theory 1967;13:21–7.

18 Kim TH, Park DC, Wou DM, Jeong T, Min SY. Multi-class classifier-based Adaboost Algorithm. In: Y Zhang, ZH Zhou, C Zhang, L Ying, eds. Intelligent science and intelligent data engineering. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 122–7.

19 Hastie T, Friedman J, Tibshirani R. Support vector machines and flexible discriminants. In: T Hastie, J Friedman, R Tibshirani, eds. The elements of statistical learning: data mining, inference, and prediction. New York, NY: Springer New York, 2001: 371–409.

20 Wang WJ, Xu ZB, Lu WZ, Zhang XY. Determination of the spread parameter in the Gaussian kernel for classification and regression. Neurocomputing 2003;55:643–63.

21 Rennie JDM, Shih L, Teevan J, Karger DR. Tackling the poor assumptions of naive Bayes text classifiers. In: Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003); Washington, DC, USA: AAAI Press, 2003;20:616.

22 Breiman L. Random Forests. Machine Learn 2001;45:5–32.

23 Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge, MA: The MIT Press, 2005.

24 Hikkaduwa Withanage MH, Gomez Hernandez MP, Starman EE, Davis AB, Zeng E, Lieberman SM. Dataset – chemokines, cytokines, and biomarkers in the saliva of children with Sjögren's syndrome. Data Brief 2021.

25 Ruacho G, Kvarnström M, Zickert A *et al.* Sjögren's syndrome in Systemic Lupus Erythematosus: a subset characterized by a systemic inflammatory state. J Rheumatol 2020;47:865–75.

26 Kumar KJ, Kudakesseril AS, Sheeladevi CS, Sowmya HV. Primary Sjögren's syndrome in a child. Iran J Pediatr 2015;25:e254.

27 Lieberman SM. Childhood Sjögren syndrome: insights from adults and animal models. Curr Opin Rheumatol 2013;25:651–7.

28 Means C, Aldape MA, King E. Pediatric primary Sjögren syndrome presenting with bilateral ranulas: a case report and systematic review of the literature. Int J Pediatr Otorhinolaryngol 2017;101:11–9.

29 Basiaga ML, Stern SM, Mehta JJ *et al.* Childhood Sjögren syndrome: features of an international cohort and application of the 2016 ACR/EULAR classification criteria. Rheumatology (Oxford) 2020; Advance Access published 6 December 2020, doi: 10.1093/rheumatology/keaa757

30 Jonsson R, Theander E, Sjöström B, Brokstad K, Henriksson G. Autoantibodies present before symptom onset in primary Sjögren's syndrome. JAMA 2013;310: 1854–5.

31 Theander E, Jonsson R, Sjostrom B *et al.* Prediction of Sjögren's syndrome years before diagnosis and identification of patients with early onset and severe disease course by autoantibody profiling. Arthritis Rheumatol 2015;67:2427–36.

32 Cohen PL, Shao WH. Gas6/TAM receptors in systemic lupus erythematosus. Dis Markers 2019;2019:7838195.

33 Smirne C, Rigamonti C, De Benedittis C *et al.* Gas6/TAM signaling components as novel biomarkers of liver fibrosis. Dis Markers 2019;2019:2304931.

34 Li D, Wang Y, Xu N *et al.* Follistatin-like protein 1 is elevated in systemic autoimmune diseases and correlated with disease activity in patients with rheumatoid arthritis. Arthritis Res Ther 2011;13:R17.

35 Prince HE. Biomarkers for diagnosing and monitoring autoimmune diseases. Biomarkers 2005;10: 44–9.

36 Asatsuma M, Ito S, Watanabe M *et al.* Increase in the ratio of matrix metalloproteinase-9 to tissue inhibitor of metalloproteinase-1 in saliva from patients with primary Sjögren's syndrome. Clin Chim Acta 2004;345:99–104.

37 Ito K, Funayama S, Hitomi Y *et al.* Proteome analysis of gelatin-bound salivary proteins in patients with primary Sjögren's syndrome: identification of matrix metalloproteinase-9. Clin Chim Acta 2009;403:269–71.

38 Ronnblom L. The importance of the type I interferon system in autoimmunity. Clin Exp Rheumatol 2016;34: 21–4.

39 Marketos N, Cinoku I, Rapti A, Mavragani CP. Type I interferon signature in Sjögren's syndrome: pathophysiological and clinical implications. Clin Exp Rheumatol 2019;37(Suppl 118):185–91.

40 Thorlacius GE, Wahren-Herlenius M, Ronnblom L. An update on the role of type I interferons in systemic lupus erythematosus and Sjögren's syndrome. Curr Opin Rheumatol 2018;30:471–81.

41 Kapsogeorgou EK, Christodoulou MI, Panagiotakos DB *et al.* Minor salivary gland inflammatory lesions in Sjögren syndrome: do they evolve? J Rheumatol 2013; 40:1566–71.

42 Christodoulou MI, Kapsogeorgou EK, Moutsopoulos HM. Characteristics of the minor salivary gland infiltrates in Sjögren's syndrome. J Autoimmun 2010;34:400–7.

43 Roescher N, Tak PP, Illei GG. Cytokines in Sjögren's syndrome. Oral Dis 2009;15:519–26.

44 Vignali DA, Kuchroo VK. IL-12 family cytokines: immunological playmakers. Nat Immunol 2012;13:722–8.

45 Lim SA, Nam DH, Lee JH *et al.* Association of IL-21 cytokine with severity of primary Sjögren syndrome dry eye. Cornea 2015;34:248–52.

46 Emamian ES, Leon JM, Lessard CJ *et al.* Peripheral blood gene expression profiling in Sjögren's syndrome. Genes Immun 2009;10:285–96.

47 Der SD, Zhou A, Williams BR, Silverman RH. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. Proc Natl Acad Sci USA 1998;95:15623–8.

48 Karzai W, Oberhoffer M, Meier-Hellmann A, Reinhart K. Procalcitonin—a new indicator of the systemic response to severe infections. Infection 1997;25:329–34.

49 Zheng X, Luo Y, Li Q *et al.* Two gene set variation index as biomarker of bacterial and fungal sepsis. Biomed Res Int 2020;2020:1–9.

50 Liu F, Li L, Xu M *et al.* Prognostic value of interleukin-6, C-reactive protein, and procalcitonin in patients with COVID-19. J Clin Virol 2020;127:104370.

51 Mitrofan CG, Appleby SL, Nash GB *et al.* Bone morphogenetic protein 9 (BMP9) and BMP10 enhance tumor necrosis factor-alpha–induced monocyte recruitment to the vascular endothelium mainly via activin receptor–like kinase 2. J Biol Chem 2017;292: 13714–26.

52 Lv Z, Yang D, Li J *et al.* Bone morphogenetic protein 9 overexpression reduces osteosarcoma cell migration and invasion. Mol Cells 2013;36:119–26.

53 Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. Nat Rev Cancer 2002;2:161–74.

54 Hiscott J, Beauparlant P, Crepieux P *et al.* Cellular and viral protein interactions regulating I kappa B alpha activity during human retrovirus infection. J Leukoc Biol 1997;62:82–92.

55 Willerford DM, Chen J, Ferry JA *et al.* Interleukin-2 receptor α chain regulates the size and content of the peripheral lymphoid compartment. Immunity 1995;3: 521–30.

56 Blesch A, Bosserhoff AK, Apfel R *et al.* Cloning of a novel malignant melanoma-derived growth-regulatory protein, MIA. Cancer Res 1994;54:5695–701.

57 Kolanczyk M, Mautner V, Kossler N *et al.* MIA is a potential biomarker for tumour load in neurofibromatosis type 1. BMC Med 2011;9:82.

58 Fagerberg L, Hallstrom BM, Oksvold P *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics 2014;13:397–406.

59 Wang R, Sun PD. Natural killer cell–mediated shedding of ULBP2. PLoS One 2014;9:e91133.

60 Gannage M, Buzyn A, Bogiatzi SI *et al.* Induction of NKG2D ligands by gamma radiation and tumor necrosis factor-alpha may participate in the tissue damage during

acute graft-versus-host disease. Transplantation 2008; 85:911–5.

61 Kubin M, Cassiano L, Chalupny J *et al.* ULBP1, 2, 3: novel MHC class I-related molecules that bind to human cytomegalovirus glycoprotein UL16, activate NK cells. Eur J Immunol 2001;31:1428–37.

62 Cosman D, Mullberg J, Sutherland CL *et al.* ULBPs, novel MHC class I–related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. Immunity 2001;14:123–33.

63 Jachimowicz RD, Fracasso G, Yazaki PJ *et al.* Induction of *in vitro* and *in vivo* NK cell cytotoxicity using high-avidity immunoligands targeting prostate-specific membrane antigen in prostate carcinoma. Mol Cancer Ther 2011;10:1036–45.

64 Delaleu N, Mydel P, Kwee I *et al.* High fidelity between saliva proteomics and the biologic state of salivary glands defines biomarker signatures for primary Sjögren's syndrome. Arthritis Rheumatol 2015;67: 1084–95.

65 Hu S, Gao K, Pollard R *et al.* Preclinical validation of salivary biomarkers for primary Sjögren's syndrome. Arthritis Care Res (Hoboken) 2010; 62:1633–8.