RHEUMATOLOGY

Original article

Are interferon-related biomarkers advantageous for monitoring disease activity in systemic lupus erythematosus? A longitudinal benchmark study

Thomas Rose[1], Andreas Grützkau[2], Jens Klotsche[2], Philipp Enghard[2], Alexandra Flechsig[1], Johannes Keller[3], Gabriela Riemekasten[4], Andreas Radbruch[2], Gerd-Rüdiger Burmester[1], Thomas Dörner[1], Falk Hiepe[1] and Robert Biesen[1]

Abstract

Objective. To determine the clinical value of six traditional and three IFN-related biomarkers in monitoring disease activity (DA) in SLE.

Methods. Prospective longitudinal study of IFNα, IFNγ-inducible protein 10 (IP-10) and sialic acid-binding Ig-like lectin 1 (SIGLEC1) *vs* antibodies against dsDNA (ELISA and Farr radioimmunoassay), dsDNA-complexed nucleosomes (anti-dsDNA-NcX: ELISA), nucleosomes (ANuA: ELISA) and complement C3/C4 for correlation with DA (measured by BILAG 2004 index) in 26 SLE patients (77 visits). Optimal upper and lower longitudinal thresholds for the biomarkers and their accuracies for reflecting clinically relevant changes in DA (flares and remission) were determined by receiver operating characteristic and Youden index analysis.

Results. Increases in IP-10, SIGLEC1 and ANuA to + 101.6 pg/ml, +5.01 relative mean fluorescence intensity and +16.20 IU/ml above the calculated upper longitudinal threshold significantly reflected lupus flares, with a sensitivity and specificity of 50 and 95% for IP-10, 83 and 90% for SIGLEC1 and 58 and 95% for ANuA. Decreases in anti-dsDNA (ELISA), IFNα and anti-dsDNA (Farr assay) to − 64.7 IU/ml, −16.69 pg/ml and −3.3 IU/ml below lower longitudinal thresholds, respectively, best reflected remission, with sensitivity and specificity of 75 and 95%, 62 and 90%, and 75 and 90%, respectively.

Conclusion. IP-10, SIGLEC1 and ANuA emerged as advantageous biomarkers for monitoring disease activity. This is the first study in SLE that provides longitudinal biomarker thresholds and test accuracies for SLE flares and remitting disease. In the context of IFN-directed therapies, chemokines and fluorescence-activated cell sorting-based IFN biomarkers for monitoring SLE activity should be further studied.

Key words: Systemic lupus erythematosus, longitudinal, biomarker, interferon, IP-10, sialic acid-binding Ig-like lectin 1, autoimmunity

BASIC SCIENCE

Rheumatology key messages

- IFNα and its response proteins IP-10 and SIGLEC1 correlate longitudinally with SLE disease activity.
- Longitudinal biomarker thresholds and test accuracies were calculated for SLE.
- Biomarker analysis revealed that IP-10 and SIGLEC1 could be useful for monitoring lupus activity.

---

[1]Department of Rheumatology and Clinical Immunology, Charité Universitätsmedizin Berlin, [2]German Rheumatism Research Center Berlin-Leibniz Institute, [3]Center for Musculoskeletal Surgery, Charité Universitätsmedizin Berlin, Berlin and [4]Department of Rheumatology, University of Lübeck, Lübeck, Germany

Correspondence to: Falk Hiepe, Department of Rheumatology and Clinical Immunology, Charité University Hospital, Charitéplatz 1; D-10117 Berlin, Germany.
E-mail: Falk.Hiepe@charite.de

Introduction

Expert knowledge and experience as well as accurate disease activity (DA) indices and biomarkers are essential for monitoring DA and improving clinical care in SLE patients. The consumption of complement components such as C3 and C4 and the appearance of antibodies against anti-dsDNA have represented the gold standard

biomarkers of SLE for more than five decades. However, their use regarding longitudinal assessment is limited [1]. To our knowledge, only the IFNγ-inducible protein 10 chemokine (IP-10), erythrocyte-bound C3d/C4d and anti-nucleosome antibodies (ANuA) have been investigated longitudinally and compared with the gold standard laboratory tests for lupus [2–5]. While IP-10 outperformed complement factors and anti-dsDNA antibodies in two independent and large longitudinal studies [2, 3], erythrocyte-bound C3d and C4d did not exceed the gold standards [5]. In 2011, we reported that ANuA and dsDNA-complexed nucleosomes (dsDNA-NcX) were more specific for SLE than anti-dsDNA antibodies and detected changes in DA over time better than complement factors or anti-dsDNA antibodies [4].

IFNα and its surrogate markers are cross-sectionally and longitudinally correlated to DA in SLE [3, 6–10]. In contrast, the IFN type I induced gene signature in whole blood did not correlate to longitudinal changes in DA in SLE patients [11, 12]. In 1979, Hooks *et al.* [6] reported a correlation between elevated IFN levels and DA in SLE; their study of samples from nine SLE patients demonstrated a longitudinal correlation between antiviral effects (indicating IFN bioactivity) and DA. In 2000, Bengtsson *et al.* [10] showed that IFNα levels in serial samples from 30 SLE patients, as measured by a dissociation-enhanced lanthanide fluorescence immunoassay (DELFIA), increased during flares [10]. Two recent studies demonstrated that IP-10, inducible by IFNα, correlated longitudinally with lupus activity [2, 3]. However, two independent studies analysing the whole blood IFN signature did not find any such longitudinal correlation with SLE activity [11, 12]. Landolt-Marticorena *et al.* [11] described a lack of association between the IFN signature and longitudinal changes in SLEDAI scores in 27 SLE patients. Petri *et al.* [12] confirmed this in a longitudinal analysis of IFN response scores over consecutive pairs of visits in different subsets of SLE patients.

Although many studies provided evidence of longitudinal correlations of biomarkers with DA, they did not affect clinical decision-making or provide the level of evidence needed to include these biomarkers as endpoints in clinical trials. To determine the usefulness of a biomarker to monitor lupus activity, the following clinical questions have to be answered: what is the optimal threshold for a longitudinal biomarker to reflect clinically relevant changes in DA? How sensitive is the biomarker of interest for detecting a flare or remitting disease? How predictive is a significant change in a biomarker for a substantial change in SLE activity?

In addition to reliable laboratory tests, longitudinal test accuracy data are needed to answer these questions. So far, these issues have only been addressed in a single study including 14 patients with lupus flares [13]. Knowledge of the longitudinal test accuracies of biomarkers enables their benchmarking and improves their interpretation; this is necessary to guide decision-making based on clinically relevant changes.

Recently, we compared the ability of IFNα and the two dependent proteins IP-10 and sialic acid-binding Ig-like lectin 1 (SIGLEC1) *vs* the traditional biomarkers anti-dsDNA-NcX and ANuA to detect and predict DA in SLE [9, 14]. Here, we report on a longitudinal extension of this previous study designed to benchmark IFN-related biomarkers (IFNα, IP-10, SIGLEC1) in comparison with standard laboratory tests for lupus (anti-dsDNA by ELISA and Farr radioimmunoassay, C3 and C4) and other assays (ANuA and anti-dsDNA-NcX by ELISA).

Methods

Study participants and detection of biomarkers

Twenty-six SLE patients (23 female and 3 male, 25 Caucasian and 1 Asian, mean age 38 years, age range 21–69 years) meeting at least 4 of the 11 components of the ACR criteria for SLE [15] in combination with a previous or present history of positive ANuA tests were prospectively recruited at the Department of Rheumatology and Clinical Immunology of the Charité University Hospital Berlin. Seventy-seven visits by these 26 patients were included in the analysis. Written informed consent according to the Declaration of Helsinki was obtained from all participants. The ethics committee of the Charité University Hospital Berlin approved the study. General information about the patients and their medication is provided in supplementary Table S1, available at *Rheumatology* Online. Patients with an increase in prednisolone dose within the last 4 weeks because of inhibitory effects of IFN were excluded from the analysis. Patients receiving dialysis therapy (n = 1) as well as those with psychiatric symptoms (n = 1) or infections (n = 7) were also excluded. SLE patients were seen 3-monthly at their regular followup visits or earlier if disease was worsening. There were no fixed time points between the visits. Lupus DA was assessed throughout the study by the same physician. BILAG 2004 categories were converted into numerical scores as follows: A = 12, B = 5, C = 1, D = 0, E = 0 [16, 17]. The individual scores were then summed to yield an overall DA score. DA was also assessed by SLEDAI [18].

Biomarker detection methods

All sera were stored at −80 °C. IP-10 levels were analysed by using a Quantikine ELISA (R&D Systems, Minneapolis, MN, USA) based on anti-human IP-10 antibodies and were processed according to the manufacturer's instructions. Miltenyi Biotec (Bergisch Gladbach, Germany) analysed serum IFNα levels by using a time-resolved, fluorimetry-based DELFIA non-commercial immunoassay. Auto-antibodies were determined with commercially available tests by Euroimmun (Lübeck, Germany) (anti-dsDNA ELISA, Farr assay, anti-dsDNA-NcX ELISA, anti-Nuc ELISA). All measurements, except for C3 and C4, were performed in duplicate.

Flow cytometric detection of SIGLEC1 on monocytes

Whole blood samples, of 90 μl, were stained with anti-CD14-FITC (Miltenyi Biotec) and anti-CD169 RPE (clone 7-329) (AbD Serotec, Oxford, UK) mAb for flow cytometric analysis for 15 min on ice. As a control, the same staining omitting anti-CD169 antibody was performed. Erythrocytes were lysed on ice for 15 min by using fluorescence-activated cell sorting (FACS) Lysing Solution (Becton Dickinson, Heidelberg, Germany). Cells were washed twice using PBS containing 2% BSA and 5 mmol/l EDTA. Cells were fixed using Cellfix (Becton Dickinson) and analysis was done within 24 h using an LSR II cytometer (Becton Dickinson) with DIVA software 6.0 (Becton Dickinson). Relative median fluorescence intensity (rMFI) for SIGLEC-1 was calculated by dividing the SIGLEC1 expression on CD14$^+$ monocytes by the intensity of negative controls.

Longitudinal biomarker threshold analysis for classification of changes in DA in SLE patients

The clinical relevance of changes in biomarker levels in lupus is unknown. In order to use a biomarker to prospectively identify clinically relevant changes in DA, reliable upper and lower thresholds for biomarkers are needed. The results of the present analysis could be applied to biomarkers that are assessed during a fluctuating or therapy-modified disease (e.g. tumour markers in cancer).
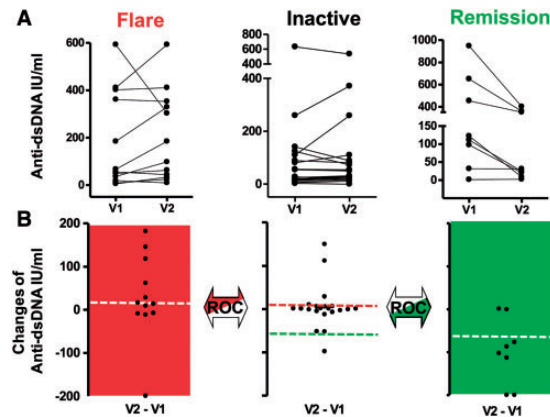
In this study, between-visit changes in lupus activity were categorized into three groups: increasing (flare), stable (inactive) or decreasing (remitting) (Fig. 1A). The first group consisted of 12 consecutive pairs of visits showing increasing DA (flare), which was defined as a new BILAG A or B score [19] and no improvement by A or B between visits. The second group consisted of 21 pairs of visits showing stable (inactive) disease, which was defined by two consecutive visits with a C, D or E but no A or B in the BILAG systems. The third group included eight pairs of visits characterized by decreasing (remitting) DA between visits, which was defined by an improvement of a prior documented A or B in the BILAG systems but no new A or B. Consecutive visits (n = 10) that did not meet the criteria of flare, remitting or inactive disease were excluded.

The change (Δ) in biomarker levels (BLs) between each pair of consecutive visits was calculated for all biomarkers and patients as follows: $\Delta BL = BL_{visit2} - BL_{visit1}$ (Fig. 1B). Receiver operating characteristic (ROC) analysis for upper threshold determination for each studied biomarker compared differences in ΔBL between the groups Flare and Stable (Inactive). ROC analysis for determination of the respective lower thresholds included the Stable (Inactive) and Remission groups. The upper and lower thresholds were defined as the highest Youden index values (sensitivity + specificity − 1). The Youden index is a well-investigated summary measure for selecting optimum cut-off points in statistics [20].

Statistical analysis

Statistical analyses were performed with GraphPad Prism 6.0 (GraphPad, La Jolla, CA, USA) and STATA 12 software

Fig. 1 Longitudinal biomarker analysis

(**A**) Changes in lupus activity between two consecutive visits were categorized as worsening (flare), stable disease (inactive) or improvement (remission). Representative plots for anti-dsDNA between visits (V) 1 and 2 are shown. (**B**) The dots for serial visits were connected by a line, and differences in biomarker changes were calculated to yield the dot plot. ROC analysis for determination of upper thresholds compared differences in biomarker changes between the Flare and Inactive groups (red). ROC analysis for determination of the lower threshold included the Inactive and Remission groups (green).

(StataCorp., College Station, TX, USA). The change (delta, Δ) in each parameter (P), including biomarkers, laboratory values, BILAG-2004 scores, BILAG-2004 systems and mSLEDAI scores, between all included serial visits of each SLE patient were calculated as follows: $\Delta P_{[n-1]} = parameter\ visit_{[n]} - parameter\ visit_{[n-1]}$. Analyses were done by using a linear mixed model (see supplementary Fig. S1, available at *Rheumatology* Online) and an analytically weighted linear regression model with weighting for the number of included visits between Δbiomarker and Δparameter to take into account the different numbers of consecutive visits per patient.

The definition of active disease characterized by an A or B category in BILAG is recommended [21]. The Youden index provides a widely accepted method for estimating thresholds. The threshold associated with the Youden index corresponds to the point with the maximum risk difference in the outcome (e.g. flare) and the dichotomized biomarker [22]. This method results in reliable threshold estimates in small sample sizes [22]. Bias-corrected confidence intervals for the area under the curve (AUC) in receiver operating characteristics were estimated by bootstrapping based on 1000 bootstrap replications [23]. The one-tailed significance test in ROC analyses achieves a statistical power of > 80% when assuming an area under curve of at least 0.7, and a balanced sample size in the two groups of patients with the biomarker below or above the selected limit. The statistical power decreases to 65% in case of an unbalanced sample size assuming 20% of the patients in the group with the biomarker above the limit.

Results

IFNα and SIGLEC1 reflect longitudinal changes in DA

The clinical course of a lupus patient is characterized by cycles of stable, relapsing and remitting disease, which often result in changes in lupus medication. This might ultimately affect the type I IFN system, represented here by serum levels of IFNα and SIGLEC1 expression levels on monocytes. Therefore, we plotted these parameters against the BILAG-2004 and medication characteristics of nine lupus patients over time (total of 51 visits) to test for relationships (supplementary Fig. S2, available at *Rheumatology* Online).

Statistical analysis of changes in biomarker levels *vs* changes in SLE activity

Seventy-seven total visits by 26 patients led to 51 delta visits (consecutive visits). Changes in BILAG-2004 scores were plotted against changes in all nine selected biomarkers (supplementary Fig. S3, available at *Rheumatology* Online). All measured IFN biomarkers—IP-10 (r = 0.50, P = 0.0002), SIGLEC1 (r = 0.45, P = 0.001) and IFNα (r = 0.42, P = 0.0022)—correlated with DA scores (BILAG-2004), while anti-dsDNA (ELISA)—the standard-of-care biomarker—was only weakly correlated to DA over time (r = 0.29, P = 0.0424). Anti-dsDNA by Farr assay (P = 0.5188), C3 (P = 0.0507) and C4 (P = 0.3895) did not show correlations related to lupus activity. Anti-dsDNA-NcX (r = 0.63, P < 10^{-4}) and ANuA (r = 0.63, P < 10^{-4}) were the biomarkers with the highest correlation coefficients reflecting changes in DA.

Correlation analyses of clinical and laboratory parameters over time

To reveal further longitudinal laboratory and clinical associations, all nine biomarkers were statistically analysed for correlation to various clinical parameters (Fig. 2). Increased anti-dsDNA-NcX and ANuA titres were predominantly associated with decreased neutrophil counts and correlated significantly with all IFN-related parameters. Decreasing thrombocyte counts were mostly associated with increasing IP-10 and IFNα levels over time. Furthermore, we analysed all nine biomarkers for a relationship to a specific organ system involvement. Therefore, only those SLE patients with a known abnormality in the studied BILAG system were included (patients with an E in the affected BILAG system were excluded). Changes in neurological complaints showed a high correlation with changes in IFNα levels (r = 0.79, P < 0.0001). Apart from IP-10, C3 and C4, changes in all other biomarkers investigated were significantly related to changes in renal DA.

Biomarker benchmarking for monitoring SLE activity

Statistical correlation analysis of changes in biomarker levels between two consecutive visits, as described in the previous section (see Fig. 2), is a simple and frequently used method for studying biomarkers in SLE. In clinical practice, however, correlation analysis of biomarkers is difficult because vital longitudinal diagnostic test accuracy data on parameters such as sensitivity and specificity are not available so far. Therefore, we performed ROC analysis to test for differences between the flare, stable (inactive) and remitting disease groups in terms of the change in biomarker levels (ΔBL) in these disease states (cf. Fig. 1A). The Youden index was used to identify optimal lower and upper thresholds between these groups. These thresholds were then used to calculate the sensitivities and specificities of the respective biomarkers. The results for all nine biomarkers are presented in Table 1.

For discrimination between flare and stable disease, IP-10, SIGLEC1 and ANuA achieved significant P-values in ROC analysis. Increasing IP-10 and SIGLEC1 discriminated best between the two groups: the AUC for these biomarkers was 0.75 and 0.78, respectively. Figure 3 shows the sensitivities of all nine longitudinal biomarkers for discriminating between the three disease states based on the calculated thresholds. Regarding the detection of disease flares, an increase in SIGLEC1 expression by >5.01 rMFI showed a sensitivity of 83%, while an increase in IP-10 by >101.6 pg/ml had a sensitivity of 50%. The gold standard biomarkers did not reach the level of significance in ROC analysis, as reflected by low AUC values. C3 showed high but clinically irrelevant fluctuation in stable disease, leading to low overall test performance.

Regarding discrimination between remitting *vs* stable (inactive) disease, ROC analysis led to significant P-values for IFN- and anti-dsDNA-related biomarkers. IFNα and anti-dsDNA (ELISA) achieved high AUC values (>0.8). The highest sensitivity for remitting disease was obtained with the calculated cutoff points (lower thresholds) for SIGLEC1 and anti-dsDNA (by Farr assay and ELISA), with a sensitivity of 75%.

Furthermore, we wanted to find out if it is possible to utilize the biomarker kinetics data (increasing/decreasing biomarker levels) to discriminate between the disease states (flare/remission) based on the specific biomarker thresholds. Therefore, we calculated the positive predictive value (PPV) for each biomarker. The results are shown in Table 2.

IFNα elevation above the calculated threshold was found to be highly predictive of flares, with a PPV of 100% in our lupus cohort. In contrast, below-threshold decreases in anti-dsDNA-NcX antibodies were 100% predictive for remission (cf. Table 2). SIGLEC1 predicted both disease states well (71 and 86%, respectively). Stable biomarker levels predicted stable (inactive) disease with a PPV of 90% for SIGLEC1, followed by 76% for IFNα and 76% for anti-dsDNA (ELISA). C3 and C4 were the biomarkers with the lowest PPV (60 and 62%, respectively).

Discussion

Few biomarkers are used in lupus practice and not all provide sufficient reliability for clinical decision-making [1]. While almost all longitudinal studies show correlations between DA and the investigated biomarkers, physicians need data on the accuracy of these tests for useful clinical interpretation. Valuable longitudinal biomarker information

Fig. 2 Laboratory and clinical significance of changing biomarker levels

			No. of visits	Statistics	Δ IFNα	Δ IP-10	Δ SIGLEC1	Δ Farr assay	Δ Anti-dsDNA-NcX	Δ Anti-dsDNA	Δ Anti-nucleosome	Δ Complement C3	Δ Complement C4
All patients (n=26)	Δ Blood values	Haemoglobin	77	r	-0.1072	-0.1935	-0.1678	-0.0316	-0.3641	-0.044	-0.3754	0.1451	-0.0442
				CI	-0.401 ; -0.043	-0.271 ; 0.104	-0.245 ; 0.132	-0.341 ; 0.028	-0.482 ; -0.142	-0.294 ; 0.079	-0.555 ; -0.238	0.143 ; 0.482	-0.229 ; 0.148
				p	0.4542;	0.1736	0.2393	0.8257	0.0086	0.759	0.0066	0.3098	0.7579
		Leucocytes	77	r	-0.2223	-0.0896	-0.2967	-0.2071	-0.5003	-0.2923	-0.4591	0.2245	0.2026
				CI	-0.306 ; 0.066	-0.391 ; -0.031	-0.419 ; -0.064	-0.305 ; 0.068	-0.575 ; -0.266	-0.357 ; 0.009	-0.573 ; -0.263	0.026 ; 0.387	-0.200 ; 0.178
				p	0.1169	0.532	0.0345	0.1449	0.0002	0.0374	0.0007	0.1133	0.1539
		Lymphocytes	77	r	0.1315	-0.0067	0.1757	0.1412	0.075	0.0826	0.0823	0.0923	0.0085
				CI	-0.063 ; 0.309	-0.285 ; 0.089	-0.295 ; 0.078	0.014 ; 0.377	-0.268 ; 0.108	-0.145 ; 0.232	-0.212 ; 0.166	-0.242 ; 0.134	-0.272 ; 0.103
				p	0.3575	0.963	0.2176	0.3229	0.6008	0.5643	0.566	0.5195	0.9527
		Monocytes	77	r	0.0104	0.1543	-0.049	-0.1433	-0.1091	-0.1218	-0.1266	0.0306	-0.0833
				CI	-0.182 ; 0.196	-0.140 ; 0.237	-0.399 ; -0.040	-0.312 ; 0.059	-0.442 ; -0.092	-0.432 ; -0.079	-0.428 ; -0.075	-0.156 ; 0.221	-0.239 ; 0.137
				p	0.9422	0.2796	0.7328	0.3158	0.4458	0.3946	0.3762	0.8314	0.5612
		Neutrophils	77	r	-0.2447	-0.109	-0.3183	-0.2108	-0.4987	-0.2898	-0.4563	0.2065	0.2126
				CI	-0.324 ; 0.047	-0.381 ; -0.018	-0.386 ; -0.025	-0.321 ; 0.050	-0.544 ; -0.223	-0.342 ; 0.026	-0.550 ; -0.232	0.030 ; 0.391	-0.184 ; 0.194
				p	0.0835	0.4466	0.0228	0.1377	0.0002	0.0392	0.0008	0.146	0.1342
		Thrombocytes	77	r	-0.3505	-0.4638	-0.2343	-0.0119	-0.1585	0.106	-0.264	0.3563	0.1558
				CI	-0.502 ; -0.168	-0.300 ; 0.073	-0.353 ; 0.014	-0.368 ; -0.004	-0.367 ; 0.002	-0.340 ; 0.029	-0.443 ; -0.093	-0.209 ; 0.169	-0.189 ; 0.189
				p	0.0117	0.0006	0.098	0.9338	0.2666	0.4591	0.0612	0.0103	0.2748
		C- reactive protein	77	r	-0.0211	0.0329	-0.3541	0.0127	-0.0584	-0.1188	0.017	0.0953	0.0125
				CI	-0.182 ; 0.195	-0.156 ; 0.221	-0.466 ; -0.121	-0.157 ; 0.220	-0.175 ; 0.240	-0.218 ; 0.159	-0.141 ; 0.236	0.019 ; 0.381	-0.307 ; 0.066
				p	0.8832	0.8186	0.0108	0.9296	0.6839	0.4065	0.9059	0.506	0.9309
	Δ IFN	IFNα	77	r	-	0.4949	0.3307	0.4911	0.6108	0.5609	0.6424	-0.3333	-0.2822
				CI		0.048 ; 0.406	0.088 ; 0.439	0.525 ; 0.747	0.473 ; 0.715	0.497 ; 0.730	0.562 ; 0.769	-0.479 ; -0.138	-0.436 ; -0.084
				p		0.0002	0.0178	0.0003	<0.0001	<0.0001	<0.0001	0.0168	0.0449
		IP-10	77	r	0.4949	-	0.2414	-0.0438	0.4892	-0.0374	0.5703	-0.1551	0.0715
				CI	0.048 ; 0.406		0.050 ; 0.407	-0.378 ; -0.016	0.300 ; 0.599	-0.382 ; -0.020	0.302 ; 0.601	-0.257 ; 0.119	-0.049 ; 0.322
				p	0.0002		0.0879	0.7603	0.0003	0.7943	<0.0001	0.2773	0.6182
		SIGLEC1	77	r	0.3307	0.2414	-	0.2794	0.4915	0.3404	0.4827	-0.1768	-0.1309
				CI	0.088 ; 0.439	0.050 ; 0.407		-0.010 ; 0.356	0.305 ; 0.602	0.104 ; 0.452	0.278 ; 0.584	-0.340 ; 0.029	-0.299 ; 0.074
				p	0.0178	0.0879		0.0471	0.0002	0.0145	0.0003	0.2146	0.3598
	Δ DAI	MSLEDAI-2k	77	r	0.4479	0.7597	0.4636	0.1021	0.5534	0.0305	0.5887	-0.1731	0.0303
				CI	0.096 ; 0.445	0.637 ; 0.813	0.186 ; 0.516	-0.203 ; 0.175	0.335 ; 0.624	-0.262 ; 0.113	0.383 ; 0.656	-0.424 ; -0.069	-0.166 ; 0.212
				p	0.001	<0.0001	0.0006	0.4758	<0.0001	0.832	<0.0001	0.2244	0.8331
		BILAG-2004 12 5 1	77	r	0.42	0.5044	0.4483	0.0924	0.6342	0.2853	0.6338	-0.2751	-0.1231
				CI	0.112 ; 0.458	0.337 ; 0.625	0.323 ; 0.616	-0.203 ; 0.175	0.449 ; 0.699	-0.027 ; 0.342	0.460 ; 0.706	-0.447 ; -0.098	-0.249 ; 0.127
				p	0.0022	0.0002	0.001	0.5188	<0.0001	0.0424	<0.0001	0.0507	0.3895
Systems of BILAG2004		Constitutional	57	r	0.3722	0.5047	0.5394	0.0556	0.5978	0.1369	0.6239	-0.1735	0.0459
				CI	0.124 ; 0.468	0.404 ; 0.670	0.229 ; 0.548	-0.180 ; 0.198	0.414 ; 0.676	-0.126 ; 0.251	0.427 ; 0.685	-0.414 ; -0.057	-0.219 ; 0.159
				p	0.0092	0.0003	0.0001	0.7075	<0.0001	0.3534	<0.0001	0.2381	0.7567
		Mucocutaneous	57	r	0.1878	0.1368	0.247	-0.0063	0.4164	0.1885	0.3785	-0.1205	0.0256
				CI	0.065 ; 0.420	-0.114 ; 0.262	0.080 ; 0.432	-0.099 ; 0.276	0.239 ; 0.555	0.075 ; 0.428	0.265 ; 0.574	-0.417 ; -0.062	-0.308 ; 0.065
				p	0.2011	0.3539	0.0906	0.966	0.0032	0.1995	0.008	0.4145	0.8629
		Neurological	43	r	0.7917	0.5199	0.254	-0.0064	0.5257	0.4917	0.5122	-0.3257	-0.3068
				CI	0.090 ; 0.441	-0.048 ; 0.322	-0.115 ; 0.261	-0.203 ; 0.175	-0.023 ; 0.345	-0.085 ; 0.289	-0.039 ; 0.331	-0.305 ; 0.067	-0.388 ; -0.027
				p	<0.0001	0.001	0.1292	0.97	0.0008	0.002	0.0012	0.0491	0.0648
		Musculosceletal	52	r	0.4919	0.4309	0.1062	-0.0849	0.5415	0.2709	0.6273	-0.2231	0.0383
				CI	0.060 ; 0.416	0.129 ; 0.471	-0.029 ; 0.340	-0.284 ; 0.090	0.158 ; 0.494	-0.137 ; 0.239	0.231 ; 0.549	-0.371 ; -0.007	-0.095 ; 0.280
				p	0.0007	0.0035	0.4927	0.5839	0.0001	0.0753	<0.0001	0.1455	0.8049
		Cardiorespiratory	38	r	-0.0144	0.4991	0.421	-0.3674	0.383	-0.306	0.4287	0.0094	0.1822
				CI	-0.245 ; 0.132	0.411 ; 0.675	0.071 ; 0.425	-0.342 ; 0.026	-0.002 ; 0.363	-0.397 ; -0.038	-0.006 ; 0.360	-0.212 ; 0.165	-0.080 ; 0.293
				p	0.9366	0.0031	0.0147	0.0354	0.0278	0.0833	0.0128	0.9585	0.3103
		Renal	49	r	0.5271	0.2016	0.4942	0.7075	0.5123	0.5645	0.5025	-0.3044	-0.2673
				CI	0.127 ; 0.470	0.133 ; 0.475	0.206 ; 0.531	0.170 ; 0.504	0.209 ; 0.534	0.059 ; 0.415	0.177 ; 0.509	-0.307 ; 0.065	-0.252 ; 0.124
				p	0.0003	0.2005	0.0009	<0.0001	0.0005	0.0001	0.0007	0.05	0.087
		Hematology	54	r	-0.0716	-0.0662	0.3318	0.0556	0.3058	0.1122	0.1898	-0.2033	-0.166
				CI	-0.412 ; -0.055	0.229 ; 0.548	0.072 ; 0.425	-0.412 ; -0.056	0.035 ; 0.395	-0.380 ; -0.018	-0.060 ; 0.311	-0.253 ; 0.123	-0.103 ; 0.272
				p	0.6325	0.6586	0.0227	0.7104	0.0366	0.4526	0.2013	0.1705	0.2649

Dependence and significance	negative	p<0.0005	p<0.005	p<0.05		p<0.05	p<0.005	p<0.0005	positive

All nine biomarkers were investigated with regard to their statistically significant relationships to whole blood values, their relationship to IFN biomarkers, overall disease activity and their manifestations in the BILAG-2004 systems. Weighted correlation coefficients in linear regression were calculated. In the category ('systems'), changes in the BILAG-2004 system (A, B, C and D score) were correlated to changes in biomarkers. The colour intensity indicates the statistical significance levels, with orange values indicating positive and blue representing negative correlations with respect to particular biomarker levels. DAI: disease activity index; mSLEDAI-2K: modified SLE disease activity index 2000; P: P-value; r: correlation coefficient.

includes the sensitivity and specificity of a given change in biomarker levels to reflect flares or remitting lupus activity. This study is the very first to provide longitudinal biomarker-specific thresholds and longitudinal diagnostic test accuracy data for SLE flares and remission. Therefore, this detailed benchmark analysis provides clinically relevant data.

First, we demonstrated the correlation of changes in IFNα and SIGLEC1 expression with changes in DA and medication in nine lupus patients. To substantiate our findings, we used a simple and frequently used correlation analysis to test for correlations between our set of biomarkers and changes in DA. Anti-dsDNA-NcX, ANuA and IP-10 showed the best correlation with longitudinal

TABLE 1 Benchmark analysis of nine biomarkers

Unit	Δ IFNα pg/ml	Δ IP-10 pg/ml	Δ SIGLEC1 rMFI	Δ Farr assay IU/ml	Δ Anti-dsDNA-NcX iU/ml	Δ Anti-dsDNA IU/ml	Δ Anti-nucleosome IU/ml	Δ Complement C3 mg/dl	Δ Complement C4 mg/dl
Flare (new A or B in BILAG)									
Limit (YI)	**17.32**	**101.60**	**5.01**	**7.45**	**36.27**	**7.67**	**16.20**	**−8.00**	**−1.50**
YI	0.42	0.45	**0.74**	0.54	0.44	0.48	0.54	0.30	0.11
AUC	0.56	0.75	0.78	0.66	0.71	0.65	0.72	0.70	0.61
95% CI by bootstrap	0.31, 0.81	0.55, 0.91	0.57, 0.96	0.41, 0.88	0.48, 0.89	0.40, 0.87	0.47, 0.91	0.51, 0.86	0.42, 0.76
P-value	0.600	**0.017**	**0.008**	0.121	0.052	0.166	**0.040**	0.054	0.312
Sensitivity	41.67	50.00	83.33	58.33	58.33	66.67	58.33	58.33	25.00
Specificity	100.00	95.24	90.48	95.24	85.71	80.95	95.24	71.43	85.71
Remitting disease (improving of A or B in BILAG)									
Limit (YI)	**−16.69**	**−346.70**	**−13.59**	**−3.30**	**−147.10**	**−64.70**	**−17.27**	**11.00**	**2.50**
YI	0.53	0.58	0.70	0.65	0.63	0.70	0.45	0.36	0.33
AUC	0.84	0.75	0.75	0.76	0.74	0.86	0.57	0.65	0.61
95% CI by bootstrap	0.52, 0.97	0.49, 0.97	0.42, 1.00	0.45, 0.97	0.45, 1.00	0.64, 1.00	0.24, 0.89	0.37, 0.88	0.29, 0.85
P-value	**0.005**	**0.040**	**0.040**	**0.036**	**0.045**	**0.003**	0.592	0.205	0.354
Sensitivity	62.50	62.50	75.00	75.00	62.50	75.00	50.00	50.00	37.50
Specificity	90.48	95.24	95.24	90.48	100.00	95.24	100.00	85.71	95.24

Clinically relevant changes in disease activity (flare/remission) were evaluated by receiver operating characteristic (ROC) and Youden index (YI) analysis to calculate optimal upper and lower thresholds and the respective test accuracies for the biomarkers. $P < 0.05$ was considered significant and is shown in bold. AUC: area under the ROC curver; rMFI: relative median fluorescence intensity.

TABLE 2 Relevance of changing biomarker levels

Predictive value Unit	Δ IFNα pg/ml	Δ IP-10 pg/ml	Δ SIGLEC1 rMFI	Δ Farr assay IU/ml	Δ Anti-dsDNA-NcX iU/ml	Δ Anti-dsDNA IU/ml	Δ Anti-nucleosome IU/ml	Δ Complement C3 mg/dl	Δ Complement C4 mg/dl
PPV of increasing, % BL (for flare)	100.00	75.00	71.43	77.78	70.00	66.67	77.78	50.00	37.50
PPV of stable, % BL (for stable diseae)	76.00	73.08	90.00	78.26	69.23	76.19	70.37	60.00	62.07
PPV of decreasing, % BL (for remission)	45.45	71.43	85.71	44.44	100.00	75.00	80.00	57.14	60.00

The positive predictive value (PPV) of changes in biomarker level (BL) for each disease state was calculated for all nine biomarkers. rMFI: relative median fluorescence intensity.

changes of disease indices, followed by SIGLEC1 and IFNα. Anti-dsDNA, C3 and C4 did not correlate longitudinally. This is in line with earlier reports. [2–4, 10, 24] Here, we showed for the first time that SIGLEC1 and lupus activity correlate longitudinally. These data provide additional evidence of a longitudinal association between the type I IFN system and lupus activity, as reflected by IFNα [6, 10], IP-10 [2, 3] and SIGLEC1, but contradict the findings of two studies using the whole blood IFN signature [11, 12].

It becomes more and more difficult to attribute symptoms to SLE activity in clinical assessment as the number of factors, such as organ damage, comorbidities (e.g. infections and psychological disorders) or drug side effects, increases. Valuable biomarkers that reflect DA with known longitudinal test accuracies might facilitate decision-making in these cases. Hereby, increasing levels of IP-10 and SIGLEC1 expression, followed by ANuA (ELISA), indicated flares longitudinally better than traditional biomarkers of lupus. Notably, increasing SIGLEC1 expression had the highest sensitivity for detecting flares (83%). Until now, longitudinal biomarker thresholds (e.g. ⩾25% change) have been defined arbitrarily [13, 25–28]. Only Hillebrandt *et al.* [13] investigated the longitudinal diagnostic test accuracies of changes in anti-dsDNA, as determined by Farr assay, for detecting lupus flares in 14 SLE patients. Using ⩾25% change in anti-dsDNA as a threshold, they calculated a sensitivity of 86% but with a specificity of only 13% for far assay and attributed the low test performance to the threshold [13]. Although not directly comparable, the use of Youden index and ROC analysis for threshold calculation in the present study resulted in improved test performance for the Farr assay (Table 1).

Remitting disease in response to treatment is determined by clinical and laboratory methods. Laboratory parameters indicating the success of treatment are evaluated regularly [29, 30], but rarely by longitudinal threshold analysis. Hui-Yuen *et al.* [28] defined a laboratory response as a ⩾ 25% increase in C3 or decrease in anti-dsDNA antibodies. Longitudinal test accuracies for disease remission have not been calculated so far. Our analysis showed that a remitting disease cycle was reflected significantly by most of our biomarkers except C3, C4 and ANuA. All IFN-related biomarkers, anti-dsDNA and anti-dsDNA-NcX indicated disease remission longitudinally, with sensitivities ranging from 62.5 to 75% and specificities of > 90%. Most notably, decreasing anti-dsDNA-NcX and SIGLEC1 levels were highly predictive for disease remission (PPV of 100 and 85.7%, respectively). Achieving a significant reduction of biomarker levels might serve as a criterion for treatment success that could be used as an end point in clinical trials or observational studies to objectify treatment response [30].

This study has certain limitations that should be considered when interpreting the results. First, as sensitivity and specificity are interrelated, optimal comparison of the sensitivities of all nine biomarkers at a set specificity level was hampered by the small number of patient visits

Fɪɢ. 3 Calculated thresholds for each biomarker indicating clinically relevant changes in disease activity
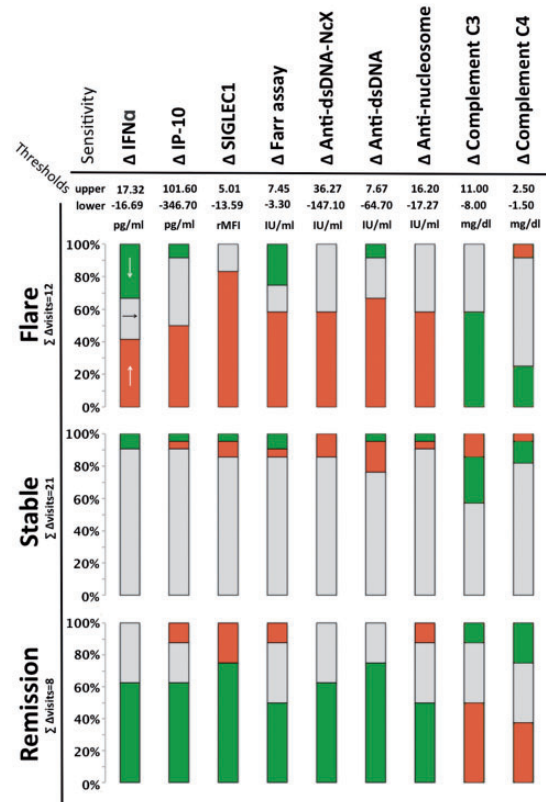
Colour coding of biomarker changes: red indicates biomarker increases above the upper threshold, green indicates a biomarker decreases below the lower threshold, and grey indicates biomarker changes within the threshold range.

available for analysis. We reported confidence intervals to show the degree of uncertainty in parameter estimates. Performing this analysis in a larger multicentre cohort study could overcome this more statistically relevant limitation. Second, we included SLE patients fulfilling the ACR criteria and who were positive for ANuA. These criteria were used earlier [25] and account for roughly two-thirds of SLE patients [31], which means the analysis might not be applicable to ANuA negative patients. Third, the use of relative MFI for the longitudinal assessment of SIGLEC1 expression could be further optimized by performing bead calibration, as done by Kikuchi-Taura *et al.* [32], to yield absolute numbers of SIGLEC1 molecules per monocyte. This would allow for more standardized measurement of SIGLEC1 expression independent of the cytometer used and its respective configuration.

In conclusion, this is the first study to benchmark biomarkers by calculating practical longitudinal thresholds and test accuracies that are needed to improve lupus diagnostics. We showed that IFN-related biomarkers and ANuA assays provide advantages over the current standard biomarkers. The extension of this analysis to a

larger validation cohort could increase the diagnostic value of these biomarkers and provide new possibilities for improving clinical care of SLE patients.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Mosca M, Tani C, Aringer M *et al*. European League Against Rheumatism recommendations for monitoring patients with systemic lupus erythematosus in clinical practice and in observational studies. Ann Rheum Dis 2010;69:1269–74.

2 Bauer JW, Petri M, Batliwalla FM *et al*. Interferon-regulated chemokines as biomarkers of systemic lupus erythematosus disease activity: a validation study. Arthritis Rheum 2009;60:3098–107.

3 Kong KO, Tan AW, Thong BYH *et al*. Enhanced expression of interferon-inducible protein-10 correlates with disease activity and clinical manifestations in systemic lupus erythematosus. Clin Exp Immunol 2009;156:134–40.

4 Biesen R, Dähnrich C, Rosemann A *et al*. Anti-dsDNA-NcX ELISA: dsDNA-loaded nucleosomes improve diagnosis and monitoring of disease activity in systemic lupus erythematosus. Arthritis Res Ther 2011;13:R26.

5 Kao AH, Navratil JS, Ruffing MJ *et al*. Erythrocyte C3d and C4d for monitoring disease activity in systemic lupus erythematosus. Arthritis Rheum 2010;62:837–44.

6 Hooks JJ, Moutsopoulos HM, Geis SA *et al*. Immune interferon in the circulation of patients with autoimmune disease. N Engl J Med 1979;301:5–8.

7 Kirou KA, Lee C, George S *et al*. Activation of the interferon-α pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. Arthritis Rheum 2005;52:1491–503.

8 Feng X, Wu H, Grossman JM *et al*. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. Arthritis Rheum 2006;54:2951–62.

9 Rose T, Grutzkau A, Hirseland H *et al*. IFN and its response proteins, IP-10 and SIGLEC-1, are biomarkers of disease activity in systemic lupus erythematosus. Ann Rheum Dis 2013;72:1639–45.

10 Bengtsson AA, Sturfelt G, Truedsson L *et al*. Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies. Lupus 2000;9:664–71.

11 Landolt-Marticorena C, Bonventi G, Lubovich A *et al*. Lack of association between the interferon-α signature and longitudinal changes in disease activity in systemic lupus erythematosus. Ann Rheum Dis 2009;68:1440–6.

12 Petri M, Singh S, Tesfasyone H *et al*. Longitudinal expression of type I interferon responsive genes in systemic lupus erythematosus. Lupus 2009;18:980–9.

13 Hillebrand JJG, Moens HB, Mulder AHL. Changes in Farr radioimmunoassay and EliA fluorescence immunoassay anti-dsDNA in relation to exacerbation of SLE. Lupus 2013;22:1169–73.

14 Biesen R, Demir C, Barkhudarova F *et al*. Sialic acid-binding Ig-like lectin 1 expression in inflammatory and resident monocytes is a potential biomarker for monitoring disease activity and success of therapy in systemic lupus erythematosus. Arthritis Rheum 2008;58:1136.

15 Tan EM, Cohen AS, Fries JF *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.

16 Isenberg DA, Rahman A, Allen E *et al*. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology 2005;44:902–6.

17 Cresswell L, Yee C-S, Farewell V *et al*. Numerical scoring for the Classic BILAG index. Rheumatology 2009;48:1548–52.

18 Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.

19 Gordon C. Definition and treatment of lupus flares measured by the BILAG index. Rheumatology 2003;42:1372–9.

20 Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J 2005;47:458–72.

21 Yee C-S, Isenberg DA, Prabu A *et al*. BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. Ann Rheum Dis 2008;67:873–6.

22 Klotsche J, Ferger D, Pieper L, Rehm J, Wittchen H-U. A novel nonparametric approach for estimating cut-offs in continuous risk indicators with application to diabetes epidemiology. BMC Med Res Methodol 2009;9:63.

23 Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000;19:1141–64.

24 Li T, Prokopec SD, Morrison S *et al*. Anti-nucleosome antibodies outperform traditional biomarkers as

longitudinal indicators of disease activity in systemic lupus erythematosus. Rheumatology 2015;54:449–57.

25 Ter Borg EJ, Horst G, Hummel EJ, Limburg PC, Kallenberg CGM. Measurement of increases in anti-double-stranded dna antibody levels as a predictor of disease exacerbation in systemic lupus erythematosus. Arthritis Rheum 1990;33:634–43.

26 ter Borg EJ, Horst G, Limburg PC, Kallenberg CG. Changes in plasma levels of interleukin-2 receptor in relation to disease exacerbations and levels of anti-dsDNA and complement in systemic lupus erythematosus. Clin Exp Immunol 1990;82:21–6.

27 Bootsma H, Spronk P, Borg E et al. The predictive value of fluctuations in IgM and IgG class anti-dsDNA antibodies for relapses in systemic lupus erythematosus. A prospective long term observation. Ann Rheum Dis 1997;56:661–6.

28 Hui-Yuen JS, Reddy A, Taylor J et al. Safety and efficacy of belimumab to treat systemic lupus erythematosus in academic clinical practices. J Rheumatol 2015;42:2288–95.

29 Vollenhoven RF, van, Petri MA, Cervera R, Roth DA et al. Belimumab in the treatment of systemic lupus erythematosus: high disease activity predictors of response. Ann Rheum Dis 2012;71:1343–9.

30 Alexander T, Sarfert R, Klotsche J et al. The proteasome inhibitior bortezomib depletes plasma cells and ameliorates clinical manifestations of refractory systemic lupus erythematosus. Ann Rheum Dis 2015;74:1474–8.

31 Bizzaro N, Villalta D, Giavarina D, Tozzoli R. Are anti-nucleosome antibodies a better diagnostic marker than anti-dsDNA antibodies for systemic lupus erythematosus? A systematic review and a study of metanalysis. Autoimmun Rev 2012;12:97–106.

32 Kikuchi-Taura A, Yura A, Tsuji S et al. Monocyte CD64 expression as a novel biomarker for the disease activity of systemic lupus erythematosus. Lupus 2015;24:1076–80.