ABSTRACT NUMBER 2183:

Evolution of Lymphoma Predictors in Primary Sjogren's Syndrome by Data Driven Analysis in Harmonized Patients

Andreas Goules[1], Loukas Chatzis[2], Vasileios pezoulas[3], Francesco Ferro[4], Valeria Manfrè[5], Dimitrios Fotiadis[3], Luca Quartuccio Quartuccio[6], Chiara Baldini[7], Salvatore De Vita[5] and Athanasios Tzioufas[1], [1]Dept. of Pathophysiology, School of Medicine, University of Athens, Athens, Greece, [2]Dept. of Pathophysiology, School of Medicine, University of Athens, Cholargos Athens, Greece, [3]Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, Ioannina, Greece, [4]Clinical and Experimental Medicine Department, Azienda Ospedaliero-Universitaria Pisana, Pisa, Italy, [5]Clinic of Rheumatology, Dept. of Medical and Biological Sciences, Udine University, Udine, Italy, [6]Rheumatology Division, Department of Medicine, University of Udine, Udine, Italy, [7]University of Pisa, Pisa, Italy

Background/Purpose: Lymphoma is the most serious complication of primary Sjogren's syndrome (pSS), occurring as a late sequel during disease course. So far, lymphoma predictors have been identified only at the time of pSS diagnosis in single center cohorts using classical logistic regression models. The current work aims to study the evolution of lymphoma predictors towards lymphoproliferation by employing data driven analysis on totally harmonized pSS patients.

Methods: Two curated and harmonized datasets with pSS patients from3 centers (Athens, Piza, Udine) were constructed to be analyzed for lymphoma predictors at two different time points before lymphoma diagnosis. Each dataset incorporated the same set of 30 clinically useful features covering the major aspects of the disease (Table 1, 2). The first dataset included 80 lymphoma patients and non-lymphoma controls (1:1 ratio), representing the time point of pSS diagnosis and the second dataset 68 lymphoma and non-lymphoma controls (1:1 ratio) representing the time point of 3-4 years before lymphoma diagnosis. All included patients fulfilled the 2016 ACR/EULAR criteria and non-lymphoma controls were matched according to age, gender, disease duration from pSS diagnosis to last follow up and treatment modalities. Systemic manifestations were defined as described in the ESSDAI domains and for those not included in the ESSDAI system, either by tissue biopsy or by applying international consensus criteria. A Fast-Correlation based feature selection (FCBF)/logistic regression (LR) model with lymphoma as an outcome, was applied on both datasets as described previously (1).

Results: Regarding the time point of pSS diagnosis, 5 prominent features in terms of magnitude of order were identified by the FCBF algorithm as potential predictors including rheumatoid Factor (RF), cryoglobulinemia, ACA pattern, ESSDAI≥5 and lacrimal gland enlargement of which only RF was finally proven an independent lymphoma predictor (Table 1). For the second time point of 3-4 years before lymphoma diagnosis, a similar set of

prominent features was found by the FCBF algorithm with the exception of lacrimal gland enlargement and cryoglobulinemia while RF and ESSDAI≥5 were identified as independent lymphoma risk factors (Table 2). Both FCBF/LR models had good overall performance after 10-fold cross validation strategy (pSS diagnosis: accuracy=63%, sensitivity=63%, specificity=63%, AUC=76% and for the second time point of 3-4 years before lymphoma diagnosis: accuracy=65%, sensitivity=65%, specificity=71%, AUC=75%).

Conclusion: Rheumatoid factor is the earliest and more persistent pSS associated lymphoma predictor while high ESSDAI≥5 in combination with RF connote an advanced stage across the lymphomagenesis process, predicting the occurrence of lymphoma in approximately 4 years before.

Table 1. FCBF-based multivariable logistic regression analysis for lymphoma predictors at pSS diagnosis*

Prominent feature*	Regression coefficient	Odds ratio	p-value	CI upper	CI low
Rheumatoid Factors**	**1.2**	**3.332**	**<0.001**	**5.644**	**1.967**
Cryoglobulinemia	1.112	3.071	0.193	16.374	0.581
ACA pattern	1.441	4.237	0.084	21.565	0.842
ESSDAI≥5	0.693	2.022	0.072	4.048	1.01
Lacrimal Gland enlargement	0.885	2.438	0.524	37.321	0.163

* Features/Variables analysed by the FCBC algorithm: Gender, age at pSS diagnosis, disease duration from SS to lymphoma diagnosis, disease duration from SS onset to SS diagnosis, dry mouth, dry eyes, salivary gland swelling, lacrimal gland enlargement, Raynaud's phenomenon, arthritis, arthralgias, palpable purpura, lymphadenopathy, renal disease-glomerulopathy, renal disease-tubulointerstitial nephritis, pulmonary disease-small airways disease, interstitial lung disease, liver disease-autoimmune hepatitis, primary biliary cirrhosis, peripheral nervous system disease, central nervous system disease, autoimmune thyroiditis, ANA, anti-La antibodies, anti-Ro antibodies, rheumatoid factors, cryoglobulinemia, low C4, ACA, ESSDAI≥5

** < 0.05 (95% confidence interval)

FCBF-based multivariable logistic regression analysis for lymphoma predictors at pSS diagnosis

Table 2. FCBF-based multivariable logistic regression analysis for lymphoma predictors at 3-4 years before lymphoma diagnosis*

Prominent feature*	Regression coefficient	Odds ratio	p-value	CI upper	CI low
ESSDAI≥5**	**1.342**	**3.871,**	**0.002**	**8.852,**	**1.694,**
Rheumatoid Factors**	**1.294**	**3.683,**	**<0.01**	**6.467,**	**2.097**
ACA pattern	1.382	3.995	0.107	21.246	0.763

* Features/Variables analysed by the FCBC algorithm: Gender, age at pSS diagnosis, disease duration from SS to lymphoma diagnosis, disease duration from SS onset to SS diagnosis, dry mouth, dry eyes, salivary gland swelling, lacrimal gland enlargement, Raynaud's phenomenon, arthritis, arthralgias, palpable purpura, lymphadenopathy, renal disease-glomerulopathy, renal disease-tubulointerstitial nephritis, pulmonary disease-small airways disease, interstitial lung disease, liver disease-autoimmune hepatitis, primary biliary cirrhosis, peripheral nervous system disease, central nervous system disease, autoimmune thyroiditis, ANA, anti-La antibodies, anti-Ro antibodies, rheumatoid factors, cryoglobulinemia, low C4, ACA, ESSDAI≥5

** < 0.05 (95% confidence interval)

FCBF-based multivariable logistic regression analysis for lymphoma predictors at 3_4 years before lymphoma diagnosis

Disclosures: A. Goules: None; **L. Chatzis**: None; **V. pezoulas**: None; **F. Ferro**: None; **V. Manfrè**: None; **D. Fotiadis**: None; **L. Quartuccio**: None; **C. Baldini**: GlaxoSmithKlein(GSK), 6, Horizon, 6, Sanofi, 6; **S. De Vita**: None; **A. Tzioufas**: None.