STEPWISE DEVELOPMENT OF A MULTI-PROTEIN BIOMARKER INDEX OF RA DISEASE ACTIVITY

Y. Shen[1], N. Knowlton[2], M. Turner[3], C. Sutton[4], D. Smith[4], D. Chernoff[1], L. Hesterberg[1], R. Roubenoff[5], N.A. Shadick[6], M. E. Weinblatt[6], G. Cavet[1] and M. Centola[3]

[1]Crescendo Bioscience, So. San Francisco, CA, [2]NSK Statistical Solns, LLC, Oklahoma City, OK, [3]OMRF, Oklahoma City, OK, [4]Crescendo Bioscience, Oklahoma City, OK, [5]Biogen-Idec, Cambridge, MA, [6]Brigham & Women's Hosp, Boston, MA

ABSTRACT # 0001122

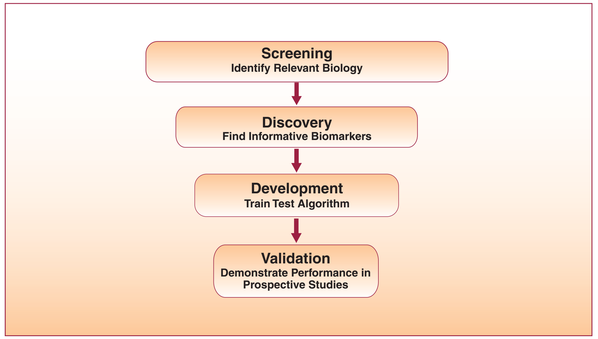
SYNOPSIS

To complement existing symptom-focused disease activity assessment tools, we are developing a multi-protein biomarker index of rheumatoid arthritis (RA) disease activity using a rigorous, stepwise development program that comprehensively surveys the biological pathways underlying RA. Candidate serum protein biomarkers were selected from an extensive screen of literature, databases, and experimental data and assays were optimized for quantitative protein assessment in serum from RA patients. Three serial studies were performed to determine the associations between candidate biomarker levels and disease activity. Statistical models with 4-11 protein biomarkers outperformed any individual biomarker at estimating disease activity. These models achieved average accuracy of >70% for assigning patients into low and high disease activity categories, and average correlations of 0.6 with DAS28 in 100 iterations of cross validation. Models developed in one cohort performed well in independent cohorts. Multi-protein biomarker index assays have the potential to elucidate RA disease biology for individual patients and improve patient assessment.

INTRODUCTION

- Tight control studies in RA such as TICORA, CAMERA, FinRACO suggest that frequent quantitative monitoring of disease activity with consequential treatment changes improves patient outcomes[1-3]
- ACR and EULAR also recommend ongoing disease activity assessment in RA[4]
- However, current disease activity assessment tools are suboptimal:
  - ESR and CRP are non-specific and do not capture the heterogeneous biology of RA
  - Clinical indices contain subjective components, which result in variability within and across assessors
- To complement existing tools, we are developing a quantitative, objective, multi-protein biomarker index of RA disease activity
- We are using a stepwise approach to develop a robust assay (Figure 1) that eliminates issues found in previous work. Our methodology includes:
  - Assessment of a large number of biomarkers from multiple biological pathways to represent RA heterogeneity
  - Technical assay optimization including rheumatoid factor (RF) blocking prior to clinical studies
  - Assessment of multiple, distinct clinical cohorts comprising:
    - Large numbers of patient samples to increase statistical power
    - Broad patient ranges to determine applicability across the entire RA population and to enable analysis of clinical covariates (eg, CCP status, RF status, gender, age)
  - Multiple studies using independent cohorts to reduce type I and type II errors and ensure biomarker reproducibility

FIGURE 1: STEPWISE DEVELOPMENT APPROACH



METHODS

Figure 2. Candidate Biomarker Selection And Optimization

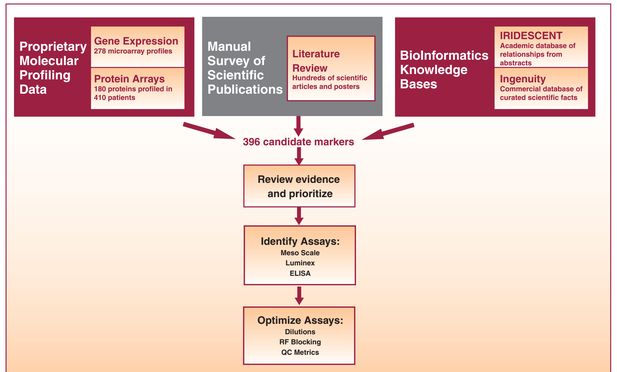


TABLE 1: 121 CANDIDATE BIOMARKERS FOR CLINICAL STUDIES AFTER TECHNICAL OPTIMIZATION

INTERLEUKINS	RECEPTORS	HORMONES	OTHERS
IL10	ager	follicle stimulating hormone	adiponectin
IL12	EGFR	gastric inhibitory polypeptide	adrenomedullin
IL12B	IL2RA	ghrelin	amyloid P component, serum
IL13	IL4R	GLP-1	bone morphogenetic
IL15	IL6R	growth hormone 1	protein 6
IL17	interleukin 1	insulin	CALCB
IL1A	receptor,	leptin	calprotein
IL1B	type I	NT-proBNP	cartilage glycoprotein-39
IL2	interleukin 1	pancreatic polypeptide	CD40 ligand
IL3	receptor,	POMC	COMP
IL4	type II	prolactin	CRP
IL5	KIT	PYY	cystatin C
IL6	sFLT4	resistin	fibrinogen
IL7	sKDR		FLT3 ligand
IL8	TNFRSF1A		fital cell derived
IL9			neurotrophic
IL1RA			gp130
			haptoglobin
			IGFBP1
			neurotrophin 4
			osteocalcin
			osteonectin
			osteopontin
			pentraxin 3
			SAA1
			sclerostin
			SERPINE1
			sFLT1
			SLPI
			thrombomodulin

Growth Factors	TNF Superfamily	TNFR Superfamily	Other Cytokines
FGF2	APRIL	CD30	EPO
EGF	LIGHT	FAS	GCSF
HGF	LTA	osteoprotegerin	GMCSF
NGF	RANKL	TNFRSF1A	IFNA1
PDGF-AA	TNF-alpha	TNFRSF1B	IFNA2
PDGF-AB	TNFSF18	TNFRSF9	IFNA2
PIGF	TWEAK		LIF
TGFA			MCSF
VEGFA			CCL22

Selectins	Adhesion Molecules	Enzymes	Apolipoproteins	Matrix Metalloproteinases
selectin E	ICAM1	alkaline	APOA1	MMP1
selectin L	ICAM3	phosphatase	APOA2	MMP10
selectin P	VCAM1	lysozyme	APOB	MMP2
		myeloperoxidase	APOC2	MMP3
		thyroid	APOC3	MMP9
		peroxidase	APOE	

Data Analysis

- Three clinical studies were performed of the cohorts of RA patients described in Table 2.
- Cohorts included patients from:
  - Oklahoma Research Foundation (OMRF), a clinical study collection from several community clinics located in and around Oklahoma City, OK
  - Brigham and Woman's Rheumatoid Sequential Study (BRASS) Registry, an observational study managed by Brigham and Women's Hospital in Boston, MA

Table 2: CHARACTERISTICS OF THE CLINICAL COHORTS FOR ALL 3 STUDIES

Categorical Variables (%)	Study 1	Study 2	Study 3
Female	82	80	91
CCP+	63	62	62
RF+	83	83	64
Smoker	Na	13	4
Concurrent medications			
• Methotrexate	53	61	48
• Non-biologic DMARDs*	69	76	64
• Biologics	65	53	43
• Corticosteroids	24	27	27
Continuous Variables (Mean±SD),(min, max)			
Age	(60±13.1),(28,88)	(59±13.8),(22,94)	(59±12.7),(29,85)
DAS28 CRP	(60±13.1),(28,88)	(4.1±1.7),(1.2,8.2)	(3.8±1.6),(1,7.9)

Analytical Methods – 3 Studies

- Univariate analysis of individual markers and disease activity
  - DAS28CRP, DAS28ESR, CDAI, SDAI, S/TJC and patient global assessment
  - Pearson and Spearman correlations
- Multivariate modeling
  - Forward stepwise linear regression, Penalized Regression, Random Forests
- Internal validation of multivariate prototype algorithms developed to estimate disease activity using multiple biomarker concentrations
  - Random subset of 70% of the total study population selected without replacement
  - Model fitted using this subset then evaluated in the remaining 30%
  - Repeated 100 times with mean result used to estimate future performance
  - Primary measure of algorithm performance was classification of patients into low vs high disease activity using DAS28CRP – models using median and 2.67
  - Models were also evaluated excluding patients with low joint counts (<3 = SJC + TJC)

RESULTS

- Statistical models with 4-11 protein biomarkers outperformed any individual biomarker at estimating disease activity
- In internal validation, the average ROC AUC was
  - 0.78 using the median DAS28CRP as a cutoff, and excluding patients with low joint counts (Figure 3a)
  - 0.89 using DAS28CRP of 2.67 as a cutoff, and including all patients (Figure 3b)
  - Average correlation with DAS28CRP was 0.6 for both analyses.

FIGURE 2: ROC ANALYSIS OF MULTIVARIATE MODELS OF DISEASE ACTIVITY USING BIOMARKERS

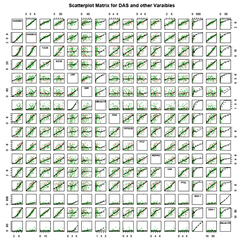


FIGURE 3: ROC ANALYSIS OF MULTIVARIATE MODELS OF DISEASE ACTIVITY USING BIOMARKERS

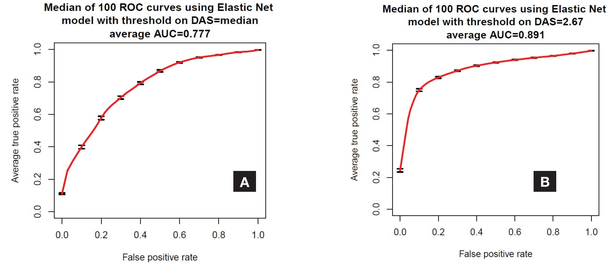
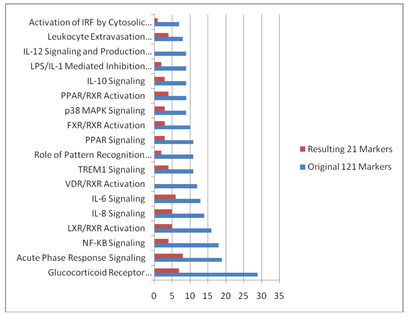


FIGURE 4: BIOLOGICAL PATHWAYS REPRESENTED BY TOP 21 MARKERS



Note: each biomarker can fall under multiple categories

- Models developed using single vs multiple time points demonstrated similar performance (0.581 vs 0.626, respectively)
- Models can track change in DAS over multiple time points
- Models developed in one cohort had similar performance in other cohorts
- 21 candidate biomarkers were prioritized for further examination; markers delineate relevant RA biology (Figure 4)

CONCLUSIONS

- A robust, stepwise development path can identify biomarkers strongly predictive of RA disease activity, which are reproducible across multiple studies and cohorts
  - Model Disease Activity Assay Prototypes have ROC AUC of 0.78-0.89
- These biomarkers are linked to several key inflammatory pathways reflecting the biological heterogeneity of RA
- This work completes the feasibility investigation of a multi-protein biomarkers index of RA disease activity and warrants study continuation:
  - Development: Further refinement of key protein biomarkers and development of a final algorithm in a prospective study of ~500 patients
  - Verification: Testing of the developed model in a separate prospective study of ~200 patients to verify the performance of the algorithm from the Development Phase
  - Validation of the resulting assay will be completed in a separate prospective study of ~300 patients

REFERENCES

1. Grigor C, et al. *Lancet.* 2004;364:263-269.
2. SM, et al. *Ann Rheum Dis.* 2007;66:1443-1149.
3. Puolakka K, et al. *Arthritis Rheum.* 2005;52:36-41.
4. Aletaha D, et al. *Arthritis Rheum.* 2008;59:1371-1377.

CRESCENDO BIOSCIENCE