Title

Prognostic biomarkers of IFNb therapy in multiple sclerosis patients.

Permalink

https://escholarship.org/uc/item/4t3870fr

Journal

Multiple sclerosis (Houndmills, Basingstoke, England), 21(7)

ISSN

1352-4585

Authors

Baranzini, Sergio E
Madireddy, Lohith R
Cromer, Anne
et al.

Publication Date

2015-06-01

DOI

10.1177/1352458514555786

Peer reviewed

Original Research Paper

Prognostic biomarkers of IFNb therapy in multiple sclerosis patients

Sergio E Baranzini, Lohith R Madireddy, Anne Cromer, Mauro D'Antonio, Lorenz Lehr, Manolo Beelke, Pierre Farmer, Marco Battaglini, Stacy J Caillier, Maria L Stromillo, Nicola De Stefano, Emmanuel Monnet and Bruce AC Cree

Abstract

Background: Interferon beta (IFNb) reduces relapse frequency and disability progression in patients with multiple sclerosis (MS).

Objectives: Early identification of prognostic biomarkers of IFNb-treated patients will allow more effective management of MS.

Methods: The IMPROVE study evaluated subcutaneous IFNb versus placebo in 180 patients with relapsing–remitting MS. Magnetic resonance imaging scans, clinical assessments, and blood samples were obtained at baseline and every 4 weeks from every participant. Thirty-nine biomarkers (32 transcripts; seven proteins) were studied in 155 patients from IMPROVE. Therapeutic response was defined by absence of new combined unique lesions, relapses, and sustained increase in Expanded Disability Status Scale over 1 year. A machine learning approach was used to examine the association between biomarker expression and treatment response.

Results: While baseline levels of individual genes were relatively poor predictors, combinations of three genes were able to identify subjects with sub-optimal therapeutic responses. The triplet CASP2/IRF4/IRF6, previously identified in an independent dataset, was tested among other combinations. This triplet showed acceptable predictive accuracy (0.68) and specificity (0.88), but had relatively low sensitivity (0.22) resulting in an area under the curve (AUC) of 0.63. Other combinations of biomarkers resulted in AUC of up to 0.80 (e.g. CASP2/IL10/IL12Rb1).

Conclusions: Baseline expression, or induction ratios, of specific gene combinations correlate with future therapeutic response to IFNb, and have the potential to be prognostically useful.

Correspondence to:
Sergio E. Baranzini
Department of Neurology, University of California, San Francisco (UCSF), 675 Nelson Rising Lane, 2nd floor Suite 242, Box 3206, San Francisco, CA 94158, USA.
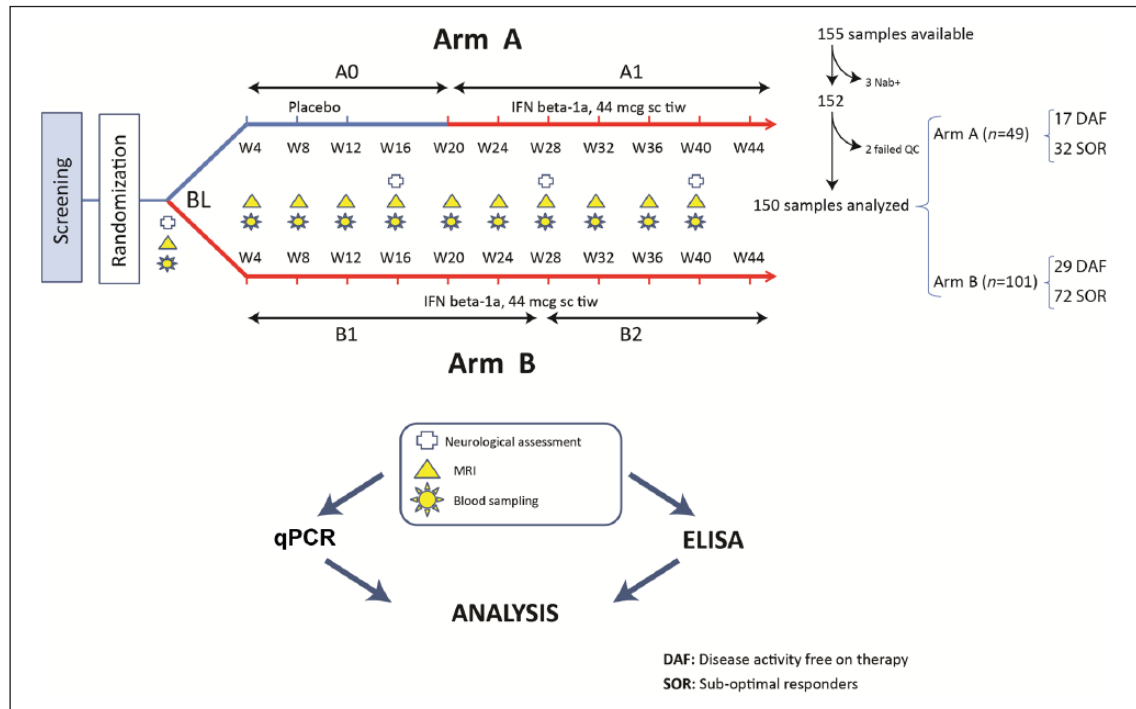**sebaran@cgl.ucsf.edu**

Sergio E Baranzini
Lohith R Madireddy
Stacy J Caillier
Bruce AC Cree
Department of Neurology, University of California, San Francisco (UCSF), San Francisco, USA/ Equal contribution

Anne Cromer
Lorenz Lehr
Manolo Beelke
Pierre Farmer
Emmanuel Monnet
Merck Serono S.A. – Geneva, Switzerland#/During the completion of this study, Merck Serono closed its Geneva operations. These authors are no longer with the company.

Mauro D'Antonio
Merck Serono RBM S.p.A– Colleretto Giacosa, Turin, Italy

Marco Battaglini
University of Siena, Siena, Italy

Introduction

Since its approval in the USA in 1993, interferon beta (IFNb) has been widely used to reduce relapse frequency in patients with multiple sclerosis (MS).[1–3] Although its precise mechanism of action has not been fully elucidated, IFNb is thought to modulate immune responses by shifting the Th1/Th2 balance, inducing T-cell apoptosis and altering expression of cell adhesion molecules.[4,5] Although IFNb has demonstrated efficacy and an excellent long-term safety profile, a proportion of patients experience ongoing disease activity despite treatment. In particular, the presence of new lesions on magnetic resonance imaging (MRI) scans or the occurrence of clinical relapses in IFNb treated can be associated with long-term disability.[6,7] The identification of biomarkers predictive of therapeutic response would be clinically useful both for identifying patients who will optimally respond to treatment and those who are at risk for ongoing disease activity.

Using a data mining approach, specific transcription-based signatures were associated with therapeutic response in a previous study.[8] A goal of the present study was to determine whether these observations could be validated in an independent cohort.

Figure 1. IMPROVE study design.
ELISA: enzyme-linked immunosorbent assay; qPCR: quantitative polymerase chain reaction.

Furthermore, the present study includes high-frequency correlations and longitudinal analyses aimed at determining the added value of biomarkers to the more established metrics of disease activity based solely on clinical and radiological examinations. Here we tested the expression of a select group of transcripts and proteins in blood samples from subjects participating in the IMPROVE study, a trial to evaluate the efficacy of a serum-free preparation (originally known as Rebif® New Formulation).[9] This 11-month study (18 months total follow-up period) enrolled 180 subjects randomized into two arms with monthly MRI and blood sampling and periodic neurological examinations. This unique study design provided an opportunity to assess molecular correlates that were predictive of both clinical and radiological disease activity.

Materials and methods

Subjects and clinical parameters

Subjects participating from the IMPROVE study were randomized into two treatment arms and followed for 40 weeks with monthly MRI scans and quarterly neurological evaluations (an extended follow-up was done at 18 months) (Figure 1).[9,10] A subgroup of 155 of the 180 original IMPROVE subjects underwent monthly blood draws for biomarker analysis. Subjects in Arm A received placebo for the first 16 weeks (Arm A0) and then active drug for the remaining period (Arm A1). Individuals in Arm B received active drug from baseline (BL). MRI metrics included the number of new T1 lesions (T1L), new T2 (T2L), new gadolinium-DPTA (Gd)-enhancing (GDL), new ring-enhancing (REL), and combined unique (new T2 lesions and/or Gd-enhancing lesions) active lesions (CU), as well as volume changes of Gd-enhancing lesions (ΔVGDL). Neurological examinations included relapse and Expanded Disability Status Scale (EDSS) assessments. Clinical characteristics of subjects in this study are described in Table 1.

Samples and biomarkers

Whole blood for RNA (Paxgene tubes) and anticoagulated plasma were obtained from each available participant at baseline and at each of the 10 subsequent time points. Total RNA was obtained from each sample using Blood RNA kit (Qiagen) and subsequently purified with RNeasy columns (Qiagen) to obtain quantitative polymerase chain reaction (qPCR)-grade RNA. Reactions were performed in an ABI7900 sequence analyzer. Expression of protein biomarkers was determined by validated enzyme-linked immunosorbent assay

Table 1. Clinical characteristics of subjects in study.

Parameter	Value
Female:male ratio (*n*)	2.6 (112:43)
Mean age at entry into IMPROVE (SD)	33.8 (8.97)
Median EDSS at baseline (range)	2.5 (0–5.5)
EDSS: Expanded Disability Status Scale.	

(ELISA) or colorimetric assays. See Supplementary Materials and Methods for more details.

Statistical analysis

The average expression levels of each biomarker at each time point after initiation of therapy was compared with the average of their expression in the absence of treatment by means of a two-sided independent two-sample *t*-test. Reported *p*-values are uncorrected unless otherwise noted.

The response of each subject to treatment was classified as being either "disease activity free on therapy" (DAF) or sub-optimal responders (SOR) according to presence of clinical and/or radiographic disease activity criteria. Because anti-IFNb neutralizing antibodies (NAbs) abrogate IFN-induced gene transcripts and inhibit the impact of IFNb on MS relapsing activity, subjects who tested positive for NAbs at baseline (*n*=3) were excluded from analysis.[11] Two samples failed quality control. In total, 49 samples from Arm A and 101 from Arm B were subjected to further analysis.

Normalized expression data at baseline was used to compute random forests with the MLInterfaces package within the *R* statistical software. Induction ratios between baseline (average of W4-W16 in Arm A) and 4 weeks after initiation of treatment were also computed and used to predict response in a similar fashion. Random forests are a type of recursive partitioning method particularly well-suited to small n large p problems.[12] Furthermore, the results of an ensemble of classification/regression trees have been shown to produce better predictions than the results of one classification tree on its own.[12]

See Supplementary Materials and Methods for more details.

Correlation of biomarker expression with MRI activity, relapse rate and EDSS

This analysis aimed at identifying biomarkers that correlated with any of the six MRI parameters measured in the absence of IFNb treatment. Only samples from subjects in Arm A collected at baseline (ABL) and through the placebo stage (A0) (i.e. ABL + A0) were considered for this analysis. For each biomarker, the expression level was correlated to the profiles of all MRI and clinical parameters measured during the same time period.

For every patient the correlation coefficient (*R*) between biomarker expression and each MRI or clinical parameter (expressed as vectors of values over time points) was computed. The distribution of *R* obtained over all patients was compared to zero (null hypothesis of no correlation) by means of a two-sided one-sample *t*-Test. Reported *p*-values are uncorrected unless otherwise noted.

Results

Blood samples from 155 subjects who participated in IMPROVE were available to test correlations between select protein and RNA biomarkers and MRI metrics and clinical parameters over time (Figure 1). Three subjects with NAbs present at baseline and samples from two subjects who failed quality control were excluded from the remainder of the study. In total, 46 subjects (17 in Arm A and 29 in Arm B) met DAF on therapy criteria, and 104 subjects were defined as SOR (32 in Arm A and 72 in Arm B). Of the DAF group, 26 individuals (seven in Arm A and 19 in Arm B) did not show evidence of disease activity at baseline or at subsequent examinations. It is possible that these subjects had either an optimal therapeutic response or mild disease.

The expression of 32 transcripts, six proteins, and nitric oxide (NO) was tested in all available subjects after quality control (*n*=49 from Arm A and *n*=101 from Arm B) at baseline and at each of the following 10 months. Transcripts for analysis were selected based on a previous study[8] and included those coding for components of the canonical type-I IFN signaling pathway, cell cycle control, apoptosis, and cytokines and receptors in lymphocyte differentiation pathways. The housekeeping gene GAPDH was used as a control to normalize all other transcripts.
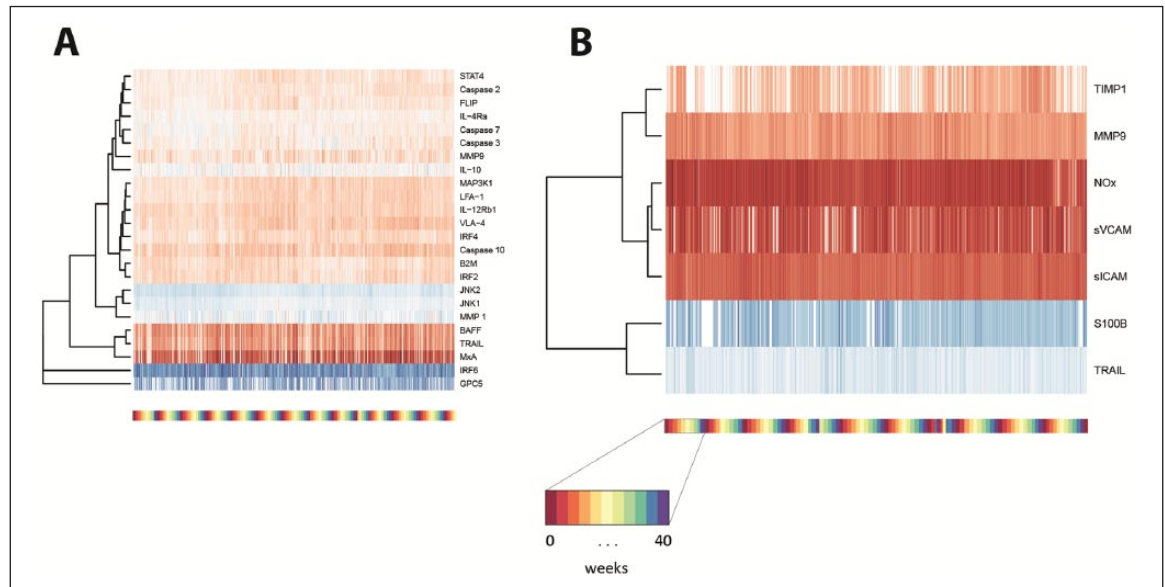
Figure 2. Heatmap of gene and protein expression in 1705 samples from IFNb-treated MS patients. Each row corresponds to a biomarker (A: transcripts, B: proteins) and each column represents a sample. For each patient, samples are sorted by time point (baseline through week 40). The color of each cell represents the normalized expression (blue=low, red=high).

Normalized levels of each biomarker (24 transcripts, six proteins and NO) at all time points were organized (by patient) and visualized in heatmaps, in which genes (or proteins) were clustered by similarity across samples (Figure 2). Transcription profiles for MXA, TRAIL and BAFF were closely correlated and showed the highest expression, an expected finding given that these are IFNb-responsive genes whose expression increases within hours following IFNb administration.[13–19] Protein levels of sVCAM were similarly up-regulated by IFNb.[20–24] These expected changes in response to IFNb administration were reassuring and served as useful controls of sample handling, processing, and overall data quality.

The associations between biomarker expression and treatment in all individuals from Arm B (*n*=101) were tested next. Specifically, the difference in expression between baseline and the average of all time points after drug administration was computed. Notably, the expression of 19/24 transcripts showed a statistically significant association with treatment (two-sided *t*-test, $p<0.05$) (Figure 3(b)). Of these, 13 were significantly increased (up to four-fold) while only six were decreased (by at least 1.5-fold). Similarly, five proteins were significantly regulated by treatment (sICAM, sVCAM, TIM-1 and TRAIL were up-regulated whereas MMP-9 was down-regulated). Similar results were observed for individuals from Arm A (*n*=53) by computing the difference between the

average of all time points before and after treatment (Figure 3(a)). Consistent with prior studies, the two transcripts that were most significantly induced by treatment were MxA (>16-fold, $p<10^{-25}$)[14,18,25,26] and TRAIL (> 16–fold, p < 10^{-21}, Figure 3(a)).[26,27] The correlation between transcript and protein levels of TRAIL was moderately high (R^2=0.632) and was within the reported range of global concordance between RNA and protein levels.[28,29] The individual patient data for TRAIL levels before and after treatment is visualized in Figure 3(c), where a clear separation between baseline (red) and treatment (blue) samples was seen.

To determine whether these gene transcripts and proteins could predict therapeutic response we employed a machine learning approach (random forest classification algorithm, see methods) to identify combinations of biomarkers that, when measured at baseline, correlate with therapeutic response measured over the course of the study. In a previous report that defined therapeutic response by the absence of clinical relapses or neurological worsening over 2 years of follow-up, nine gene triplets with a predictive accuracy of at least 80% (range 80–87%) were identified.[8] The predictive accuracy of the same triplets in this cohort ranged from 59–68%) and the area under the receiver operating characteristic curve (AUC) was up to 63% (Table 2). It is worth noticing, however, that in the previous study, response was defined using only clinical (not
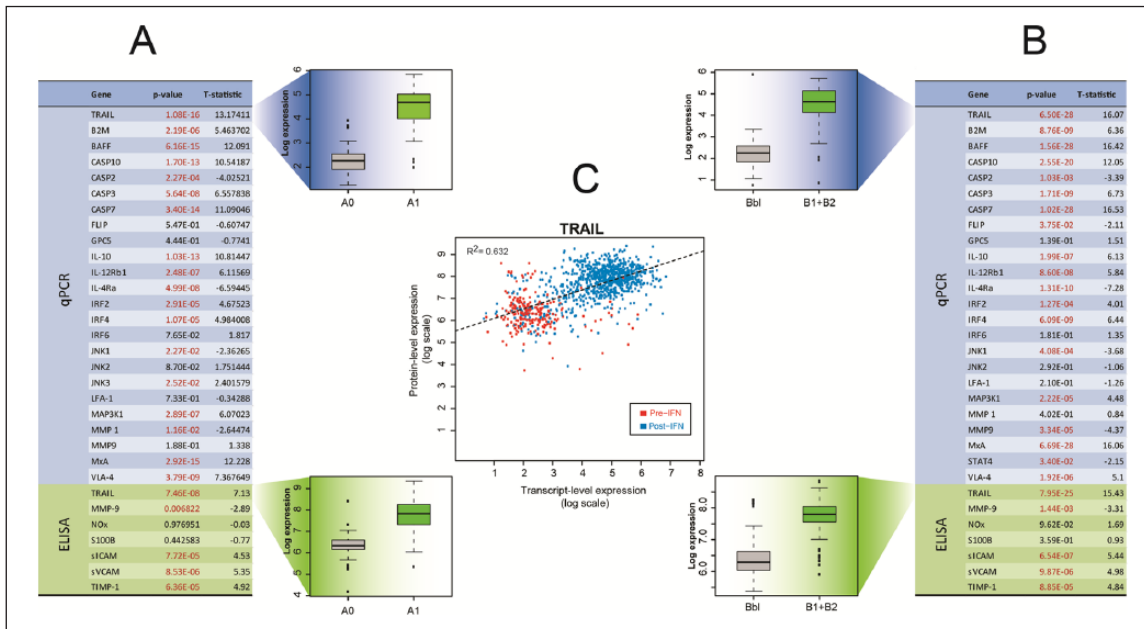
Figure 3. Correlation between biomarker expression and treatment. For each Table the *p*-value and *t*-statistic is provided. The *t*-statistic provides an indication of whether the biomarker is increased (positive) or decreased (negative) with treatment. **A**, Arm A; **B**, Arm B; **C**. correlation between qPCR and ELISA. Red: untreated; blue: treated.
ELISA: enzyme-linked immunosorbent assay; qPCR: quantitative polymerase chain reaction.

imaging) information. When only clinical criteria were used to define DAF and SOR in the present study, the predictive accuracy of those same triplets was 76–78% (Table 2). Sub-clinical brain MRI lesions will occur more frequently than clinical measures of disease activity. Thus when MRI information is incorporated, fewer patients will meet the DAF criteria than when only clinical measures are used.

A recent report suggested that measuring biomarker induction ratios after IFNb administration provides a more sensitive predictor of long-term response than baseline measurements alone.[30] We used induction ratios computed from baseline and week 4 (the next available time point) to predict response and observed an improvement in the predictive accuracy for the same biomarkers (up to 72%). However, the AUC remained unchanged at 62% (Table 3).

We next explored the performance of additional combinations using the same set of biomarkers in all subjects for whom all transcripts could be reliably measured (*n* ≥ 59). Several gene triplets were identified whose predictive accuracy and AUC performed numerically better than the gene combinations listed in Tables 2 and 3. The top-scoring classifier was CASP2/IL-10/IL12Rb1, yielding a predictive accuracy of 82% and AUC of 0.80. Of these individual transcripts, only the expression of CASP2 resulted in

a significant difference between the DAF and SOR groups. The expression of IL10 was borderline and that of IL12Rb1 was not significant (Figure 4(a)). Correspondingly, the individual AUC were 0.71, 0.68 and 0.54 (Figure 4(b)). This example highlights the power of gene combinations over single gene predictors. Seven other combinations resulted in AUC ≥75% (Table 4), and yet another set of 54 triplets yielded a predictive accuracy ≥80% (Supplementary Table 1).

To further evaluate that the overall performance of these classifiers was greater than would occur by chance association, we generated 100 datasets using the same expression values, but in which the DAF and SOR groups were randomly assigned (permuted). We then compared the AUC and predictive accuracy of the real dataset against the average of the randomly generated datasets (Supplementary Figure 1). Both AUCs and predictive accuracies of the original dataset were greater than those obtained with the permuted data, indicating that the identified classifiers held predictive power. This is a remarkable result given the small number of genes tested in this independent cohort.

The longitudinal nature of IMPROVE enabled the evaluation of correlations between biomarker expression and clinical and imaging parameters. The expression of four transcripts (Caspase -3, -7 and-10, and IL12Rb1)

**Table 2.** Validation of triplet biomarkers using baseline expression as predictor.

Gene triplets	AUC	Predictive accuracy[a]	Sensitivity	Specificity	PPV	NPV	Overall error rate	Non-responder error rate	Responder error rate	Sample size
Caspase2 IRF4 IRF6	0.63	0.68 (0.77)	0.22	0.88	0.47	0.72	0.32	0.12	0.78	113
Caspase2 Caspase7 IRF4	0.59	0.68 (0.77)	0.32	0.85	0.51	0.72	0.32	0.15	0.68	136
Caspase10 Caspase2 FLIP	0.52	0.64 (0.76)	0.24	0.83	0.41	0.69	0.36	0.17	0.76	136
Caspase2 Caspase3 IRF4	0.50	0.63 (0.76)	0.23	0.83	0.39	0.69	0.37	0.17	0.77	135
Caspase2 IRF2 STAT4	0.55	0.63 (0.77)	0.25	0.81	0.40	0.69	0.37	0.19	0.75	136
Caspase2 IL4Ra IRF4	0.53	0.63 (0.77)	0.23	0.82	0.38	0.69	0.37	0.18	0.77	136
Caspase10 IL12Rb1 MAP3K1	0.50	0.62 (0.77)	0.21	0.82	0.36	0.68	0.378	0.18	0.79	136
Caspase2 Caspase3 IL4Ra	0.49	0.62 (0.78)	0.20	0.82	0.35	0.68	0.38	0.178	0.80	135
Caspase10 IRF4 MAP3K1	0.52	0.59 (0.78)	0.18	0.79	0.29	0.67	0.41	0.21	0.82	136

[a]The PA of each triplet based on the same definition of response used in Baranzini et al (2005) is indicated in brackets.
AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value.

Table 3. Triplet biomarkers using induction ratios as predictors.

Gene triplets	AUC	Predictive accuracy	Sensitivity	Specificity	PPV	NPV	Overall error rate	Non-responder error rate	Responder error rate	Sample size
Caspase10 Caspase2 FLIP	0.62	0.72	0.31	0.89	0.56	0.75	0.29	0.11	0.69	120
Caspase2 IRF4 IRF6	0.64	0.70	0.22	0.86	0.35	0.77	0.30	0.14	0.78	93
Caspase2 Caspase7 IRF4	0.68	0.69	0.27	0.87	0.48	0.74	0.31	0.13	0.73	119
Caspase10 IRF4 MAP3K1	0.55	0.69	0.26	0.87	0.49	0.74	0.31	0.13	0.74	120
Caspase2 IL4Ra IRF4	0.63	0.69	0.25	0.87	0.47	0.73	0.31	0.13	0.75	119
Caspase2 IRF2 STAT4	0.64	0.68	0.32	0.84	0.48	0.74	0.32	0.16	0.68	119
Caspase2 Caspase3 IL4Ra	0.54	0.67	0.21	0.86	0.40	0.72	0.33	0.14	0.79	119
Caspase10 IL12Rb1 MAP3K1	0.56	0.66	0.26	0.85	0.47	0.71	0.34	0.15	0.74	121
Caspase2 Caspase3 IRF4	0.58	0.65	0.20	0.84	0.34	0.71	0.35	0.16	0.80	118

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value.
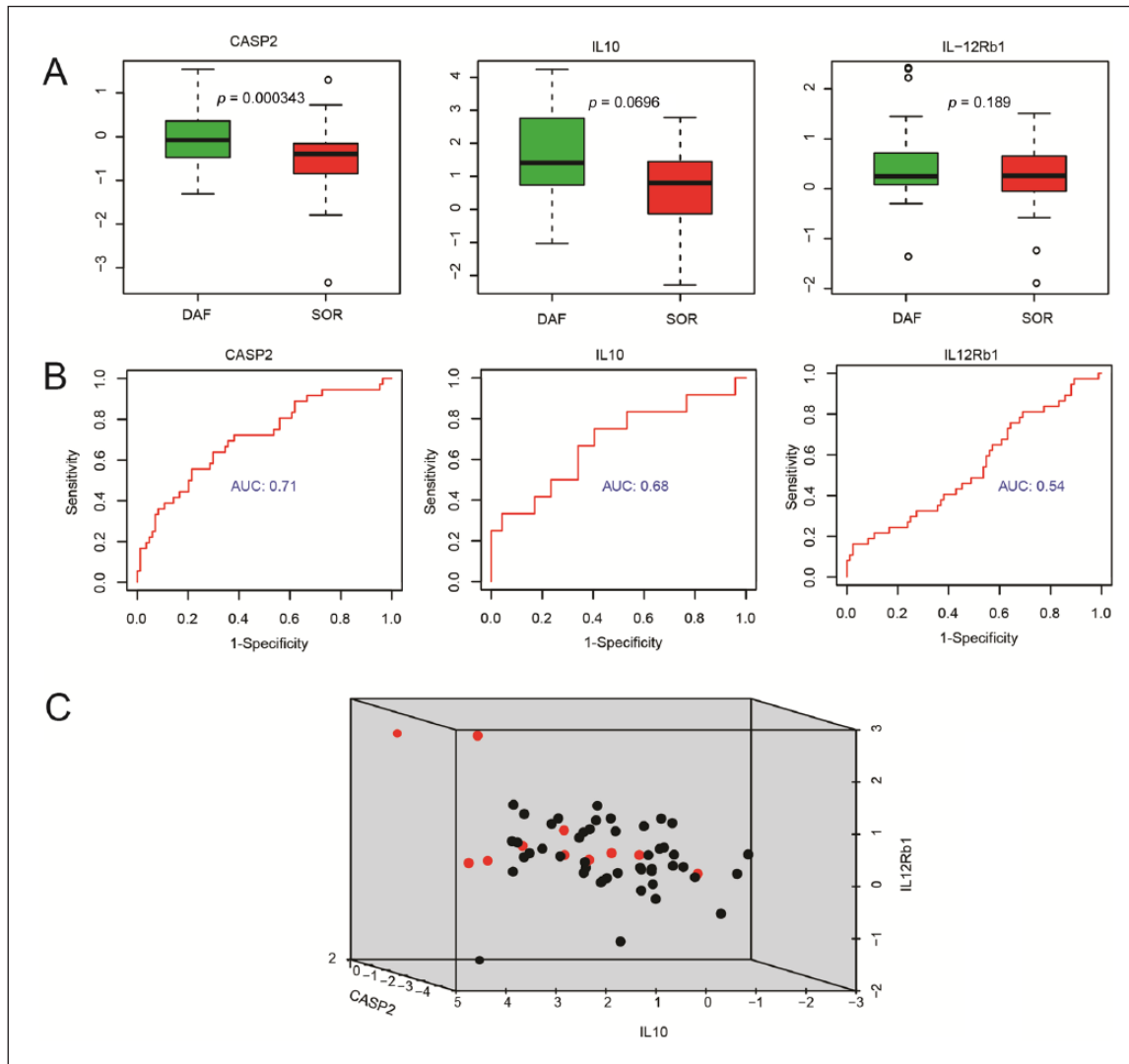
Figure 4. Performance of top classifier. A. box plots of the expression of each transcript in the classifier in disease activity free on therapy (DAF) vs. sub-optimal responders (SOR). B. Receiver operating characteristic curves showing the individual predictive power of each transcript. C. A tree-dimensional plot showing the expression values for all three transcripts of the CASP-2, IL-10, IL12Rb1 triplet. DAF (red); SOR (black).
AUC: area under the curve.

was negatively correlated with clinical attacks ($p>0.001$). Another five transcripts (BAFF, JNK2, MAP3K1, MXA and TRAIL) also showed negative correlation albeit at a lower level of significance ($p<0.05$). Only MMP-9 was positively correlated with relapses ($p<0.05$) (Table 5). At the protein level, TIMP-1 ($p<0.001$) and NO and S100B ($p<0.05$) were negatively correlated with the presence of relapses (data not shown). None of the tested biomarkers correlated with sustained EDSS increase or six MRI metrics (T1L, T2L, GDL, REL, NCU, and ΔVGDL) measured at each of the 11 time points. This is likely due to the narrow list of markers tested, which was originally designed around the biological effects of IFNb.

Discussion

The modest size of analyzed cohorts as well as heterogeneous criteria used to define therapeutic response thus far have limited identification of validated biomarkers that are predictive of therapeutic response in MS.[31,32] In the present study, we analyzed samples from 155 subjects from the IMPROVE study, constituting the largest and most rigorously characterized MS cohort for biomarker discovery to date. We reasoned that using a strict definition of therapeutic response would be of greatest clinical relevance and thus performed our study using the DAF criteria. The DAF criteria are emerging as an increasingly important metric in recent clinical trials because they incorporate both radiographic and

Table 4. Top gene combination using induction ratio predictors.

Gene triplets	AUC	Predictive accuracy	Sensitivity	Specificity	PPV	NPV	Overall error rate	Non-responder error rate	Responder error rate	Sample size[a]
Caspase2 IL10 IL12Rb1	0.80	0.82	0.25	0.96	0.44	0.84	0.19	0.04	0.75	59
Caspase2 IL10 IRF4	0.79	0.78	0.17	0.93	0.29	0.82	0.22	0.07	0.83	59
Caspase2 MMP1 MMP9	0.79	0.79	0.48	0.91	0.67	0.83	0.21	0.09	0.52	58
Caspase2 LFA1 MMP1	0.77	0.77	0.45	0.89	0.61	0.82	0.23	0.11	0.55	58
Caspase2 JNK2 MMP1	0.77	0.79	0.43	0.92	0.68	0.82	0.21	0.08	0.57	58
Caspase2 IL10 STAT4	0.76	0.78	0.17	0.93	0.27	0.82	0.22	0.07	0.83	59
Caspase2 IL10 IRF2	0.75	0.80	0.30	0.93	0.47	0.84	0.20	0.07	0.70	59
Caspase2 IRF4 MMP1	0.75	0.77	0.44	0.89	0.61	0.82	0.23	0.11	0.56	58

[a]The number of samples with complete information (the machine learning approach can only be employed if no data is missing). AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value.

Table 5. Transcripts correlated with relapses.

Gene symbol	*p*-value	*t*-statistic
BAFF	0.03	−2.24
CASP10	0.005	−2.98
CASP3	0.005	−2.97
CASP7	7.43E-05	−4.40
IL12Rb1	0.002	−3.26
JNK2	0.04	−2.08
MAP3K1	0.01	−2.69
MMP9	0.04	2.09
MxA	0.01	−2.69
TRAIL	0.005	−2.93
VLA-4	0.0002	−3.99

clinical measures of disease activity and are thought to possibly represent disease remission.[33,34]

Particular combinations of transcripts that are predictive of IFNb response when measured before initiation of therapy were previously reported.[8] In that earlier report, the top classifier (CASP2, CASP10 and FLIP) achieved 87% predictive accuracy. Here, we report validation of this triplet and other predictors in an independent cohort with predictive accuracy up to 68%. The numerically lower predictive accuracy is in part because here we use a more sensitive indicator of SOR that incorporated monthly MRI scans. When only clinical measures of disease activity were used to define DAF and SOR then the predictive accuracy of the previously identified triplets increased up to 78%.

Based on a prior report[30] in which an exaggerated response to IFNb was associated with sub-optimal response, we tested the performance of predictors computed from 1-month induction ratios rather than baseline expression. In this analysis, the use of induction ratios on the previously reported triplet CASP2/IRF4/IRF6 (Table 2) resulted in a predictive accuracy of 72%, a modest improvement over the accuracy obtained using baseline values alone, mostly due to an increase in sensitivity. One difference between the present study and that of Rudick et al.[30] is the way in which induction ratios were computed. In that study, samples were obtained at baseline and 12 h after the first administration of IFNb, while samples in the present study were collected at 4-week intervals. However, in order to measure the stability of induction ratios, Rudick et al. also performed similar measurements at 6 and 24 months showing that the induction ratios were largely independent of the time from first injection. We thus inferred that computing induction ratios at 4 weeks was a valid strategy.

When a search for new combinations of gene triplets was performed using the 1-month transcript induction ratio, the triplet CASP2/IL10/IL12Rb1 achieved a predictive accuracy of 82% with an AUC of 0.80. This observation suggests that the addition of IL10 to the gene expression profile could further improve the predictive value of this assay. However, the IL10 transcript is usually expressed at very low levels and thus was reliably observed in only a subset of the overall cohort. Technical improvements in the IL10 transcription assay might allow verification of the utility of this biomarker in a more representative sample of this dataset.

Subjects without baseline activity (*n*=22), or no activity during treatment (*n*=26), were included in the DAF group. However, these subjects could either be optimally responsive to treatment or might simply have very mild MS. Unfortunately, due to the short follow-up period, this study cannot distinguish between those two groups. It is possible that biomarkers of therapeutic response could be different from those associated with quiescent disease, and therefore the inclusion of patients with spontaneously quiescent MS would reduce the predictive accuracy of markers associated with true therapeutic response. However, the baseline transcription profiles of the 20 DAF subjects (10 in Arm A and 10 in Arm B) who had baseline Gd-enhancing lesions was not substantially different from that of the 26 patients who had no disease activity at any time during the course of the study (data not shown), suggesting these two groups of subjects might share similar underlying biology.

Another limitation of the present study is that the DAF criteria resulted in a disproportion between DAF and SOR patients. Because MRI disease activity occurs more frequently than either clinical relapses or disability progression, for a partially effective therapy the number of SOR patients would be expected to be greater than that of DAF patients. Indeed, only 46 of the 150 evaluable study subjects met DAF criteria in this study. This unbalanced partition has direct consequences on the performance of the predictors. Specifically, if a classifier is trained with more SOR than DAF, the error rate in predicting SOR will be lower than DAF as can be seen in Tables 2–4. Although it may be clinically desirable to identify the subgroup of optimal therapeutic responders, to do so will require a larger number of study subjects. Nevertheless, the biomarkers employed in the present study do have potential clinical utility in identifying SOR. Such patients may be at risk for disease progression and could be monitored more closely for disease activity.[6]

In summary, here we use a well-powered and characterized patient cohort to test previously reported prognostic biomarkers of IFNb-treated MS patients. Several of the original biomarker combinations achieved acceptable predictive accuracy in this independent study. However, new combinations of these transcripts that showed increased predictive accuracy were discovered. These gene expression signatures are more robust at predicting patients who will experience disease activity despite treatment with IFNb than the group of patients who will be free from disease activity on treatment. Although larger studies are warranted, these biomarkers have the potential to be clinically useful.

Conflict of interest
Sergio E. Baranzini and Bruce AC Cree have participated in Advisory boards from EMD-Serono.

References

1. The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* 1993; 43: 655–661.

2. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. *Lancet* 1998; 352: 1498–1504.

3. Interferon beta-1A for relapsing multiple sclerosis. *Med Lett Drugs Ther* 1996; 38: 63–64.

4. Dhib-Jalbut S and Marks S. Interferon-beta mechanisms of action in multiple sclerosis. *Neurology* 2010; 74(Suppl 1): S17–S24.

5. Yong VW, Chabot S, Stuve O, et al. Interferon beta in the treatment of multiple sclerosis: Mechanisms of action. *Neurology* 1998; 51: 682–689.

6. Prosperini L, Gallo V, Petsas N, et al. One-year MRI scan predicts clinical response to interferon beta in multiple sclerosis. *Eur J Neurol* 2009; 16: 1202–1209.

7. Bermel RA, You X, Foulds P, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann Neurol* 2013; 73: 95–103.

8. Baranzini SE, Mousavi P, Rio J, et al. Transcription-based prediction of response to IFNbeta using supervised computational methods. *PLoS Biol* 2005; 3: e2.

9. De Stefano N, Curtin F, Stubinski B, et al. Rapid benefits of a new formulation of subcutaneous interferon beta-1a in relapsing-remitting multiple sclerosis. *Mult Scler* 2010; 16: 888–892.

10. De Stefano N, Sormani MP, Stubinski B, et al. Efficacy and safety of subcutaneous interferon beta-1a in relapsing-remitting multiple sclerosis: Further outcomes from the IMPROVE study. *J Neurol Sci* 2012; 312: 97–101.

11. Goodin DS, Frohman EM, Hurwitz B, et al. Neutralizing antibodies to interferon beta: Assessment of their clinical and radiographic impact: An evidence report. Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology* 2007; 68: 977–984.

12. Strobl C, Malley J and Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods* 2009; 14: 323–348.

13. Gandhi KS, McKay FC, Schibeci SD, et al. BAFF is a biological response marker to IFN-beta treatment in multiple sclerosis. *J Interferon Cytokine Res* 2008; 28: 529–539.

14. Gilli F, Marnetto F, Caldano M, et al. Biological markers of interferon-beta therapy: Comparison among interferon-stimulated genes MxA, TRAIL and XAF-1. *Mult Scler* 2006; 12: 47–57.

15. Hedegaard CJ, Sellebjerg F, Krakauer M, et al. Interferon-beta increases systemic BAFF levels in multiple sclerosis without increasing autoantibody production. *Mult Scler* 2011; 17: 567–577.

16. Hesse D, Sellebjerg F and Sorensen PS. Absence of MxA induction by interferon beta in patients with MS reflects complete loss of bioactivity. *Neurology* 2009; 73: 372–377.

17. Krumbholz M, Faber H, Steinmeyer F, et al. Interferon-beta increases BAFF levels in multiple sclerosis: Implications for B cell autoimmunity. *Brain* 2008; 131: 1455–1463.

18. Pachner A, Narayan K, Price N, et al. MxA gene expression analysis as an interferon-beta bioactivity measurement in patients with multiple sclerosis and the identification of antibody-mediated decreased bioactivity. *Mol Diagn* 2003; 7: 17–25.

19. Roers A, Hochkeppel HK, Horisberger MA, et al. MxA gene expression after live virus vaccination: a sensitive marker for endogenous type I interferon. *J Infect Dis* 1994; 169: 807–813.

20. Calabresi PA, Prat A, Biernacki K, et al. T lymphocytes conditioned with Interferon beta induce membrane and soluble VCAM on human brain endothelial cells. *J Neuroimmunol* 2001; 115: 161–167.

21. Hua LL, Kim MO, Brosnan CF, et al. Modulation of astrocyte inducible nitric oxide synthase and cytokine expression by interferon beta is associated with induction and inhibition of interferon gamma-activated sequence binding activity. *J Neurochem* 2002; 83: 1120–1128.

22. Matusevicius D, Kivisakk P, Navikas VV, et al. Influence of IFN-beta1b (Betaferon) on cytokine mRNA profiles in blood mononuclear cells and plasma levels of soluble VCAM-1 in multiple sclerosis. *Eur J Neurol* 1998; 5: 265–275.

23. Ossege LM, Sindern E, Patzold T, et al. Immunomodulatory effects of interferon-beta-1b in vivo: Induction of the expression of transforming growth factor-beta1 and its receptor type II. *J Neuroimmunol* 1998; 91: 73–81.

24. Yao SY, Ljunggren-Rose A, Stratton CW, et al. Regulation by IFN-beta of inducible nitric oxide synthase and interleukin-12/p40 in murine macrophages cultured in the presence of *Chlamydia pneumoniae* antigens. *J Interferon Cytokine Res* 2001; 21: 137–146.

25. Bertolotto A, Gilli F, Sala A, et al. Evaluation of bioavailability of three types of IFNbeta in multiple sclerosis patients by a new quantitative-competitive-PCR method for MxA quantification. *J Immunol Methods* 2001; 256: 141–152.

26. Kayagaki N, Yamaguchi N, Nakayama M, et al. Type I interferons (IFNs) regulate tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) expression on human T cells: A novel mechanism for the antitumor effects of type I IFNs. *J Exp Med* 1999; 189: 1451–1460.

27. Guo Y, Xiao P, Lei S, et al. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochim Biophys Sin (Shanghai)* 2008; 40: 426–436.

28. Gry M, Rimini R, Stromberg S, et al. Correlations between RNA and protein expression profiles in 23 human cell lines. *BMC Genomics* 2009; 10: 365.

29. Shankavaram UT, Reinhold WC, Nishizuka S, et al. Transcript and protein expression profiles of the NCI-60 cancer cell panel: An integromic microarray study. *Mol Cancer Ther* 2007; 6: 820–832.

30. Rudick RA, Rani MR, Xu Y, et al. Excessive biologic response to IFNbeta is associated with poor treatment response in patients with multiple sclerosis. *PLoS ONE* 2011; 6: e19262.

31. Rio J, Nos C, Tintore M, et al. Defining the response to interferon-beta in relapsing-remitting multiple sclerosis patients. *Ann Neurol* 2006; 59: 344–352.

32. Rudick RA, Lee JC, Simon J, et al. Defining interferon beta response status in multiple sclerosis patients. *Ann Neurol* 2004; 56: 548–555.

33. Giovannoni G, Cook S, Rammohan K, et al. Sustained disease-activity-free status in patients with relapsing-remitting multiple sclerosis treated with cladribine tablets in the CLARITY study: A post-hoc and subgroup analysis. *Lancet Neurol* 2011; 10: 329–337.

34. Havrdova E, Galetta S, Hutchinson M, et al. Effect of natalizumab on clinical and radiological disease activity in multiple sclerosis: A retrospective analysis of the Natalizumab Safety and Efficacy in Relapsing-Remitting Multiple Sclerosis (AFFIRM) study. *Lancet Neurol* 2009; 8:254–260.

Visit SAGE journals online
http://msj.sagepub.com

Ⓢ SAGE journals