



**UNIVERSIDADE FEDERAL DA BAHIA**  
**FACULDADE DE MEDICINA DA BAHIA - FMB**  
**Programa de Pós-graduação em Ciências da Saúde**

---



**Uso de biomarcadores para diagnóstico e predição de tuberculose ativa**  
**"Use of biomarkers for diagnosis and prediction of active tuberculosis"**

Betânia Mara de Freitas Nogueira

Tese de Doutorado

Salvador (Bahia), 2025

## SUMÁRIO

**Introdução:** Apesar de prevenível e curável, a Tuberculose (TB) é um problema de saúde pública, com cerca de 10.8 milhões de novos casos e 1,25 milhão de mortes somente em 2023 (15.5/100,000 habitantes). É estimado que 1/4 da população esteja infectada com TB e cerca de 5 a 15% destes evoluam para doença em algum momento da vida. A Organização Mundial da Saúde (OMS) busca reduzir os casos em 80% até 2030, mas enfrenta desafios, incluindo limitações em testes diagnósticos e de predição para a progressão da infecção para a TB ativa. **Objetivos:** Identificar citocinas nos sobrenadantes do QuantiFERON que possam prever a progressão de infecção para TB ativa em contatos de pacientes com TB pulmonar, em coortes do Brasil e da Índia. **Métodos:** Estudo de caso-controle conduzido com dados da coorte prospectiva RePORT-Brazil, que seguiu 1930 contatos próximos de pacientes com TB por 24 meses, entre 2015 e 2021. 20 contatos IGRA positivos que progrediram para TB ativa foram pareados por sexo, idade e tempo de exposição à TB com 40 contatos IGRA positivos que não progrediram para doença. Medimos as citocinas nos sobrenadantes do QuantiFERON e utilizamos análises descritivas e de *random forest* para identificação de biomarcadores capazes de prever progressão para doença. Em seguida, avaliamos esses biomarcadores em uma coorte indiana de contactantes com 8 casos de progressão e 12 controles. **Resultados:** Na coorte brasileira identificamos IL-8, IL-10 e CCL3 como uma assinatura preditiva importante. A análise da curva ROC desta combinação apresentou AUC de 0,75 (IC 95%: 0,61-0,90), com sensibilidade de 80% (IC 95%: 0,62-0,88) e especificidade de 85% (IC 95%: 0,69-1,00),  $p=0,001$ . Na coorte indiana, o mesmo modelo apresentou AUC de 0,804 (IC 95%: 0,67-0,94), sensibilidade de 79% (IC 95%: 0,57-1,00) e especificidade de 85% (IC 95%: 0,69-1,00),  $p=0,01$ . **Conclusão:** Este estudo identificou uma nova assinatura que pode servir como

ferramenta para prever a progressão da infecção por TB em contatos próximos. Nossos achados atendem aos critérios da OMS para ferramentas preditivas de TB.

**Palavras-chave:** Tuberculose; biomarcadores; citocinas; QuantiFERON; Tuberculose ativa; Infecção por tuberculose

## ABSTRACT

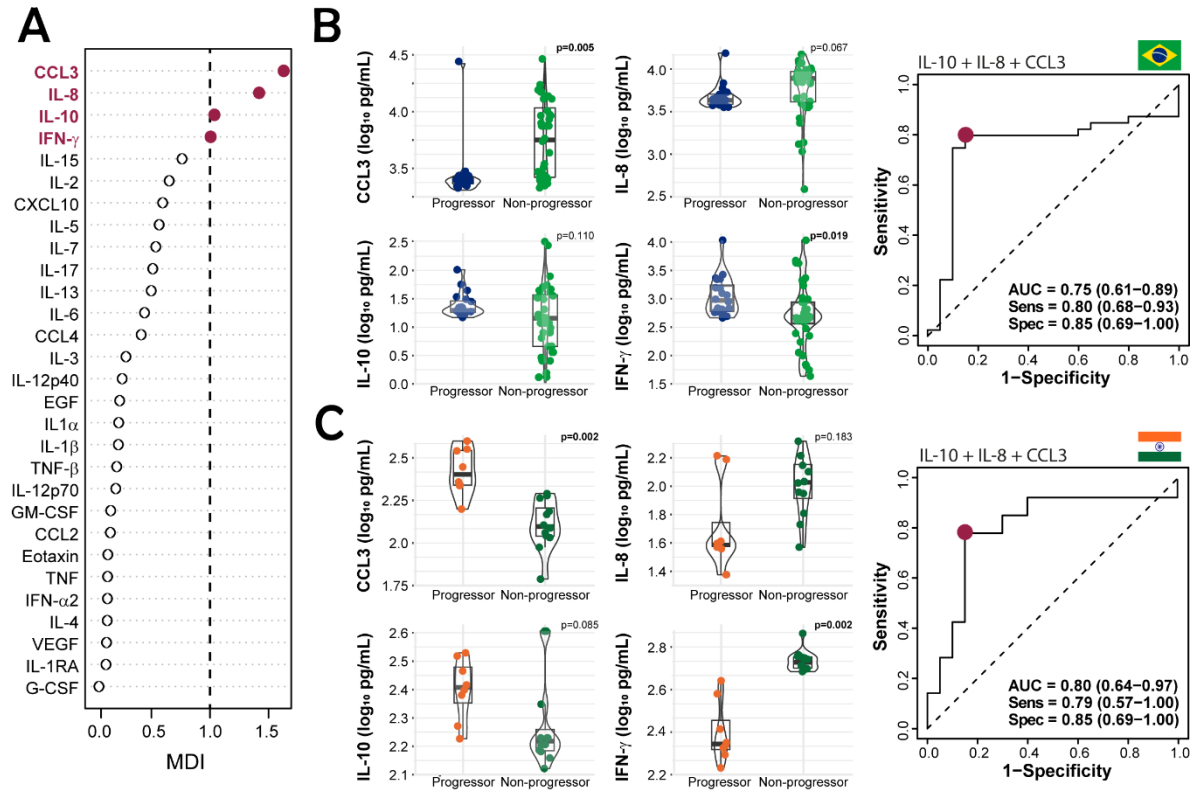
**Introduction:** Despite being preventable and curable, tuberculosis (TB) remains a public health problem, with approximately 10.8 million new cases and 1.25 million deaths in 2023 alone (15.5 per 100,000 population). It is estimated that one-quarter of the world's population is infected with TB, and about 5–15% of those infected will progress to active disease at some point in their lives. The World Health Organization (WHO) aims to reduce TB cases by 80% by 2030, but this goal faces challenges, including limitations in current diagnostic tests and in predicting progression from infection to active TB. **Objectives:** To identify cytokines in QuantiFERON supernatants that can predict progression from TB infection to active disease among close contacts of pulmonary TB patients in cohorts from Brazil and India. **Methods:** This was a nested case-control study conducted within the RePORT-Brazil prospective cohort, which followed 1,930 close contacts of pulmonary TB patients for 24 months between 2015 and 2021. Twenty IGRA-positive contacts who progressed to active TB were matched by sex, age, and exposure time with 40 IGRA-positive contacts who did not develop disease. Cytokines were measured in QuantiFERON supernatants, and descriptive statistics and random forest analysis were used to identify biomarkers predictive of disease progression. The identified biomarkers were then evaluated in an Indian cohort of TB contacts that included 8 progressors and 12 matched controls. **Results:** In the Brazilian cohort, IL-8, IL-10, and CCL3 were identified as a predictive biosignature. The ROC curve analysis of this combination yielded an AUC of 0.75 (95% CI: 0.61–0.90), with a sensitivity of 80% (95% CI: 0.62–0.88) and specificity of 85% (95% CI: 0.69–1.00),  $p=0.001$ . In the Indian cohort, the same model showed an AUC of 0.804 (95% CI: 0.67–0.94), sensitivity of 79% (95% CI: 0.57–1.00), and specificity of 85% (95% CI: 0.69–1.00),  $p=0.01$ . **Conclusion:** This study identified a novel

biosignature that may serve as a predictive tool for progression from TB infection to active disease among close contacts. Our findings meet WHO criteria for predictive TB tools.

**Key words:** Tuberculosis; biomarkers; cytokines; QuantiFERON; Active Tuberculosis; Tuberculosis infection

## FIGURAS E TABELAS

## FIGURES AND TABLES



**Figura 1. Identificação e Validação de uma Assinatura Imune Preditiva da Progressão da TB.** (A) O gráfico MDI classifica a importância das citocinas e quimiocinas na distinção entre progressores e não progressores. CCL3, IL-8, IL-10 e IFN- $\gamma$  se destacam como as variáveis mais importantes, com valores de MDI superiores a 1. (B) Os gráficos de violino mostram a distribuição dos níveis dessas citocinas (transformados em log) na coorte brasileira, comparando indivíduos que progrediram para TB com aqueles que não progrediram. A curva ROC demonstra o poder preditivo do modelo combinado com IL-8, IL-10 e CCL3 na coorte brasileira. (C) Avaliação da assinatura imune em uma coorte independente na Índia. Os gráficos de violino mostram a distribuição das mesmas citocinas, evidenciando um padrão semelhante de expressão diferencial e métricas ROC entre progressores e não progressores, comparável ao observado no Brasil.

**Figure 1. Identification and Validation of an Immune Signature Predicting TB Progression.** (A) The MDI plot ranks the importance of cytokines and chemokines in distinguishing progressors from non-progressors. CCL3, IL-8, IL-10, and IFN- $\gamma$  are highlighted as the most important variables with a MDI greater than 1. (B) Violin plots display the distribution of these cytokine

levels (log-transformed) in the Brazilian cohort, comparing TB progressors and non-progressors. The ROC curve demonstrates the predictive power of the combined IL-8, IL-10, and CCL3 model in the Brazilian cohort. (C) Evaluation of the immune signature in an independent Indian cohort. Violin plots display the distribution of the same cytokines, showing a similar pattern of differential expression and ROC metrics between progressors and non-progressors if compared to Brazil.

Tabela 1. Características clínicas e demográficas da população do estudo RePORT-Brazil.

Table 1. Clinical and demographic characteristics of the RePORT-Brazil study population

	<b>Progressors (n=20)</b>	<b>Non progressors (n=40)</b>	<b>P value</b>
<b>Age (years), median (IQR)</b>	26.3 (18.3-49.1)	26.4 (17.9-49.9)	0.856
<b>Sex (male), n (%)</b>	6 (30.0)	12 (30.0)	1.000
<b>Race, n (%)</b>			0.305
<b>White</b>	5 (25.0)	2 (5.00)	
<b>Black</b>	6 (30.0)	15 (37.5)	
<b>Brown</b>	9 (45.0)	22 (55.0)	
<b>Other</b>	0 (0.00)	1 (2.50)	
<b>Literacy, n (%)</b>	19 (95.0)	37 (92.5)	1.000
<b>Monthly Income, n (%)</b>			0.405
<b>Over minimal wage</b>	3 (15)	14 (35.0)	
<b>Minimal wage or less</b>	9 (45.0)	17 (42.5)	
<b>No income</b>	6 (30.0)	7 (17.5)	
<b>Missing</b>	2 (10)	2 (5)	
<b>BCG scar, n (%)</b>	17 (85.0)	37 (92.5)	0.575
<b>HIV status (positive), n (%)</b>	0 (0.00)	0 (0.00)	-
<b>Smoker, n (%)</b>	3 (15.0)	10 (25.0)	0.735

<i>Secondary smoking, n (%)</i>	8 (40.0)	17 (42.5)	1.000
<i>Alcohol consumption, n (%)</i>	6 (30.0)	22 (55.0)	0.289
<i>Illicit drug use, n (%)</i>	1 (5.00)	4 (10.0)	0.828
<i>Diabetes, n (%)</i>	0 (0.00)	4 (10.0)	0.453
<i>Hours of contact per month),median (IQR)</i>	23.0 (8.00-24.2)	18.0 (5.00-64.5)	0.83
<i>TB Preventive Therapy, n (%)</i>	5 (25.0)	20 (50.0)	0.06

Nota da tabela: Os valores são apresentados como n (%) para variáveis categóricas e mediana (intervalo interquartil, IQR) para variáveis contínuas. As comparações entre progressores e não progressores foram realizadas utilizando o teste de Mann-Whitney U para variáveis contínuas e o teste exato de Fisher para variáveis categóricas. Os valores de p indicam o nível de significância estatística das diferenças entre os grupos. BCG = Bacilo Calmette-Guérin; HIV = Vírus da Imunodeficiência Humana.

Table note: Values are presented as n (%) for categorical variables and median (interquartile range, IQR) for continuous variables. Comparisons between progressors and non-progressors were conducted using the Mann-Whitney U test for continuous variables and Fisher's exact test for categorical variables. p-values indicate the level of statistical significance for differences between the groups. BCG = Bacillus Calmette-Guérin; HIV = Human Immunodeficiency Virus;

Tabela Suplementar 2 – Características clínicas e demográficas dos casos-índice na coorte do estudo RePORT-Brasil.

Supplementary Table 2 - Clinical and demographic characteristics of the Index cases in the RePORT-Brazil study cohort.

	<b>Progressor Index Cases (n=20)</b>	<b>Non progressor Index Cases (n=40)</b>	<b>P value</b>
<b><i>Age (years), median (IQR)</i></b>	34.0 (22.0-52.2)	33.3 (24.0-44.2)	0.830
<b><i>Sex (male), n, n (%)</i></b>	14 (70.0)	24 (60.0)	1.000
<b><i>Race, n (%)</i></b>			1.000
<b>White</b>	3 (15.0)	6 (15.0)	
<b>Black</b>	8 (40.0)	12 (30.0)	
<b>Brown</b>	9 (45.0)	21 (52.5)	
<b>Other</b>	0 (0.00)	1 (2.50)	
<b><i>HIV status (positive), n (%)</i></b>	0 (0.00)	1 (2.50)	1.000
<b><i>Sputum smear result (%)</i></b>			0.897
<b>1 to 9</b>	3 (15.0)	7 (17.5)	
<b>+</b>	1 (5.00)	3 (7.50)	
<b>++</b>	3 (15.0)	12 (30.0)	
<b>+++</b>	8 (40.0)	5 (12.5)	
<b>Negative</b>	3 (15.0)	7 (17.5)	
<b><i>Cavitations on X-ray (%)</i></b>			1.00
<b>Yes</b>	13 (65.0)	22 (56.4)	
<b>No</b>	7 (35.0)	17 (43.6)	

Nota da tabela: Os valores são apresentados como n (%) para variáveis categóricas e mediana (intervalo interquartil, IQR) para variáveis contínuas. As comparações entre progressores e não progressores foram realizadas utilizando o teste de Mann-Whitney U para variáveis contínuas e o teste exato de Fisher para variáveis categóricas. Os valores de p indicam o nível de significância estatística das diferenças entre os grupos. BCG = Bacilo Calmette-Guérin; HIV = Vírus da Imunodeficiência Humana.

Table note: Values are presented as n (%) for categorical variables and median (interquartile range, IQR) for continuous variables. Comparisons between progressors and non-progressors were conducted using the Mann-Whitney U test for continuous variables and Fisher's exact test for categorical variables. p-values indicate the level of statistical significance for differences between the groups. HIV = Human Immunodeficiency Virus;

Tabela Suplementar 3 – Níveis medianos de citocinas (pg/mL) em progressores e não progressores (Brasil).

Supplementary Table 3 - Median levels cytokines (pg/mL) in progressors and non-progressors (Brazil)

	<b>Progressors (n=20)</b>	<b>Non progressors (n=40)</b>	<b>P value</b>
<b>EGF, median (IQR)</b>	128 (72.9-196)	107 (88.4-149)	0.973
<b>Eotaxin, median (IQR)</b>	385 (277-476)	325 (246-652)	1.000
<b>GCSF, median (IQR)</b>	106 (81.9-94)	98.6 (69.8-202)	0.775
<b>GM-CSF, median (IQR)</b>	38.0 (16.6-55.4)	18.2 (12.1-33.6)	0.156
<b>IFN-<math>\alpha</math>, median (IQR)</b>	63.6 (31.3-86.0)	58.1 (38.4-84.5)	1.000
<b>IFN-<math>\gamma</math>, median (IQR)</b>	970 (592-1961)	444 (335-879)	<b>0.019</b>
<b>IL-10, median (IQR)</b>	19.8 (18.3-29.0)	14.4 (4.60-37.1)	0.161
<b>IL-12p40, median (IQR)</b>	10.9 (8.55-16.0)	9.44 (6.10-11.0)	0.135
<b>IL-12p7, median (IQR)</b>	1.87 (0.27-2.62)	2.25 (1.63-4.52)	0.324
<b>IL-13, median (IQR)</b>	16.4 (8.04-31.3)	7.00 (1.49-11.2)	<b>0.021</b>
<b>IL-15, median (IQR)</b>	7.45 (6.46-9.99)	3.39 (2.55-7.92)	0.067
<b>IL-17, median (IQR)</b>	5.27 (3.77-6.93)	7.21 (5.29-14.5)	0.067
<b>IL-1-ra, median (IQR)</b>	758 (297-1363)	377 (255-663)	0.273
<b>IL-1a, median (IQR)</b>	93.8 (65.0-215)	60.7 (36.8-89.8)	0.090
<b>IL-1b, median (IQR)</b>	225 (136-677)	394 (310-667)	0.405
<b>IL-2, median (IQR)</b>	90.6 (59.3-208)	33.4 (25.7-84.9)	<b>0.021</b>
<b>IL-3, median (IQR)</b>	0.44 (0.27-0.85)	0.24 (0.14-0.44)	0.067
<b>IL-4, median (IQR)</b>	266 (239-333)	246 (203-276)	0.243
<b>IL-5, median (IQR)</b>	6.89 (5.82-10.4)	5.53 (3.70-7.54)	0.067
<b>IL-6, median (IQR)</b>	603 (417-2091)	898 (618-1327)	0.751
<b>IL-7, median (IQR)</b>	12.2 (9.94-14.4)	8.56 (6.70-12.1)	0.067
<b>IL-8, median (IQR)</b>	3786 (3362-4569)	7368 (3618-9090)	0.067
<b>CXCL-10, median (IQR)</b>	12925 (10773-15906)	11246 (8579-13110)	0.067
<b>CCL-2, median (IQR)</b>	1292 (1102-1775)	1447 (1037-2144)	1.000
<b>CCL-3, median (IQR)</b>	1778 (1626-1910)	4475 (1892-9401)	<b>0.005</b>
<b>CCL-4, median (IQR)</b>	1932 (1470-17180)	1557 (1401-1775)	0.090
<b>TNF, median (IQR)</b>	673 (244-1039)	470 (165-861)	0.405
<b>TNF-<math>\beta</math>, median (IQR)</b>	148 (143-164)	165 (146-204)	0.325
<b>VEGF, median (IQR)</b>	176 (131-216)	176 (108-311)	1.000

Nota da tabela: Os dados são apresentados como mediana e intervalo interquartil (IQR) ou frequência (percentual). Os dados categóricos foram comparados entre os grupos clínicos utilizando o teste do qui-quadrado. Os dados contínuos foram comparados entre os grupos clínicos utilizando o teste de Mann-Whitney U. Fonte em negrito e letras na coluna “valor de p” indicam significância estatística ( $p < 0,05$ ).

Table note: Data are shown as median and interquartile (IQR) range or frequency (percentage).  
Categorical data were compared between the clinical groups using the Chi-squared tests.  
Continuous data were compared between the clinical groups using the Mann-Whitney U test.  
Bold font and letters in “p value” column indicate statistical significance ( $p < 0.05$ ).

Tabela Suplementar 4 – Análise ROC de cada marcador e de combinações na coorte brasileira para distinguir progressores e não progressores.

Supplementary Table 4 – ROC analysis for each marker and combinations in the Brazilian cohort to distinguish progressors and non-progressors.

	AUC (95% CI)	P value	Sensitivity (95% CI)	Specificity (95% CI)
<b>IL-10</b>	0.35 (0.21-0.49)	0.973	0.35 (0.20-0.50)	0.80 (0.63-0.98)
<b>IFN-<math>\gamma</math></b>	0.25 (0.12-0.37)	0.999	0.10 (0.01-0.19)	0.95 (0.85-1.00)
<b>IL-8</b>	0.69 (0.55-0.84)	<b>0.008</b>	0.65 (0.50-0.80)	0.95 (0.85-1.00)
<b>CCL3</b>	0.79 (0.67-0.91)	<b>&lt;0.001</b>	0.65 (0.50-0.80)	0.90 (0.77-1.00)
<b>IL-10+IFN-<math>\gamma</math></b>	0.25 (0.13-0.37)	0.999	0.10 (0.01-0.19)	0.95 (0.85-1.00)
<b>IL-10+IL-8</b>	0.69 (0.55-0.84)	<b>0.007</b>	0.65 (0.50-0.80)	0.95 (0.85-1.00)
<b>IL-10+CCL3</b>	0.77 (0.65-0.90)	<b>&lt;0.001</b>	0.63 (0.48-0.78)	0.95 (0.85-1.00)
<b>IL-8+CCL3</b>	0.75 (0.61-0.89)	<b>0.001</b>	0.80 (0.68-0.92)	0.85 (0.69-1.00)
<b>IL-10+IFN-<math>\gamma</math> +IL-8</b>	0.64 (0.48-0.79)	0.958	0.68 (0.53-0.82)	0.90 (0.77-1.00)
<b>IL-10+IFN-<math>\gamma</math> +CCL3</b>	0.58 (0.43-0.73)	0.839	0.53 (0.37-0.68)	0.90 (0.77-1.00)
<b>IL8+IL-10+CCL3</b>	0.75 (0.61-0.89)	<b>0.001</b>	0.80 (0.62-0.88)	0.85 (0.69-1.00)
<b>IFN-<math>\gamma</math> +IL-8+CCL3</b>	0.71 (0.56-0.86)	<b>0.004</b>	0.75 (0.62-0.88)	0.85 (0.69-1.00)
<b>IL-10+IFN-<math>\gamma</math>+IL-8+CCL3</b>	0.29 (0.14-0.44)	0.997	1.00 (1.00-1.00)	0.10 (0.07-0.19)

Tabela Suplementar 5 – Características clínicas e demográficas da coorte de validação (Índia).

Supplementary Table 5 - Clinical and demographic characteristics of the validation cohort (India)

	<b>Progressors (n=8)</b>	<b>Non progressors (n=12)</b>	<b>P value</b>
<b>Age (years), median (IQR)</b>	32.0 (24.0-36.2)	26.0 (23.0-37.0)	0.907
<b>Sex (male), n (%)</b>	3 (37.5)	6 (50.0)	0.804
<b>Literacy, n (%)</b>	19 (95.0)	37 (92.5)	1.000
<b>HIV status (positive), n (%)</b>	1 (12.5)	0 (0.00)	0.730
<b>Diabetes, n (%)</b>	0 (0.00)	0 (0.00)	-

Nota da tabela: Os valores são apresentados como n (%) para variáveis categóricas e mediana (intervalo interquartil, IQR) para variáveis contínuas. As comparações entre progressores e não progressores foram realizadas utilizando o teste de Mann-Whitney U para variáveis contínuas e o teste exato de Fisher para variáveis categóricas. Os valores de p indicam o nível de significância estatística das diferenças entre os grupos. BCG = Bacilo Calmette-Guérin; HIV = Vírus da Imunodeficiência Humana.

Table note: Values are presented as n (%) for categorical variables and median (interquartile range, IQR) for continuous variables. Comparisons between progressors and non-progressors were conducted using the Mann-Whitney U test for continuous variables and Fisher's exact test for categorical variables. p-values indicate the level of statistical significance for differences between the groups. HIV = Human Immunodeficiency Virus;

Tabela Suplementar 6 – Análise ROC de cada marcador e de combinações na coorte indiana para distinguir progressores e não progressores.

Table 6 – ROC analysis for each marker and combinations in the Indian cohort to distinguish progressors and non-progressors.

	AUC (95% CI)	P value	Sensitivity (95% CI)	Specificity (95% CI)
<b>IL-10</b>	0.78 (0.49-0.89)	0.020	0.88 (0.65-1.00)	0.75 (0.50-0.99)
<b>IFN-<math>\gamma</math></b>	0.01 (0.00-0.01)	1.000	1.00 (1.00-1.00)	0.00 (0.00-0.00)
<b>IL-8</b>	0.25 (0.01-0.52)	1.000	0.13 (0.00-0.30)	0.92 (0.76-1.00)
<b>CCL3</b>	0.97 (0.90-1.00)	<b>&lt;0.001</b>	0.88 (0.65-1.00)	1.00 (1.00-1.00)
<b>IL-10+IFN-<math>\gamma</math></b>	0.06 (0.00-0.15)	1.000	1.00 (1.00-0.59)	0.00 (0.00-0.00)
<b>IL-10+IL-8</b>	0.56 (0.29-0.83)	0.688	0.88 (0.65-0.95)	0.42 (0.14-0.70)
<b>IL-10+CCL3</b>	0.93 (0.81-1.00)	<b>&lt;0.001</b>	0.93 (0.79-1.00)	0.83 (0.63-0.98)
<b>IL-8+CCL3</b>	0.73 (0.54-0.93)	<b>0.011</b>	0.71 (0.48-0.95)	0.80 (0.63-0.98)
<b>IL-10+IFN-<math>\gamma</math>+IL-8</b>	0.06 (0.00-0.17)	1.000	1.00 (1.00-1.00)	0.00 (0.00-0.00)
<b>IL-10+IFN-<math>\gamma</math>+CCL3</b>	0.27 (0.09-0.45)	0.990	1.00 (1.00-1.00)	0.00 (0.00-0.00)
<b>IL8+IL-10+CCL3</b>	0.80 (0.67-0.94)	<b>0.001</b>	0.79 (0.51-1.00)	0.85 (0.69-1.00)
<b>IFN-<math>\gamma</math>+IL-8+CCL3</b>	0.08 (0.00-0.16)	1.000	1.00 (1.00-1.00)	0.00 (0.00-0.00)
<b>IL-10+IFN-<math>\gamma</math>+IL-8+CCL3</b>	0.19 (0.03-0.35)	0.999	1.00 (1.00-1.00)	0.00 (0.00-0.00)