European Journal of Cancer 120 (2019) 97–106

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Prognostic value of transcriptomic determination of tumour-infiltrating lymphocytes in localised breast cancer

Caroline Truntzer [a,b,*], Nicolas Isambert [c], Laurent Arnould [c,d], Sylvain Ladoire [a,b,c,e,f], Francois Ghiringhelli [a,b,c,e,f]

[a] *Platform of Transfer in Biology of Cancer, Georges Francois Leclerc Cancer Center, 1 Rue Du Professeur Marion, 21000 Dijon, France*
[b] *GIMI Genetic and Immunology Medical Institute, 21000 Dijon, France*
[c] *Department of Medical Oncology, Georges Francois Leclerc Cancer Center, 1 Rue Du Professeur Marion, 21000 Dijon, France*
[d] *Department of Pathology, Georges Francois Leclerc Cancer Center, 1 Rue Du Professeur Marion, 21000 Dijon, France*
[e] *INSERM U1231, Dijon, France*
[f] *Univ. Bourgogne Franche-Comté, Dijon, France*

Received 19 April 2019; received in revised form 16 July 2019; accepted 17 July 2019
Available online 6 September 2019

KEYWORDS
Breast;
Biostatistics;
Immunology;
Tumour-infiltrating lymphocytes;
Gene expression signatures

Abstract ***Purpose:*** Tumour-infiltrating lymphocyte (TIL) detection by histology is associated with outcomes in breast cancer; nevertheless, analysis standardisation is difficult. We determined whether transcriptomic data could generate a genomic signature that estimated TIL infiltrates and determined patient prognosis in localised breast cancer.

Experimental design: Using 1928 transcriptomic profiles of pure cells, we generated a genetic signature specific to lymphocyte, myeloid, stromal and cancer cells. We then computed a score based on this signature and tested the association between the score and the TILs estimated for patients in an adjuvant setting from public and private data sets. We tested the capacity of the transcriptomic RNA TIL score to predict disease-free survival (DFS) or overall survival (OS) through multivariate Cox models adjusted for classical clinical variables and PAM50 molecular classification in two public data sets (Carte d'Identité des Tumeurs [CIT], n = 530; Metabric, n = 1832).

Results: A high RNA TIL score was significantly associated with the presence of a high level of TILs as assessed by histology. The score was also associated with DFS and OS in multivariate Cox models adjusted for molecular and clinical variables (CIT: OS hazard ratio [HR] = 0.15 [0.04, 0.61], p-value = 0.007; DFS: 0.27 [0.08, 0.8] p-value = 0.02; Metabric: OS HR = 0.87 [0.77, 0.97], p-value = 0.01). The association between the RNA TIL score

* *Corresponding author*: Georges-Francois Leclerc Cancer Center, 1 rue Professeur Marion, Dijon 21000, France. Fax: +33 3 80 39 34 34.
E-mail addresses: ctruntzer@cgfl.fr (C. Truntzer), nisambert@cgf.fr (N. Isambert), larnould@cgfl.fr (L. Arnould), sladoire@cgfl.fr (S. Ladoire), fghiringhelli@cgfl.fr (F. Ghiringhelli).

https://doi.org/10.1016/j.ejca.2019.07.020

and survival was tested by univariate analysis in each molecular subgroup; the RNA TIL score was associated with survival only in basal-like tumours.
Conclusions: Determination of the TIL rate using a transcriptomic signature is feasible and has a high prognostic value in patients with basal-like tumours in an adjuvant setting.

1. Introduction

Intratumoural immune response is associated with prognosis in various types of cancer. Most studies suggest that immune infiltrates with high densities of CD3 and CD8 T cells are associated with a better outcome [1]. Currently, there is an increasing demand that quantitative histopathology is required to better assess prognosis in many cancer types [2,3]. Tumour-infiltrating lymphocytes (TILs) scored on haematoxylin and eosin slides (HESs) or after immunohistological staining were shown to have prognostic or therapeutic relevance in breast cancer, non–small-cell lung cancer and melanoma [2,3]. In breast cancer, evaluation of TILs on HESs is prognostic in triple-negative (TN) cancers, but their role in luminal tumours is less clear [4]. Conflicting results have been observed for human epidermal growth factor receptor 2 (HER2)–positive tumours [5,6]. Guidelines by the International TIL working group were published to standardise the technology [7–9], and the robustness of this technique was evaluated in international trials [10]. Histological assessment requires pathologists' course formation and currently, some online tools are available to help pathologists performing such analysis (https://www.tilsinbreastcancer.org/). This is essential to avoid interobserver variability. Furthermore, TIL evaluation on HES sections does not allow differentiation of lymphoid from myeloid subsets, suggesting that quantitative and automatic evaluation might be more accurate and avoid repetitive work by the pathologist.

We aimed to determine a score that would reflect the presence of distinct underlying cell populations invading breast tumour tissue using transcriptomic measurements. We generated a transcriptomic signature that could mirror TILs, named RNA TIL score. Using two large public cohorts of patients with localised breast cancer, we tested the capacity of the RNA TIL score to predict the disease outcome.

2. Materials and methods

2.1. Patients and cohorts

Several data sets were used in this study, first to construct immune cell signatures and then to compute and validate corresponding RNA TIL scores in patient samples.

2.1.1. Signature data sets

In total, 1928 transcriptomic profiles (Supplementary Table S1) of pure immune cells (lymphocyte and myeloid cells), stromal cells (endothelial cells, adipocytes and fibroblasts) and breast cancer cells were all retrieved from the public repository Gene Expression Omnibus.

2.1.2. Microarray data sets

2.1.2.1. Carte d'Identité des Tumeurs data set. It consisted of the breast cohort [11] from the 'Cartes d'Identité des Tumeurs®' (CIT), including 530 patients for whom overall survival (OS) and disease-free survival (DFS) were available, as well as clinical and biological parameters.

2.1.2.2. Metabric data set. It consisted of both discovery and validation Metabric [12] data sets (n = 1832 patients), generated by the Molecular Taxonomy of Breast Cancer International Consortium. OS as well as clinical and biological parameters was available.

Description of the RNA sequencing (RNAseq) data sets and microarray preprocessing methods are available as Supplementary Data.

2.2. Generation of the RNA TIL immune signature

Breast tumour samples constitute 4 major cellular populations: lymphoid, myeloid, stromal and cancer cells.

To select genes characteristic of each cellular type, we used the barcode approach [13] that provides absolute gene expression estimations. In this way, each gene is described by binary values, where 1 denotes expressed genes and 0 denotes unexpressed genes. Only probesets present in at least 80% of the samples and present in less than 20% of the other cell populations were retained as characteristics of a cell population. In this way, 11 probesets were selected for lymphoid cells, 24 for myeloid cells, 13 for stromal cells and 35 for cancer cells. Reducing the number of probesets used for the signature to 10 probesets by cell type did not change the classification (Supplementary Table S2); these 10 probesets were selected as those that best differentiated each cell type from the other three. Using principal component analysis, performed on the 1928 samples of pure cell

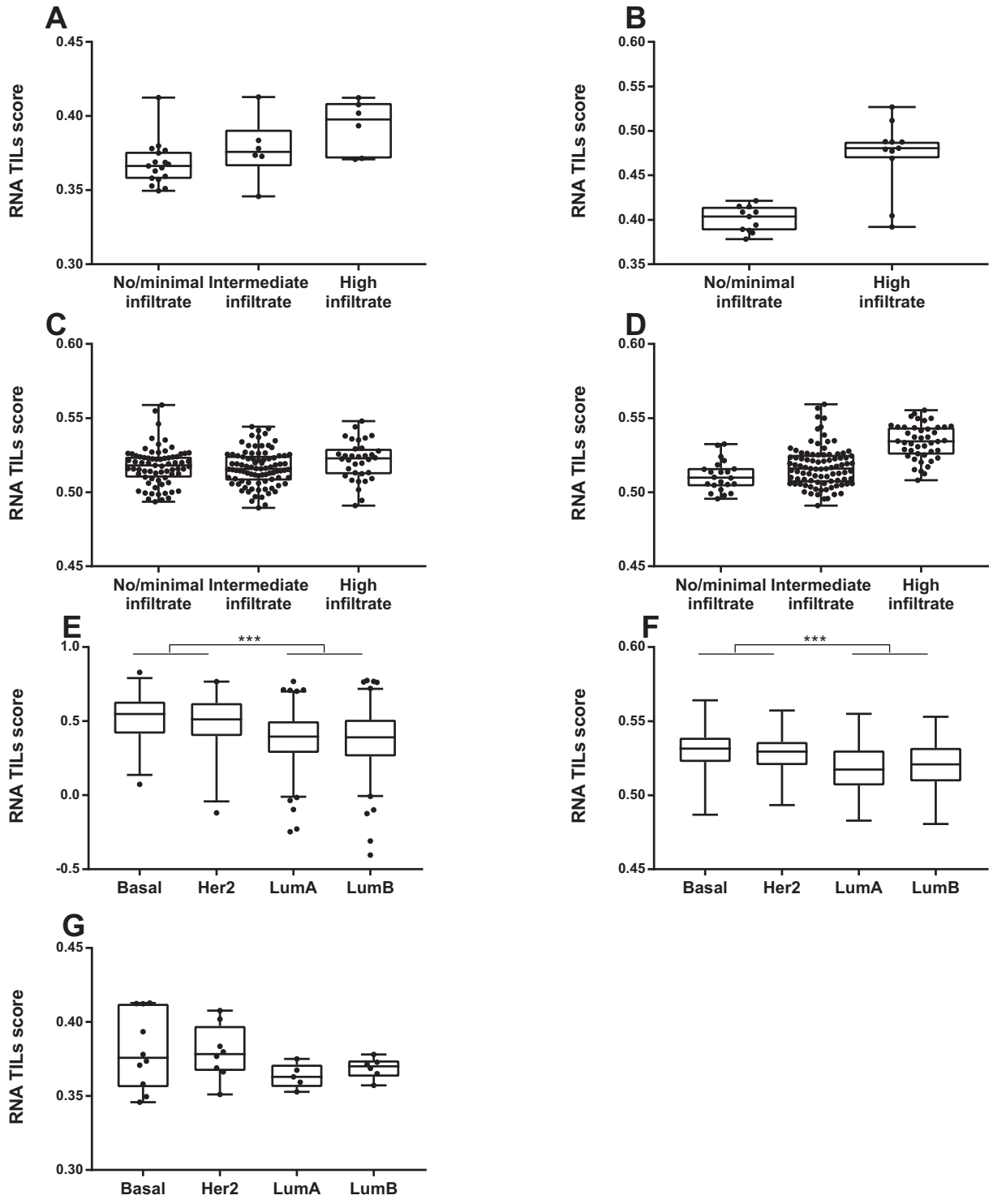

Fig. 1. **Association between the RNA TIL score and TILs determined by histology and pathological characteristics**. (A–D) Boxplots showing the RNA TIL score, given the proportions of TILS determined concordantly by 2 pathologists and categorised into 3 categories after the Salgado classification in the TIL RNAseq (A), the FinHer (C) and the Metabric (D) data sets and categorised into 2 categories (<3: no/minimal and intermediate infiltrate; ≥3: high immune infiltrate) in the TCGA data set (B). (E–G) Boxplots representing the distribution of the RNA TIL score, given the molecular subtype of breast cancer for the CIT (E), Metabric (F) and RNAseq TIL (G) data sets. RNA TIL scores were scaled to have null mean and unit variance. ***$p < 0.05$ Wilcoxon test. TIL, tumour-infiltrating lymphocyte; CIT, Carte d'Identité des Tumeurs.

lines to visualise the variability contained in the data [14], we observed the capacity of our selected probesets to characterise each cellular type (Supplementary Fig. S1A). Complements about this strategy are described in Supplementary Data.

2.3. RNA TIL score

Scores for each of the four cell types were computed by averaging the z-scores obtained for selected probesets constituting the 4 cell-type signatures for each sample. Each sample is thus described by 4 values, each summarising one cellular subtype. We considered a final score defined as the ratio of the sum of lymphoid and myeloid scores to the total of the 4 cell-type scores. Hereafter, this score is called the RNA TIL score.

2.4. Statistical analysis

Patient and disease characteristics from the 2 cohorts were compared using the chi-square test for qualitative variables and the Wilcoxon test or analysis of variance (ANOVA) for continuous variables, as appropriate. Survival analysis was performed using the survival R library [15]. The prognostic value of RNA TIL signatures was tested using univariate or adjusted multivariate Cox proportional hazards models for DFS or OS. Survivors were censored after 120 months. Survival probabilities were estimated using the Kaplan–Meier method, and survival curves were compared using the log-rank test. Statistical analyses were performed using R software (http://www.R-project.org/), and graphs were drawn using GraphPad Prism version 7.03.

3. Results

3.1. Association between the RNA TIL signature and TIL enumeration by histology

To validate the relevance of our RNA TIL score, RNAseq was performed in 29 breast cancers (5 Luminal A, 6 Luminal B, 8 HER2 and 10 TN) coming from the Centre Georges-François Leclerc (Dijon). The proportions of TILs were independently determined by two pathologists using international recommendations, and concordance was observed for all evaluated slides (29 of 29). The molecular subtype of these patients was determined using PAM50. The RNA TIL score was strongly associated with pathological classification (ANOVA test; p-value = 0.025) (Fig. 1A). Using the RNA TIL score, we observed a non-significant trend towards greater infiltration of immune cells in HER2 tumours and basal-like breast cancer than in Luminal A and Luminal B tumours (Fig. 1G). A greater variability of the RNA TIL score was observed in HER2 and basal-like breast cancer, suggesting a wider immunological heterogeneity than that previously reported using histological TIL enumeration [16,17].

These observations were also validated using public RNAseq data from the TCGA (The Cancer Genome Atlas) breast cancer cohort. In this cohort, TILs from 30 patients were enumerated by histology [18]. Similarly, the RNA TIL score was associated with the histologically enumerated TILs (ANOVA test. p-value$< 1.10^{-04}$) (Fig. 1B).

We have identified two public data sets containing array transcriptomic data and TIL evaluation by histology: the FinHer study, which includes 203 patients with HER2 breast cancer, and a subpopulation of 157 patients from the Metabric cohort. In the FinHer cohort also, the RNA TIL score was strongly associated with the histological score (ANOVA test; p-value$<1.10^{-13}$) (Fig. 1C). In the Metabric cohort, the RNA TIL score was also strongly associated with the histological score (ANOVA test; p-value$< 1.10^{-15}$) (Fig. 1D).

3.2. Distribution of RNA TIL scores according to molecular subgroups and clinicopathological variables

The distribution of the RNA TIL score according to clinical and pathological variables was evaluated in the CIT and Metabric data sets. For the first, the score was significantly different for oestrogen receptor (ER) and HER2 status, age, Scarff-Bloom-Richardson grade (1–2 vs 3) and nodal status. For the latter, age and nodal status were not related to the RNA TIL score. In both data sets, tumour size was not related to the RNA TIL score (Table 1).

Table 1
RNA TIL score according to clinical and molecular variables.

Variable	CIT		Metabric	
	Mean (SD)	t-test, p-value	Mean (SD)	t-test, p-value
ER				
+	0.383 (0.184)	$<1.10^{-7}$	0.520 (0.013)	$<1.10^{-16}$
−	0.493 (0.170)		0.530 (0.014)	
HER2				
+	0.507 (0.164)	0.0003	0.526 (0.014)	0.0002
−	0.409 (0.184)		0.522 (0.014)	
Age				
>65 years	0.340 (0.199)	$<1.10^{-4}$	0.523 (0.014)	0.21
≤65 years	0.428 (0.178)		0.522 (0.014)	
Tumour size				
T1	0.403 (0.195)	0.55	0.523 (0.015)	0.72
T2-3	0.392 (0.194)		0.522 (0.014)	
SBR grade				
1–2	0.378 (0.186)	$<1.10^{-5}$	0.520 (0.013)	$<1.10^{-11}$
3	0.453 (0.182)		0.525 (0.015)	
Nodal status				
+	0.352 (0.196)	0.002	0.520 (0.015)	0.09
−	0.441 (0.190)		0.522 (0.013)	

SBR, Scarff-Bloom-Richardson; ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2.

3.3. RNA TIL score (mean [standard deviation]) according to clinical and pathological variables evaluated in the CIT and metabric data sets

The distribution of the RNA TIL score was then evaluated, given the molecular intrinsic subtypes. In both cohorts, we observed that RNA TIL scores were significantly higher in HER2 and basal tumours than in Luminal A and B tumours (Fig. 1E and F).

We estimated the association between the RNA TIL score as a log2-transformed continuous variable and the standard clinicopathological variables, using univariate and multivariate linear regression models. In the CIT cohort, we observed by univariate analysis that higher RNA TIL values were significantly associated with higher grade, HER2 status, presence of positive lymph nodes, absence of hormone receptors and younger age (Fig. 2A). In the multivariate model, higher RNA TIL values were significantly associated with absence of hormone receptors and younger age; a strong trend was observed for presence of positive lymph nodes and HER2 status.

In the Metabric cohort, we observed by univariate analysis that higher RNA TIL values were significantly associated with absence of positive lymph nodes, higher grade, HER2 status and absence of hormone receptors (Fig. 2B). In the multivariate model, higher RNA TIL values were significantly associated with higher grade, absence of positive lymph nodes, absence of hormone receptors and older age.

Together, these data demonstrate that a high RNA TIL score is mainly associated with hormone-negative status.

3.4. Determination of patient prognosis based on the RNA TIL score

The association between the RNA TIL score and both DFS and OS was evaluated in the CIT and Metabric cohorts. Only OS was available for the Metabric data set. Clinical characteristics of the patients are summarised in Table 2. Characteristics of patients between the 2 cohorts were compared using chi-square or Wilcoxon tests, given that variables are qualitative or quantitative; the corresponding p-values are presented in the last column of Table 2.

The prognostic value of the RNA TIL score was first assessed in the CIT cohort (n = 530), for which both DFS and OS are available (Table 3 and Supplementary Table 3).

On univariate analysis, we found that the RNA TIL score was not associated with the outcome in the whole population but was significantly associated with outcome after exclusion of Luminal A tumours. On multivariate analysis, using a Cox proportional hazards model, it was found that the RNA TIL score was significantly associated with better DFS and OS (hazard ratio [HR] for a one-unit increase in the value of the RNA TIL score for DFS: 0.27 [0.08, 0.8], p-value = 0.02 and for OS: 0.15 [0.04, 0.61], p-value = 0.007). The RNA TIL score added significant prognostic information to the clinicopathological characteristics at diagnosis, as shown by the likelihood-ratio test (P = 0.008 for OS and 0.02 for DFS). This test enables the comparison of two nested survival models. The two models tested here are the multivariate survival model with all clinicopathological characteristics alone and the survival model including clinicopathological characteristics and the RNA TIL score, as described in Table 3 for DFS and Supplementary Table S3 for OS. The same conclusions were achieved based on standard pathology classification (Supplementary Tables S3–5).

These observations were validated on the Metabric data set (n = 1832). By univariate analysis, the RNA TIL score was associated with outcome only when luminal A tumours were excluded (HR = 0.88 [0.80, 0.97], p-value = 0.012). By multivariate analysis, the RNA TIL score was significantly associated with better OS (HR: 0.87 [0.77, 0.97], p-value = 0.01) (Table 4).

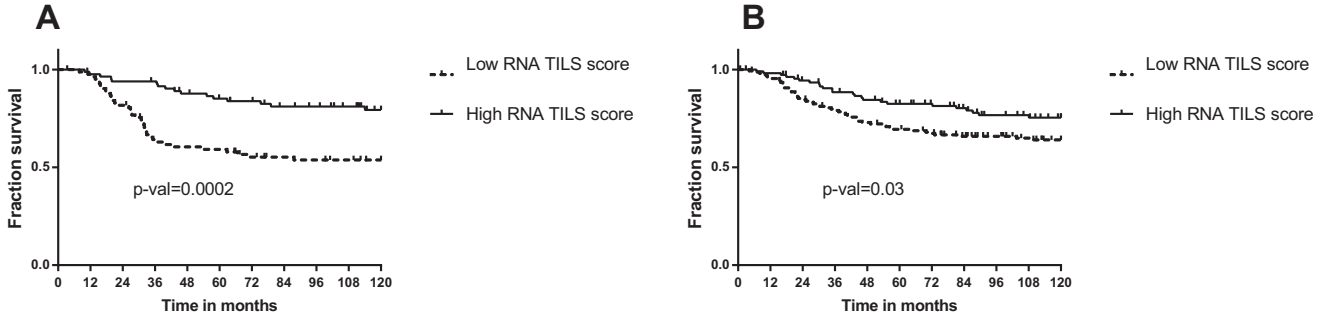

Fig. 2. **Prognostic role of RNA TILs in the different molecular subgroups for the Metabric data set**. Kaplan–Meier curves for overall survival (OS), with patients stratified according to their RNA TIL signature: low vs high score. (A) For the discovery Metabric data set, the two groups were determined using the optimal threshold of the RNA TIL score in patients with basal-like tumours. (B) For the validation Metabric data set, the two groups were determined using the threshold determined in the discovery data set. TIL, tumour-infiltrating lymphocyte.

Table 2
Patients' clinical characteristics.

Variable	CIT n (%)	Metabric n (%)	p-value
Sample size	530	1832	
Median DFS (years)	5.2	–	$<1.10^{-16}$
Median OS (years)	6.2	9.6	
Oestrogen receptor			0.10
Positive	401 (76)	1381 (75)	
Negative	107 (20)	451 (25)	
Missing	22 (4)	–	
HER2			1
Positive	49 (9)	233 (13)	
Negative	335 (63)	1599 (87)	
Missing	146 (28)		
Molecular subtype			$<1.10^{-16}$
Luminal A	214 (40)	700 (38)	
Luminal B	206 (39)	475 (26)	
HER2-enriched	52 (11)	224 (12)	
Basal-like	58 (10)	427 (23)	
Missing	–	6 (1)	
T size			$<1.10^{-4}$
T1	138 (26)	794 (43)	
T2–4	290 (55)	1022 (56)	
Missing	102 (19)	16 (1)	
Median age	55	62.4	
Tumour grade			$<1.10^{-6}$
1	46 (9)	157 (9)	
2	267 (50)	692 (38)	
3	197 (37)	904 (49)	
Missing	20 (4)	79 (4)	
Nodal status			$<1.10^{-16}$
N0	136 (26)	1238 (67)	
N+	296 (56)	121 (7)	
Missing	98 (18)	473 (26)	
Treatment type	NA		
Chemotherapy		377 (20)	
Hormone therapy		960 (52)	
No treatment		495 (28)	

DFS, disease-free survival; OS, overall survival; CIT, Carte d'Identité des Tumeurs; HER2, human epidermal growth factor receptor 2; NA, not available
p-values correspond to chi-square or Wilcoxon tests, given that variables are qualitative or quantitative.

The RNA TIL score added significant prognostic information to the clinicopathological characteristics at diagnosis, as shown by the likelihood-ratio test ($P = 0.015$) comparing the full multivariate survival model of Table 4 and the reduced model without the RNA TIL score.

To assess these results, the Metabric data set was split into one discovery (n = 957) and one validation (n = 875) data set, as described by the Molecular Taxonomy of Breast Cancer International Consortium. In patients with basal-like tumours from the discovery data set, two risk groups (low vs high) were determined of the RNA TIL score estimated in these molecular subgroups, based on the maximally selected rank statistics [19]. The corresponding Kaplan–Meier OS curves underline that the high RNA TIL score was associated with better prognosis in patients with basal-like tumours (HR = 0.34 [0.19, 0.61], p-value $<1.10^{-3}$ (Fig. 3A). The optimal threshold computed on the discovery data set was then applied to the validation data set to generate a low- and a high-risk group. Fig. 3B shows Kaplan–Meier curves obtained using these groups. The association between the RNA TIL score was validated with HR = 0.59 (0.37; 0.96), p-value = 0.0328. By this way, we were able to validate the association between overall survival and our RNA TIL score.

The effect of the RNA TIL score was also analysed according to the type of adjuvant therapy (see Supplementary Data).

In a recent article, Loi *et al.* [20] demonstrated the prognostic value of stromal TILs in TN breast cancer. To validate our method, we confirmed the findings by Loi *et al.* [20] using the RNA TIL score on patients with basal-like tumours from the Metabric data set (n = 189 patients). We showed that the RNA TIL score dichotomised by the method used by Lausen *et al.* [29] adds value in comparison with clinical variables only (likelihood-ratio test chi-square, 4.495; p-value = 0.03) (see Supplementary Data).

4. Discussion

This study shows that a high RNA TIL score was significantly associated with the presence of a high level of TILs assessed by histology. The RNA TIL score was also associated with DFS and OS in multivariate Cox models adjusted for molecular and clinical variables in two public data sets.

Accurate prediction of relapse risk in early breast cancer is a major challenge because it might avoid excessive usage of adjuvant chemotherapy and allow the identification of patients at a high risk of relapse with a strong therapeutic medical need. Although recurrence risk has been assessed by examining histological data and biomarkers (ER, progesterone receptor (PR), HER2 and Ki67), these conventional examinations are not accurate enough to select the subset of patients who are at sufficiently low risk of recurrence to be spared from adjuvant chemotherapy without comprising prognosis. Several recent clinical studies have revealed that TILs can predict prognosis and chemotherapy response, independently of conventional clinical and pathological criteria such as age, nodal status or tumour size [21–23]. In TN breast cancers, a meta-analysis revealed that high levels of TILs were significantly associated with favourable outcomes [10,24]. New technology such as automated image analysis methods or determination of immune infiltrates using transcriptomic data can improve quantification accuracy. In this context, it was interesting to determine whether RNA TIL determination could be combined with histological evaluation. A previous work tested this hypothesis in a neoadjuvant setting and underlined that the TIL signature was predictive of response to neoadjuvant

Table 3
Univariate and multivariate Cox model for CIT data set for disease-free survival and PAM50 classification.

Variable	Univariate			Multivariate		
	HR	95% CI	P-value	HR	95% CI	P-value
PAM50						
Luminal A	1			1		
Basal-like	2.71	1.54, 4.76	0.0005	1.55	0.68, 3.53	0.29
HER2	2.47	1.38, 4.42	0.002	1.42	0.63, 3.21	0.39
Luminal B	1.97	1.26, 3.10	0.003	2.09	1.28; 3.39	0.003
Age > 65 years	1.03	0.67, 1.58	0.89	1.05	0.65; 1.70	0.85
Tumour size						
T1	1			1		
T2–4	2.13	1.34, 3.40	0.006	1.85	1.34, 3.02	0.01
SBR grade						
Grade 1–2	1			1		
Grade 3	1.91	1.32, 2.74	0.0005	1.51	0.78, 1.98	0.35
ER+	0.51	0.35, 0.75	0.001	0.46	0.24, 0.85	0.01
Nodal status N ≥ 1	1.40	0.92, 2.13	0.10	1.60	1.09, 2.63	0.02
RNA TIL score						
Whole population	0.71	0.27, 1.83	0.48	0.27	0.08, 0.8	0.02
Population without Luminal A	0.24	0.08, 0.72	0.01			
Luminal A	5.43	0.63, 46.40	0.12	–	–	–
Luminal B	0.32	0.07, 1.42	0.13			
HER2	0.007	0.002, 0.22	0.004			
Basal-like	0.01	0.0003, 0.27	0.006			

HR, hazard ratio; CI, confidence interval; HER2, human epidermal growth factor receptor 2; SBR, Scarff-Bloom-Richardson; ER, oestrogen receptor; TIL, tumour-infiltrating lymphocyte.

chemotherapy but not a prognostic factor in untreated patients or patients treated with hormone therapy [25]. Our study shows, in an adjuvant setting, that TIL evaluation using transcriptomic data highlights greater TIL accumulation in basal and HER2 tumours. RNA TILs are a good biomarker to improve the estimation of survival in basal and HER2 tumour types. Most of the previous studies have evaluated the prognostic value of TILs in the context of randomised adjuvant trials for breast cancer. The results indicated

Table 4
Univariate and multivariate Cox model for Metabric data set for overall survival and PAM50 classification.

Variable	Univariate			Multivariate		
	HR	95% CI	P-value	HR	95% CI	P-value
PAM50						
Luminal A	1			1		
Basal	2.49	1.92, 3.23	$<1.10^{-12}$	1.66	1.07, 2.58	0.02
HER2	3.68	2.78, 4.86	$<1.10^{-12}$	2.31	1.49, 3.57	$<1.10^{-3}$
Luminal B	2.46	1.91, 3.16	$<1.10^{-12}$	1.92	1.43, 2.60	$<1.10^{-4}$
Age > 65 years	1.05	0.88, 1.26	0.57	1.17	0.94, 1.46	0.16
Tumour size						
T1	1			1		
T2–4	1.92	1.58, 2.33	$<1.10^{-12}$	1.43	1.13, 1.81	0.003
SBR grade						
Grade 1–2	1			1		
Grade 3	2.08	1.70, 2.82	$<1.10^{-12}$	1.51	1.17, 1.94	0.001
ER +	0.48	0.40, 0.58	$<1.10^{-12}$	0.71	0.48, 1.03	0.07
Nodal status N ≥ 1	3.3	2.53, 4.31	$<1.10^{-16}$	2.50	1.88, 3.32	$<1.10^{-9}$
RNA TIL score						
Whole population	0.97	0.88, 1.07	0.52	0.87	0.77;0.97	0.01
Population without Luminal A	0.89	0.81, 0.99	0.028			
Luminal A	0.85	0.68, 1.07	0.17	–	–	–
Luminal B	0.85	0.76, 1.09	0.18			
HER2	0.85	0.71, 1.09	0.26			
Basal-like	0.86	0.73, 1.01	0.07			

HR, hazard ratio; CI, confidence interval; TIL, tumour-infiltrating lymphocyte; HER2, human epidermal growth factor receptor 2; SBR, Scarff-Bloom-Richardson; ER, oestrogen receptor.

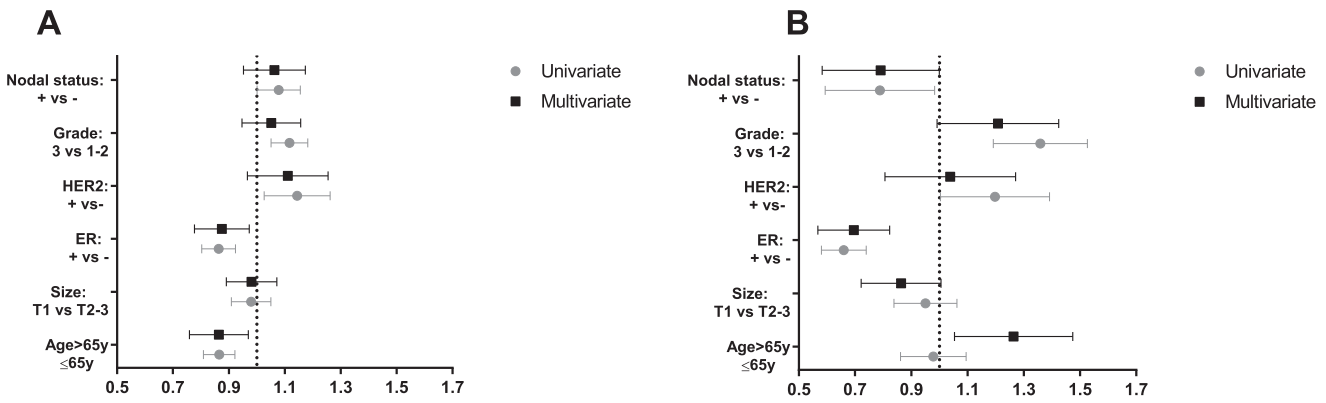

Fig. 3. **Association between the RNA TIL score and clinical characteristics**. Forest plots representing the coefficient of association between the RNA TIL score as a continuous log2-transformed variable and each of the standard clinicopathological variables, using the linear regression model for the CIT (A) and Metabric (B) data sets. Solid lines represent coefficients from univariate linear regression models, and dashed lines represent coefficients from the multivariate models. TIL, tumour-infiltrating lymphocyte; ER, oestrogen receptor; CIT, Carte d'Identité des Tumeurs.

that baseline TILs were a strong prognostic predictor for specific breast cancer subtypes, especially TN cancers [17,22]. These previous trials testing the TIL infiltrate were not always able to test the predictive value because all arms received adjuvant treatment. In our study, the RNA TIL score showed prognostic properties in HER2 and basal-like tumours. However, we cannot eliminate the possibility that such a result is biased owing to heterogeneity of data sets. To go further, we analysed patients with TN cancer from a neoadjuvant data set (GSE25055) [26] and showed that in this setting, the RNA TIL score was also associated with response to neoadjuvant chemotherapy using taxane/anthracycline, an up-to-date chemotherapy regimen (see Supplementary Data). These data add an additional argument to propose that the TIL signature may be predictive.

It is classically known that chemotherapies used for breast cancer may induce immune cell death and that such immune activation could contribute to the efficacy of adjuvant therapy [27]. Our data support the hypothesis that preexisting TIL accumulation is required for the efficacy of chemotherapy in such disease. Further studies are required to compare the ability of transcriptomic and histological TIL evaluation to address patient prognosis.

The present study has some limitations. First, the development and validation data sets were obtained from retrospective cohorts. Second, few clinical variables and little information on treatments were available owing to use of public data. In addition, it should be noted that the RNA TIL signature was associated with TIL histology assessment in only a small number of patients. One of the most significant strengths of this study is the validation of the observations in 2 different data sets. Moreover, our method can be applied independently of the technology used to access gene expression. Genes involved in the RNA TIL score computation could be used in a deconvolution strategy to quantify qualitative features of TILs and thus qualify the immune response. In another extent, multigene tests such as PAM50 are recently used in clinical practice to improve patient prognostic determination and to support adjuvant therapy decision [28]. Genes involved in the RNA TIL score could also be easily added to existing tests and could be an interesting additional tool to predict patient prognostic.

In conclusion, we believe that our RNA TIL signature is a simple tool to determine prognosis in patients with early breast cancer.

Acknowledgements

The authors wish to thank Fiona Caulfield and Isabel Gregoire for English correction and helpful comments.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Conflict of interest statement

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2019.07.020.

References

[1] Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol 2017;14(12):717–34. https://doi.org/10.1038/nrclinonc.2017.101.

[2] Mahmoud SM, Paish EC, Powe DG, Macmillan RD, Grainge MJ, Lee AH, et al. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. J Clin Oncol 2011;29(15):1949–55. https://doi.org/10.1200/JCO.2010.30.5037.

[3] Klauschen F, Müller K-R, Binder A, Bockmayr M, Hägele M, Seegerer P, et al. Scoring of tumor-infiltrating lymphocytes: from visual estimation to machine learning. Semin Cancer Biol 2018;52(Pt 2):151–7. https://doi.org/10.1016/j.semcancer.2018.07.001.

[4] Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. Lancet Oncol 2018;19(1):40–50. https://doi.org/10.1016/S1470-2045(17)30904-X.

[5] Loi S, Michiels S, Salgado R, Sirtaine N, Jose V, Fumagalli D, et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Ann Oncol Off J Eur Soc Med Oncol 2014;25(8):1544–50. https://doi.org/10.1093/annonc/mdu112.

[6] Ingold Heppner B, Loibl S, Denkert C. Tumor-infiltrating lymphocytes: a promising biomarker in breast cancer. Breast Care Basel Switz 2016;11(2):96–100. https://doi.org/10.1159/000444357.

[7] Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al. Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the international immuno-oncology biomarkers working group: Part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. Adv Anat Pathol 2017;24(6):311–35. https://doi.org/10.1097/PAP.0000000000000161.

[8] Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, et al. Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the international immunooncology biomarkers working group: Part 1: assessing the host immune response, TILs in invasive breast carcinoma and ductal carcinoma in situ, metastatic tumor deposits and areas for further research. Adv Anat Pathol 2017;24(5):235–51. https://doi.org/10.1097/PAP.0000000000000162.

[9] Criscitiello C, Esposito A, Trapani D, Curigliano G. Prognostic and predictive value of tumor infiltrating lymphocytes in early breast cancer. Cancer Treat Rev 2016;50:205–7. https://doi.org/10.1016/j.ctrv.2016.09.019.

[10] Denkert C, Wienert S, Poterie A, Loibl S, Budczies J, Badve S, et al. Standardized evaluation of tumor-infiltrating lymphocytes in breast cancer: results of the ring studies of the international immuno-oncology biomarker working group. Mod Pathol Off J U S Can Acad Pathol Inc 2016;29(10):1155–64. https://doi.org/10.1038/modpathol.2016.109.

[11] Guedj M, Marisa L, de Reynies A, Orsetti B, Schiappa R, Bibeau F, et al. A refined molecular taxonomy of breast cancer. Oncogene 2012;31(9):1196–206. https://doi.org/10.1038/onc.2011.301.

[12] Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486(7403):346–52. https://doi.org/10.1038/nature10983.

[13] McCall MN, Uppal K, Jaffee HA, Zilliox MJ, Irizarry RA. The Gene Expression Barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic Acids Res 2011;39:D1011–5. https://doi.org/10.1093/nar/gkq1259. Database issue.

[14] Lebart L, Morineau A, Piron M. Statistique exploratoire multidimensionnelle. 1995. Paris.

[15] Therneau TM. A package for survival analysis in S version 2.38. 2015. https://CRAN.R-project.org/package=survival.

[16] Harano K, Wang Y, Lim B, Seitz RS, Morris SW, Bailey DB, et al. Rates of immune cell infiltration in patients with triple-negative breast cancer by molecular subtype. PLoS One 2018;13(10):e0204513. https://doi.org/10.1371/journal.pone.0204513.

[17] Loi S, Sirtaine N, Piette F, Conte P, Delaloge S, Andre F, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol Off J Am Soc Clin Oncol 2013;31(7):860–7. https://doi.org/10.1200/JCO.2011.41.0902.

[18] Panda A, Betigeri A, Subramanian K, Ross JS, Pavlick DC, Ali S, et al. Identifying a clinically applicable mutational burden threshold as a potential biomarker of response to immune checkpoint therapy in solid tumors. JCO Precis Oncol 2017;2017. https://doi.org/10.1200/PO.17.00146.

[19] Hothorn T, Lausen B. On the exact distribution of maximally selected rank statistics. Comput Stat Data Anal 2003;43(2):121–37. https://doi.org/10.1016/S0167-9473(02)00225-6.

[20] Loi S, Drubay D, Adams S, Pruneri G, Francis PA, Lacroix-Triki M, et al. Tumor-infiltrating lymphocytes and prognosis: a pooled individual patient Analysis of early-stage triple-negative breast cancers. J Clin Oncol January 2019. https://doi.org/10.1200/JCO.18.01010.

[21] Dieci MV, Criscitiello C, Goubar A, Viale G, Conte P, Guarneri V, et al. Prognostic value of tumor-infiltrating lymphocytes on residual disease after primary chemotherapy for triple-negative breast cancer: a retrospective multicenter study. Ann Oncol Off J Eur Soc Med Oncol 2015;26(7):1518. https://doi.org/10.1093/annonc/mdv241.

[22] Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, Shulman LN, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. J Clin Oncol Off J Am Soc Clin Oncol 2014;32(27):2959–66. https://doi.org/10.1200/JCO.2013.55.0491.

[23] West NR, Milne K, Truong PT, Macpherson N, Nelson BH, Watson PH. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res 2011;13(6):R126. https://doi.org/10.1186/bcr3072.

[24] Ibrahim EM, Al-Foheidi ME, Al-Mansour MM, Kazkaz GA. The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. Breast Canc Res Treat 2014;148(3):467–76. https://doi.org/10.1007/s10549-014-3185-2.

[25] Kochi M, Iwamoto T, Niikura N, Bianchini G, Masuda S, Mizoo T, et al. Tumour-infiltrating lymphocytes (TILs)-related genomic signature predicts chemotherapy response in breast cancer. Breast Canc Res Treat 2018;167(1):39–47. https://doi.org/10.1007/s10549-017-4502-3.

[26] Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. J Am Med Assoc 2011;305(18):1873–81. https://doi.org/10.1001/jama.2011.593.

[27] Apetoh L, Ghiringhelli F, Tesniere A, Obeid M, Ortiz C, Criollo A, et al. Toll-like receptor 4-dependent contribution of the immune system to anticancer chemotherapy and radiotherapy. Nat Med 2007;13(9):1050–9. https://doi.org/10.1038/nm1622.

[28] Nicolini A, Ferrari P, Duffy MJ. Prognostic and predictive biomarkers in breast cancer: past, present and future. Semin Cancer Biol 2018;52(Pt 1):56–73. https://doi.org/10.1016/j.semcancer.2017.08.010.

[29] Lausen B, Sauerbrei W, Schumacher M. Classification and Regression Trees (CART) used for the exploration of prognostic factors measured on different scales. In: Dirschedl P, Ostermann R, editors. Computational Statistics. Heidelberg: Physica-Verlag; 1994. p. 483–96.

[30] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17(1):218. https://doi.org/10.1186/s13059-016-1070-5.