Prognostic Value of Long Non-Coding RNAs MALAT1 in Colorectal Cancer Patients: A Pooled Analysis of Two Cohorts.

Heng Li
Harbin Medical University
YuXue Zhang
Harbin Medical University
YanLong Liu
Harbin Medical University
ZhangYi Qu
Harbin Medical University
YuPeng Liu
Harbin Medical University
JiPing Qi (qijiping2003@163.com)
Harbin Medical University https://orcid.org/0000-0001-8025-9781

Research

Keywords: colorectal cancer, MALAT1, prognosis value, propensity score analysis, cohort

Posted Date: June 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-583663/v1

Abstract

Background: Previous researches have shown that the aberrant expression of Metastasis associated in lung adenocarcinoma transcript 1 (MALAT1) in tumour tissues may serve as a biomarker for colorectal cancer (CRC) prognosis. However, these previous studies have small sample sizes and lacked validation from independently external populations. We therefore aimed to clarify the prognostic value of MALAT1 expression status in CRC patients using a large cohort and validate the findings with another large external cohort.

Methods: The prognostic association between MALAT1 expression status and CRC outcomes was evaluated initially in a prospective cohort in China (n=164) and then validated in an external TCGA population (n=596). In the initial cohort, MALAT1 expression levels were quantified by quantitative reverse transcriptase polymerase chain reaction. Propensity score (PS) adjustment method was used to control potential confounding biases. The prognostic significance was reported as PS-adjusted hazard ratio (HR) and corresponding 95% confidence interval (CI).

Results: There was no statistically significant association between MALAT1 expression status and CRC patient overall survival (OS) or disease free survival (DFS) in both initial cohort and external validation cohort populations. When combining these populations together, the results did not change materially. The summarized $HR_{PS\text{-}adjusted}$ were 1.010 (95%CI, 0.752-1.355, *P*=0.950) and 1.170 (95%CI, 0.910-1.502, *P*=0.220) for OS and DFS, respectively. We performed extensive sensitivity analyses, and demonstrated a very robustness of these results.

Conclusions: The MALAT1 expression status was not associated with prognostic outcomes in CRC patients. Our findings did not support a prognostic association of MALAT1 expression with CRC outcomes.

Background

Colorectal cancer (CRC) is a leading cause of cancer-related death in the world. It is the second- and third-most commonly diagnosed cancer in females and males, respectively; and more than 1.85 million newly diagnosed CRC patients and 881,000 deaths are estimated to occur in 2018 worldwide.[1] By 2035, the incidence and mortality of CRC is predicted to increase to 2.5 million new cases and 1.1 million deaths.[2, 3] Different from majority developed countries, there is a significant upward trend in the incidence of CRC in both men and women from 2000 to 2011 in China.[4] In 2018, approximately 522,000 new cases and 248,000 deaths are estimated in China, which accounts for more than 28% of all annually diagnosed CRC cases and CRC-related deaths worldwide.[1]

Even though mortality from CRC has significantly decreased over the past two decades, the 5-year relative survival rate is about 64% in the United States and the rate remained less than 50% in low-income regions.[3, 5] Approximately half of these new cases spread as micrometastases at the time of initial

diagnosis. Consequently, about 45% of patients suffer from recurrence or metastases after lesion resection.[6] To date, pathological tumour staging system and specific histological characteristics have been reported as the most common prognostic predictors for CRC patients after surgery. However, patients with similar clinical/pathological features often experience different clinical outcomes.[7] Therefore, an effectual predictive biomarker is urgently needed in prediction of outcomes of the disease. In our previous studies, we have developed a series blood based DNA methylation biomarkers for CRC prognosis.[8, 9] Long non-coding RNAs are becoming hotspots in the research fields of tumour biomarkers. Metastasis associated in lung adenocarcinoma transcript 1 (MALAT1) is a well-studied lncRNA.[10] MALAT1 that has 8.5kb nucleotides in length locates at 11q13, and is firstly found in a study of early-stage non-small-cell lung cancer.[11] Subsequent mechanism researches have demonstrated a vital function of MALAT1 in the development and progression in various cancers, including CRC.[12–14]

Recently, several studies have shown that the aberrant expression of MALAT1 in tumour tissues may serve as a biomarker for CRC prognosis.[15–19] However, these previous studies have small sample sizes. None of these studies validated their results in external populations. In order to clarify whether the expression status of MALAT1 in tumour tissues is associated with CRC prognosis or the clinical characteristics of CRC patients, we perform this prospective cohort analysis with a relatively large sample size and a long-term follow-up period. We further used the datasets of colon and rectum adenocarcinoma from The Cancer Genome Atlas (TCGA), as an independently external cohort population, to validate the findings.

Materials And Methods

Patient samples and inclusion criteria

This study was approved by the Harbin Medical University Ethics Review Board (Harbin, China). The study design and patient selection strategy have been published previously.[8, 9] Briefly, in our initial prospective cohort of CRC patients, a total of 168 patients were included according to the inclusion criteria. All of the patients provided written informed consents. The inclusion and exclusion criteria are as follows: (1) all patients are newly diagnosed with stage I-IV primary CRC, and their diagnosis was histologically confirmed by a senior pathologist (HL); (2) fresh frozen tumour tissues were collected from all patients; (3) patients with cocurrent any other types of cancer were excluded (n = 3); (4) patients with a family history of CRC in first-degree relatives were excluded (n = 5); (5) patients who received anti-cancer therapy before surgery were excluded (n = 11).

All CRC patients were diagnosed and operated at the First Affiliated Hospital and the Third Affiliated Hospital of Harbin Medical University between May 2010 and December 2012. The tumour specimens were staged according to the 2009 seventh version of the AJCC TNM staging system. The clinical characteristics and medical records were collected. The primary outcome was overall survival (OS), defined as time from surgery to death from any cause. The secondary outcome was disease-free survival

(DFS), defined as time from surgery to a local or regional relapse, distant metastasis, or CRC-specific death, whichever came first. Outcomes were observed during the follow-up period through March 15, 2018 via an established protocol. Postoperative patients were followed up at 3–6 months intervals for the first year and then annually. We used a telephone-delivery follow-up questionnaire to collect information on the date and cause of death of CRC patients. The recorded date and cause of death of each CRC patient were validated using the medical certification of death and the Harbin Death Registration system. Four cases lacked follow-up data and were then excluded in this analysis. Of these 164 eligible CRC patients, the median follow-up period was 61.1 months (ranging from 4.9 to 80.8 months) and 75 cases died.

RNA extraction and qRT-PCR assays

Fresh tumour tissue samples were collected and immediately stored at -80℃. Total RNA were extracted from fresh frozen tissues (0.5g) using TRIzol reagent (Invitrogen). cDNA was reverse transcribed from 2 μg total RNA using MultiScribe™ reverse transcriptase (Applied Biosystems). The RNA and cDNA concentration was measured using NanoDrop 2000c (ThermoFisher, USA). cDNA was then amplified and quantified by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) with Fast SYBR® Green Master Mix (Applied Biosystems) on the LightCycler 480 platform (Roche).The housekeeping gene GAPDH was selected as an internal control. No-template control was included in each batch and all reactions were performed in triplicate. The primer sequences are as follows. MALAT1 (NR_002819.4): F-(5'- GCTCTGTGGTGTGGGATTGA – 3'), MALAT1-R-(5'- GTGGCAAAATGGCGGACTTT – 3'); GAPDH (NM_002046.7): F-(5'-GGTGGTCTCCTCTGACTTCAACA – 3'), R-(5'- CCAAATTCGTTGTCATACCAGGAAATG – 3'). Melting curve analysis was used to monitor the specificity of PCR reactions. The resulting data was analysed using the Gene Scanning and TM Calling modules (Roche). Two co-authors (HL and YXZ) blinded to outcomes and independently recorded the results. The relative expression level of MALAT1 was determined using the $2^{-\Delta Ct}$ method. The ΔCt value of each sample was calculated by subtracting the average Ct value of MALAT1 from the average Ct value of GAPDH. According to the median value of $2^{-\Delta Ct}$, patients were categorized into higher or lower MALAT1 expression groups.

External validation dataset

The colorectal dataset (CORD) from TCGA was used as external validation population. The MALAT1 expression profile data, the clinicopathologic information, and survival data were downloaded from the TCGA database and the UCSC Xena resource.[20, 21] After excluding those without MALAT1 expression data (n = 102) or survival data (n = 30), a total of 596 patients were included in our analyses, including 475 patients with colon cancer and 121 with rectal cancer. The median follow-up period for these 596 patients was 22.5 months, with a range of 0.2 to 150.1 months, and a total of 121 cases died.

The gene expression RNA-seq-HTSeq-FPKM-UQ dataset for TCGA colon and rectum adenocarcinoma was performed using the UCSC Xena website tools, and then used in our analyses. The relative quantification of MALAT1 expression level was presented as N-fold differences and termed as 'N_{Malat1}', which was

determined by dividing the value of MALAT1 expression by the value of GAPDH. Then, patients were categorized into the higher (≥ median of N_{Malat1}) or lower (< median of N_{Malat1}) groups.

Statistical analysis

We used a Cox proportional hazards regression model to calculate the sample size. Given a pre-estimated overall survival rate of 50% in this initial cohort population, a sample size of 128 cases was required to achieve 90% power to detect an estimated hazard ratio (HR) of 1.5 with a two-sided 5% level of statistical significance. Finally, we included additionally 25% more patients and targeted a total sample size of 164 patients. The sample size was estimated using PASS software (version 11.0.7, NCSS LLC., USA).

We reported means (standard deviations) and counts (frequencies) for continuous and categorical variables, respectively. To minimise covariate differences between groups, we performed a PS-based analysis.[22] Group differences were compared using the standardised differences method with a significant imbalance level of standardised difference ≥ 25%. The PS value was calculated with MALAT1 expression level as the dependent variable using a multivariate logistic regression model that included demographic factors and clinical/pathological characteristics. We used the PS-adjustment method in order to incorporate all the patients in our analysis.[23]

Survival curves were estimated by the Kaplan-Meier method, and the differences between survival rates between groups were examined with log-rank tests. The univariate and PS-adjustment multivariate Cox-proportional hazards regression models were used to assess prognostic significance and the results were reported as hazard ratios (HRs) and 95% confidence intervals (CIs). The associations between MALAT1 expression status and those clinical/pathological covariates were reported as odds ratios (ORs) and 95% CIs. Statistical significance was defined as a two-sided $P < 0.05$. All statistical analyses were conducted with SPSS Statistics (v.23.0, IBM, USA).

Sensitivity analysis

Several predesigned sensitivity analyses were performed to explore the robustness of the results. Firstly, we compared the univariate HR and the PS-adjustment HR using the confounding RR,[24] which was calculated to evaluated the relative impact of the PS adjustment on the results. Secondly, we performed a conventional multivariate Cox regression analysis as a sensitivity analysis. Additionally, for the external cohort population, we performed a post hoc sensitivity analysis by excluding those patients with a shorter follow-up duration (≤ 1 or ≤ 3 months) in order to explore the potential confounding impact. Finally, we performed extensive post hoc subgroup analysis according to clinical/pathological factors. In post hoc subgroup analyses, we used the Bonferroni adjustment method to correct the level of statistical significance.

Meta analysis

In order to better understand the current evidence for the association between MALAT1 expression and CRC prognosis, we systematically review the relevant researches and performed a meta-analysis. We

systematically searched eligible studies assessing the prognostic significance of MALAT1 expression on CRC patient outcomes in PubMed, EmBase, and ProQuest through May 25, 2020. The inclusion criteria were as follows: (1) prospective cohort studies addressing the prognostic associations of MALAT1 and CRC outcomes; (2) studies that reported effect estimates including HRs with corresponding CIs; (3) studies with the sample size more than 50 participants; (4) there was no restriction on language, race, or any other participant characteristics. Data extraction was conducted independently by two co-authors (HL and YXZ). The maximally adjusted effect sizes and 95% CIs were extracted and summarised using random-effects models. The Q test and the I^2 Statistic were used to test the between-study heterogeneity. The pooled effect estimates were presented as forest plots. We performed E-value analysis,[25] as a post-hoc sensitivity analysis, to explore whether an unmeasured confounding factor could explain the observed associations.

Results

MALAT1 and Patient Outcomes

In the initial cohort, we analysed MALAT1 expression levels in a series of 164 tumour tissues from primary CRC patients with known clinical/pathological status and long-period follow-up outcomes. After PS adjustment, all these covariates between groups reached balance (Standardised mean difference < 0.25, **eTable 1**). There was no prognostic association between MALAT1 expression status and CRC patient outcomes. The univariate HRs were 1.428 (95% CI, 0.901–2.263, $P = 0.130$) and 1.525 (95% CI, 0.975–2.387, $P = 0.065$) for OS and DFS, respectively. After PS adjustment, The $HR_{PS-adjusted}$ were 1.087 (95% CI, 0.657–1.797, $P = 0.745$) and 1.150 (95% CI, 0.709–1.865, $P = 0.570$) for OS and DFS, respectively. Subgroup analyses by clinical/pathological factors showed similar results (Table 1).

Table 1
Prognostic associations of MALAT1 expression and colorectal cancer outcomes in the initial population.

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
Overall	164	82	82	1.087 (0.657–1.797)	0.745	1.150 (0.710–1.865)	0.570
Gender							
Female	93	51	42	1.133 (0.590–2.173)	0.708	1.328 (0.689–2.559)	0.396
Male	71	31	40	0.987 (0.448–2.174)	0.974	0.965 (0.477–1.953)	0.921
Age (yr)							
< 60	90	46	44	1.108 (0.559–2.199)	0.769	1.153 (0.592–2.246)	0.675
≥ 60	74	36	38	1.025 (0.488–2.153)	0.948	1.138 (0.568–2.279)	0.716
BMI (kg/m2)							
Normal weight	87	44	43	1.268 (0.621–2.589)	0.514	1.150 (0.581–2.279)	0.688
Overweight or Obese	77	38	39	0.824 (0.397–1.711)	0.604	1.026 (0.509–2.066)	0.943
Tumor Location							
Right Colon	19	10	9	1.866 (0.407–8.547)	0.422	0.911 (0.204–4.069)	0.902
Left Colon	44	19	25	1.046 (0.322–3.395)	0.940	1.074 (0.341–3.378)	0.903
Rectum	101	53	48	1.115 (0.610–2.038)	0.725	1.332 (0.748–2.372)	0.330

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
AJCC Stage							
1	16	10	6	0.002 (0.000-40.664)	0.391	0.002 (0.000-40.664)	0.391
2	66	26	40	0.707 (0.331–1.510)	0.370	0.859 (0.420–1.755)	0.676
3	66	35	31	1.702 (0.705–4.113)	0.237	1.790 (0.782–4.097)	0.168
4	16	11	5	1.037 (0.297–3.623)	0.954	2.399 (0.672–8.565)	0.178
CEA (ng/mL)							
≤ 5	83	45	38	0.677 (0.293–1.562)	0.360	0.682 (0.296–1.572)	0.369
> 5	81	37	44	1.394 (0.717–2.711)	0.328	1.528 (0.811–2.876)	0.189
History of Cancers							
No	142	69	73	1.195 (0.688–2.077)	0.527	1.256 (0.740–2.133)	0.398
Yes	22	13	9	0.676 (0.183–2.493)	0.577	0.766 (0.233–2.629)	0.671
CA19-9 (U/mL)							
≤ 37	122	68	54	1.208 (0.620–2.353)	0.578	1.484 (0.768–2.868)	0.240
> 37	42	14	28	0.547 (0.259–1.157)	0.114	0.427 (0.207–0.880)	0.021
Tumor Size (mm)							

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
≤ 40	58	31	27	2.086 (0.896–4.859)	0.088	2.063 (0.915–4.651)	0.081
> 40	106	51	55	0.788 (0.419–1.484)	0.461	0.857 (0.471–1.560)	0.614
Histopathological Morphology							
Protruding	114	60	54	0.971 (0.494–1.908)	0.931	1.085 (0.561-2.100)	0.808
Infiltrating ulcer	50	22	28	1.336 (0.625–2.587)	0.454	1.373 (0.668–2.822)	0.389
Differentiation							
Low to Medium	103	52	51	1.525 (0.836–2.780)	0.169	1.650 (0.916–2.974)	0.095
High	61	30	31	0.568 (0.224–1.440)	0.233	0.590 (0.247–1.410)	0.235
Adjuvant Chemotherapy							
No	102	54	48	0.774 (0.379–1.579)	0.482	0.808 (0.415–1.575)	0.531
Yes	62	28	34	1.748 (0.839–3.643)	0.136	1.843 (0.888–3.825)	0.101

The results from the initial cohort study were validated by using of a large external cohort involving 596 CRC patients from TCGA. After PS adjustment, all these clinical/pathological covariates between groups were balanced (**eTable 2**). The findings were consistent with the results from the initial cohort. The univariate HRs were 1.072 (95% CI, 0.750–1.532, $P = 0.703$) and 1.266 (95% CI, 0.948–1.690, $P = 0.110$) for OS and DFS, respectively. After PS adjustment, the $HR_{PS-adjusted}$ were 0.971 (95% CI, 0.676–1.396, $P = 0.876$) and 1.177 (95% CI, 0.878–1.577, $P = 0.276$) for OS and DFS, respectively. Subgroup analyses showed similar results (Table 2).

Table 2
Prognostic associations of MALAT1 expression and colorectal cancer outcomes in the external validation population.

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
Overall	596	298	298	0.971 (0.676–1.396)	0.876	1.177 (0.878–1.577)	0.276
Gender							
Female	274	137	137	1.103 (0.646–1.881)	0.720	1.290 (0.835–1.994)	0.252
Male	322	161	161	0.839 (0.508–1.383)	0.491	1.087 (0.729–1.619)	0.683
Age (yr)							
< 60	174	92	82	1.680 (0.713–3.960)	0.235	1.513 (0.835–2.743)	0.172
≥ 60	422	206	216	0.880 (0.587–1.319)	0.536	1.099 (0.783–1.542)	0.585
BMI (kg/m^2)							
Normal weight	195	116	79	1.195 (0.654–2.186)	0.562	1.450 (0.869–2.421)	0.155
Overweight or Obese	401	182	219	0.987 (0.627–1.555)	0.956	1.136 (0.794–1.625)	0.486
Tumor Location							
Right Colon	250	130	120	0.856 (0.514–1.426)	0.551	0.896 (0.574–1.397)	0.627
Left Colon	225	111	114	1.232 (0.645–2.352)	0.528	1.477 (0.907–2.406)	0.117
Rectum	121	57	64	0.867 (0.348–2.160)	0.759	1.354 (0.673–2.727)	0.396

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
AJCC Stage							
1	106	57	49	0.942 (0.149–5.970)	0.950	1.836 (0.553-6.100)	0.321
2	222	113	109	0.445 (0.213–0.926)	0.030	0.792 (0.467–1.344)	0.388
3	179	92	87	1.144 (0.618–2.116)	0.669	1.211 (0.727–2.015)	0.462
4	89	36	53	1.082 (0.545–2.150)	0.822	1.342 (0.756–2.382)	0.314
CEA (ng/mL)							
≤ 5	263	143	120	0.837 (0.419–1.672)	0.615	1.139 (0.677–1.915)	0.623
> 5	333	155	178	1.037 (0.675–1.594)	0.869	1.172 (0.821–1.673)	0.381
History of polyps							
No	407	212	195	1.027 (0.671–1.571)	0.903	1.126 (0.800-1.586)	0.497
Yes	189	86	103	0.814 (0.403–1.647)	0.568	1.301 (0.734–2.307)	0.368

We then combined the PS-adjusted HRs from the initial and external populations together by using random effect models, and found no prognostic significance of MALAT1 expression status in CRC patient outcomes. The pooled $HR_{PS-adjusted}$ were 1.010 (95% CI, 0.752–1.355, $P = 0.950$) and 1.170 (95% CI, 0.910–1.502, $P = 0.220$) for OS and DFS, respectively. The pooled effect estimates for subgroup populations also showed similar results (Table 3).

Table 3
Prognostic associations of MALAT1 expression and colorectal cancer outcomes in the combined populations.

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
Overall	760	380	380	1.010 (0.752–1.355)	0.950	1.170 (0.910–1.502)	0.220
Gender							
Female	367	188	179	1.115 (0.737–1.685)	0.606	1.301 (0.905–1.871)	0.155
Male	393	192	201	0.879 (0.576–1.341)	0.549	1.056 (0.746–1.494)	0.759
Age (yr)							
< 60	264	138	126	1.303 (0.763–2.225)	0.332	1.341 (0.861–2.091)	0.195
≥ 60	496	242	254	0.911 (0.639-1.300)	0.609	1.106 (0.816-1.500)	0.515
BMI (kg/m^2)							
Normal weight	282	160	122	1.225 (0.773–1.942)	0.388	1.334 (0.886–2.010)	0.168
Overweight or Obese	478	220	258	0.939 (0.638–1.381)	0.748	1.112 (0.809–1.530)	0.513
Tumor Location							
Right Colon	269	140	129	0.926 (0.571–1.503)	0.756	0.897 (0.586–1.373)	0.616
Left Colon	269	130	139	1.186 (0.673–2.091)	0.556	1.407 (0.898–2.204)	0.136
Rectum	222	110	112	1.033 (0.625–1.708)	0.900	1.341 (0.859–2.093)	0.196

Factors and Subgroups	No. of Patients	MALAT1 Expression Level		Overall Survival		Disease Free Survival	
		Lower (ref.)	Higher	PS-Adjusted HR (95% CI)	P-value	PS-Adjusted HR (95% CI)	P-value
AJCC Stage							
1	122	67	55	0.928 (0.147–5.879)	0.937	1.824 (0.549–6.059)	0.326
2	288	139	149	0.556 (0.328–0.943)	0.029	0.815 (0.533–1.247)	0.346
3	245	127	118	1.303 (0.786–2.157)	0.304	1.348 (0.873–2.080)	0.178
4	105	47	58	1.071 (0.587–1.956)	0.822	1.481 (0.878–2.498)	0.141
CEA (ng/mL)							
≤ 5	346	188	158	0.768 (0.451–1.309)	0.332	0.982 (0.623–1.549)	0.939
> 5	414	192	222	1.131 (0.789–1.624)	0.503	1.249 (0.916–1.703)	0.160

Sensitivity Analysis

The results of confounding RR analyses are showed in **eTable 3**. Overall, confounding RRs demonstrated no significant change after PS adjustment. However, all these confounding RRs were small than 1, suggesting that the PS adjustment generate more conservative effect estimates. Another sensitivity analysis by using of the conventional multivariate Cox regression models found very similar results (**eTable 4**). For the external cohort validation analysis, a sensitivity analysis by excluding those fourteen patients with follow-up periods of no more than 1 month or 3 months did not materially change the results (**eTable 5**).

Meta-analysis of MALAT1 and Patient Outcomes

To further assess the robustness of the results, we performed a systematically meta-analysis. The pooled results were showed in Fig. 1. Briefly, another three eligible studies were included in this meta-analysis. By pooling these results together, we still did not find a positive prognostic association for OS, with a summarized HR of 1.683 (95% CI, 0.917–3.087; $P = 0.093$). For DFS, there was a marginally positive

association between higher MALAT1 expression and worse DFS, with a summarized HR of 1.784 (95% CI, 1.021–3.118; *P* = 0.042).

MALAT1 and Clinical/Pathological Characteristics

We sought associations between MALAT1 expression level and clinical/pathological characteristics in CRC patients (**eTable 6 and eTable 7**). We found a significantly positive association between MALAT1 overexpression and higher CA19-9 level (*P* = 0.016), and higher T stage (*P* = 0.030) in the initial cohort population. In the external cohort, a significantly strong association between MALAT1 overexpression and overweight or obese (≥ 25 kg/m^2) was observed (*P* = 0.001). As for other clinical/pathological factors, there was no positive relationship.

Discussion

In our initial cohort, there was no prognostic association between MALAT1 expression status and CRC patient outcomes. This finding was confirmed in the external TCGA cohort. Furthermore, the consistence among extensive sensitivity analyses provided the robustness of the results. To our best knowledge, this present study is the largest population cohort addressing the prognostic effect of MALAT1 on CRC patient outcomes. We initially performed a prospective cohort analysis with a long-term follow-up period of 7 years. Then we used the CORD patient cohort from TCGA as external validation datasets. No association of MALAT1 expression status with OS or DFS of CRC patients was found in our analysis, which was inconsistent with the findings from several previous studies.

A total of eight relevant studies assessed the association of MALAT1 expression and CRC prognosis.[13–19] Most of these studies reported that CRC patients with higher MALAT1 expression in tumour tissues had worse clinical outcomes with shorter OS or DFS. However, the sample sizes of these previous studies were all small, ranging from 30 to 146 cases. In addition, univariate Cox hazard ratio regression models were used in most of these studies, and only Zheng[15] and Li[16] took the potential impact of multifactor confounders into consideration. None of these studies conducted PS-based analyses. The PS-based method is a powerful statistical tool to control for confounding bias and is often more practical and statistically more efficient than conventional strategies of multivariate statistical analyses,[22, 23] which has been increasingly used to reduce the impact of confounders in observational studies, especially studies with small sample size.

Based on the inclusion criteria, three eligible studies were finally included in the meta-analysis. The pooled results supported the notion that there was no association between MALAT1 expression and OS of CRC patients. For DFS, a marginally positive prognostic significance was observed; however, E-values of both the point estimate and the lower CI limit of the pooled results were small, suggesting that a hypothetical residual confounding factor would fully explain the observed association for DFS (**eTable 8**). Future large population cohorts are needed to further validate this issue for DFS. Given the rigorousness

and better performance in controlling confounders of the PS methods used in this study, conclusions were drawn mainly according to the findings from our initial and external validation populations.

Subgroup analyses by AJCC stages revealed a marginally better OS in stage II CRC patients with higher MALAT1 expression than those with lower expression. The $HR_{PS-adjusted}$ was 0.556 (95% CI, 0.328–0.943) with a *P*-value of 0.029, which did not reach statistical significance according to the Bonferroni correction method (α = 0.0125). In the CA19-9 higher level subgroup, it is found that CRC patients with higher MALAT1 expression had a longer OS than those with lower expression. But this finding cannot be validated in the TCGA external cohort population, due to the lack of eligible data. Therefore, the findings from subgroup analyses should be interpreted with cautions.

Strengths and limitations

This study had several major strengths, including the novel PS-based analysis, a relatively large population, the validation by using of the external cohort population, and the validation by meta-analysis. However, our present study had certain limitations. Firstly, confounding bias was the major limitation due to the nature of the observational cohort study design. The findings of confounding RR analyses suggested that those confounders could overstate the prognostic association of MALAT1 with CRC patient outcomes. In a conservative manner, we used the PS-adjustment method to maximally control for the impact of potential confounders on the results. It is known that the PS method is a powerful statistical tool to reduce the likelihood of confounding bias in observational studies. Another limitation is the lack of detailed information about adjuvant chemotherapy from both our initial cohort and the external cohort.

Conclusions

In summary, our findings did not support a prognostic association of MALAT1 expression with CRC patient outcomes.

List Of Abbreviations

CI, confidence interval; CORD, colorectal dataset; CRC, colorectal cancer; DFS, disease free survival; HR, hazard ratio; MALAT1, Metastasis associated in lung adenocarcinoma transcript 1; OS, overall survival; PS, propensity score; TCGA, The Cancer Genome Atlas.

Declarations

Ethics approval and consent to participate: This work has been approved by the Medical Ethics Committee of Harbin Medical University. All participants in the initial cohort provided written informed consent.

Consent for publication: Not applicable.

Availability of data and materials: Authors are willing to share any data that are used in this work. The datasets used in the current initial study are available from the corresponding author (YPL) on reasonable request. The datasets used in the validation stage are publicly available on TCGA (https://cancergenome.nih.gov/).

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by the Natural Science Foundation of Heilongjiang Province (grant number YQ2019H021 to YPL); the China Postdoctoral Science Foundation (grant number 2018M641875 to YPL); and the Wenzhou Science & Technology Bureau Scientific Research Project (grant number Y20190191 to YPL). The funders had no role in the design of the study, analysis, interpretation of data and manuscript writing.

Authors' contributions: HL and YXZ contributed equally to this work. YPL and JPQ had full access to all of the data in this work and take responsibility for the integrity of the data and the accuracy of the data analysis. YPL and JPQ contributed to study conception and design. YPL and JPQ were responsible for study supervision. HL, YLL and JPQ contributed to sample collection. HL, YXZ and YPL contributed to RNA preparation, RT-PCR experiments, and acquisition and assembly of data. HL, YXZ, YLL and YPL contributed to analysis and interpretation of data. YXZ, HL and YPL contributed to drafting of the initial versions of the manuscript. YLL, ZYQ and JPQ contributed to revise this manuscript critically for important intellectual content. All authors contributed to the review and final approval of this manuscript.

Acknowledgements: We thank supports of our work by the Natural Science Foundation of Heilongjiang Province (grant number YQ2019H021 to YPL); the China Postdoctoral Science Foundation (grant number 2018M641875 to YPL); and the Wenzhou Science & Technology Bureau Scientific Research Project (grant number Y20190191 to YPL). We also thank American Journal Experts for English language polishing of the manuscript.

References

1. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. (2018). Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. https://gco.iarc.fr/today/home (Accessed November 22, 2020).
2. Arnold M, Sierra MS, Laversanne M, et al. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683–91.
3. Dekker E, Tanis PJ, Vleugels JLA, et al. Colorectal cancer *Lancet*. 2019;394(10207):1467–80.
4. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115–32.
5. American Cancer Society. Survival Rates for Colorectal Cancer. Available at: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html. (Accessed November 22, 2020).

6. Schmoll HJ, Van Cutsem E, Stein A, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. Ann Oncol. 2012;23(10):2479–516.
7. Bosman FT, Hamilton SR, Lambert R. Colorectal cancer. In: Stewart BW, Wild CP, editors. World Cancer Report 2014. Lyon, France, International Agency for Research on Cancer. WHO Press; 2014. pp. 560–76.
8. Sun H, Huang H, Li D, et al. PBX3 hypermethylation in peripheral blood leukocytes predicts better prognosis in colorectal cancer: A propensity score analysis. Cancer Med. 2019;8(8):4001–11.
9. Liu X, Fu J, Bi H, et al. DNA methylation of SFRP1, SFRP2, and WIF1 and prognosis of postoperative colorectal cancer patients. BMC Cancer. 2019;19(1):1212.
10. Gutschner T, Richtig G, Haemmerle M, et al. From biomarkers to therapeutic targets-the promises and perils of long non-coding RNAs in cancer. Cancer Metastasis Rev. 2018;37(1):83–105.
11. Ji P, Diederichs S, Wang W, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene. 2003;22(39):8031–41.
12. Xiong Y, Wang J, Zhu H, et al. Chronic oxymatrine treatment induces resistance and epithelialmesenchymal transition through targeting the long non-coding RNA MALAT1 in colorectal cancer cells. Oncol Rep. 2018;39(3):967–76.
13. Ji Q, Cai G, Liu X, et al. MALAT1 regulates the transcriptional and translational levels of proto-oncogene RUNX2 in colorectal cancer metastasis. Cell Death Dis. 2019;10(6):378.
14. Qiu G, Zhang XB, Zhang SQ, et al. Dysregulation of MALAT1 and miR-619-5p as a prognostic indicator in advanced colorectal carcinoma. Oncol Lett. 2016;12(6):5036–42.
15. Zheng HT, Shi DB, Wang YW, et al. High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer. Int J Clin Exp Pathol. 2014;7(6):3174–81.
16. Li P, Zhang X, Wang H, et al. MALAT1 Is Associated with Poor Response to Oxaliplatin-Based Chemotherapy in Colorectal Cancer Patients and Promotes Chemoresistance through EZH2. Mol Cancer Ther. 2017;16(4):739–51.
17. Thiele JA, Hosek P, Kralovcova E, et al. lncRNAs in Non-Malignant Tissue Have Prognostic Value in Colorectal Cancer. *Int J Mol Sci* 2018; 19(9).
18. Li Q, Dai Y, Wang F, et al. Differentially expressed long non-coding RNAs and the prognostic potential in colorectal cancer. Neoplasma. 2016;63(6):977–83.
19. Thorsteinsson M, Kirkeby LT, Hansen R, et al. Gene expression profiles in stages II and III colon cancers: application of a 128-gene signature. Int J Colorectal Dis. 2012;27(12):1579–86.
20. National Cancer Inistitute. GDC Data Portal. Available at: https://portal.gdc.cancer.gov/. (Accesed April 7, 2020).
21. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675–8.
22. Haukoos JS, Lewis RJ. The Propensity Score. JAMA. 2015;314(15):1637–8.

23. Elze MC, Gregson J, Baber U, et al. Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies. J Am Coll Cardiol. 2017;69(3):345–57.
24. Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev. 1987;9:1–30.
25. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268–74.

Figures

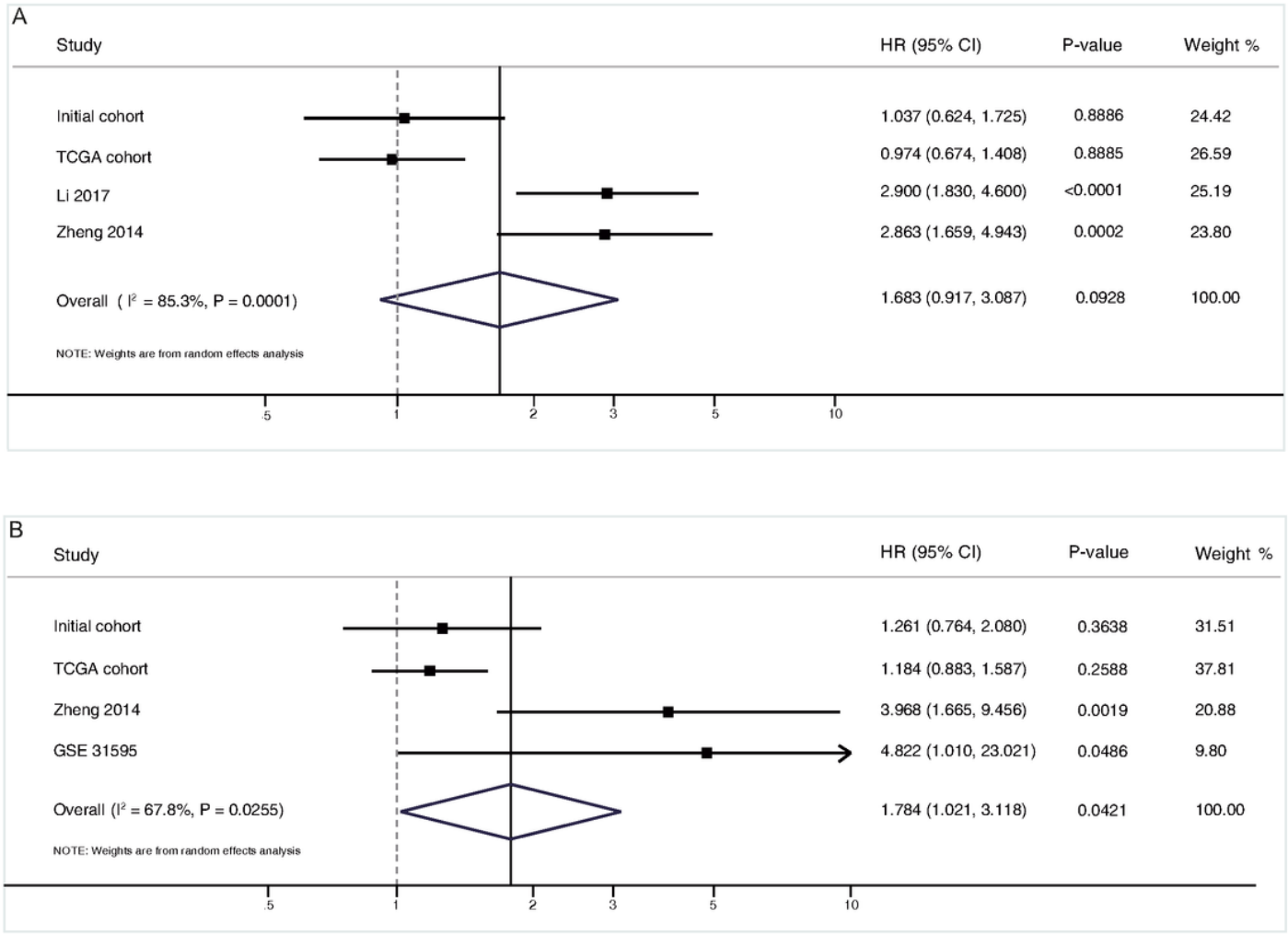

Figure 1

Pooled results of the meta-analysis for MALAT1 expression and colorectal cancer outcomes. (A) Overall survival; (B) Disease free survival. CI, confidence interval; HR, hazard ratio.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 2SupplementaryMetarials.pdf