Development of a preoperative prediction nomogram for lymph node metastasis in colorectal cancer based on a novel serum miRNA signature and CT scans

Ailin Qu, Yongmei Yang, Xin Zhang, Wenfei Wang, Yingjie Liu, Guixi Zheng, Lutao Du, Chuanxin Wang

ARTICLE INFO
Article history:
Received 15 August 2018
Received in revised form 25 September 2018
Accepted 30 September 2018
Available online 9 October 2018

Keywords:
Colorectal cancer
miRNA-based panel
LN metastasis
Prediction
Nomogram

ABSTRACT

Background: Preoperative prediction of lymph node (LN) status is of crucial importance for appropriate treatment planning in patients with colorectal cancer (CRC). In this study, we sought to develop and validate a non-invasive nomogram model to preoperatively predict LN metastasis in CRC.

Methods: Development of the nomogram entailed three subsequent stages with specific patient sets. In the discovery set (n = 20), LN-status-related miRNAs were screened from high-throughput sequencing data of human CRC serum samples. In the training set (n = 218), a miRNA panel-clinico-pathologic nomogram was developed by logistic regression analysis for preoperative prediction of LN metastasis. In the validation set (n = 198), we validated the above nomogram with respect to its discrimination, calibration and clinical application.

Findings: Four differently expressed miRNAs (miR-122-5p, miR-146b-5p, miR-186-5p and miR-193a-5p) were identified in the serum samples from CRC patients with and without LN metastasis, which also had regulatory effects on CRC cell migration. The combined miRNA panel could provide higher LN prediction capability compared with computed tomography (CT) scans (P < .0001 in both the training and validation sets). Furthermore, a nomogram integrating the miRNA-based panel and CT-reported LN status was constructed in the training set, which performed well in both the training and validation sets (AUC: 0.913 and 0.883, respectively). Decision curve analysis demonstrated the clinical usefulness of the nomogram.

Interpretation: Our nomogram is a reliable prediction model that can be conveniently and efficiently used to improve the accuracy of preoperative prediction of LN metastasis in patients with CRC.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Colorectal cancer (CRC) is the third most frequent malignant tumor worldwide, representing the second leading cause of cancer-related mortality globally [1]. Lymph node (LN) metastasis is an important determining factor for the outcome of CRC [2]. Accurate preoperative prediction of LN status in CRC is of crucial importance for appropriate therapeutic decisions, such as the utilization of neoadjuvant and/or adjuvant chemotherapy for patients with LN metastasis, or the implementation of a more conservative approach to keep bowel resection to a minimum for patients without LN metastasis [3,4]. To date, imaging modalities, including computed tomography (CT), are frequently used to predict LN involvement before surgery in clinical practice. These modalities, however, usually have limited sensitivity and specificity in predicting LN metastasis [5–7]. High-risk histopathologic features, including lymphatic or vascular invasion and poor differentiation, are also known to be predictors of LN metastasis [8]; however these data can only be obtained postoperatively. Therefore, there is an unmet need to develop novel non-invasive biomarkers to complement and improve current strategies for preoperative prediction of LN metastasis in patients with CRC.
Recent advancements in transcriptome profiling have highlighted the potential of small non-coding RNAs (ncRNAs) as tumor biomarkers [9]. MicroRNAs (miRNAs) are the most abundant and best characterized class of small ncRNAs, which regulate gene expression post-transcriptionally in multiple cancer-related processes including metastasis [10–12]. Accumulating evidence has provided insight into the role of dysregulated miRNAs as potential tumor markers to predict disease progression and metastasis [13,14]. More importantly, circulating miRNAs originating from primary tumor tissues, are stably detectable in human body fluids [15–17]. These characteristics make circulating miRNAs ideal noninvasive indicators for tumor detection and prognosis [18]. Recently, differential expressions of circulating miRNAs have been reported to predict LN metastasis in various cancers [19–21]. In CRC, circulating miRNAs, such as miR-203 and miR-200c, have been reported to be independent predictors of LN metastasis [22,23]. Although many studies have proposed circulating miRNAs to be predictors of metastasis, very few have attempted to identify circulating miRNA-based signatures for prediction of LN status before surgery [21,24,25]. Our group has recently identified a serum four-miRNA signature to preoperatively predict LN status in gastric cancer [24]. However, to date, there is no direct evidence as to whether a serum miRNA signature would enable superior prediction of LN status in other tumors, including CRC.

The combined analysis of multiple factors, rather than just a single biomarker—as could be provided by a nomogram—would be able to yield more powerful and accurate information in the clinical setting [26–28]. In the current study, we identified serum miRNAs that were significantly associated with LN metastasis by comprehensive miRNA profiling and RT-qPCR analysis, and then developed a serum miRNA-based panel in our CRC sample set. The serum miRNA-based panel was further combined with clinical risk factors to build a non-invasive nomogram for the preoperative prediction of LN status. Additionally, we assessed the predictive accuracy and clinical usefulness of the nomogram and validated it in an independent cohort.
(radiotherapy, chemotherapy or chemoradiotherapy) or suffered from other tumor diseases at the same time were excluded.

Relevant clinical information including sex, age, preoperative histological differentiation and carcinoembryonic antigen (CEA) levels were collected from medical records. CEA was obtained from routine blood test before surgery, and the cutoff value was 5 ng/mL. CT scans were reviewed by two radiologists with >10 years of experience, who were blinded to clinical characteristics and postoperative pathological findings. Patients with regional LN of ≥1 cm and/or clusters of ≥3 lymph nodes were identified as clinically LN-positive, and patients without enlarged or clustered lymph nodes were regarded as clinically LN-negative [30]. Any disagreement was resolved by consultation.

Whole blood samples (5 mL) were collected from each participant by venipuncture. Serum was separated within 1 h of blood collection by centrifugation at 1500 × g for 10 min at 4 °C to completely remove cell debris. The supernatant (serum) was then transferred to RNase/DNase-free tubes and stored at -80 °C until further processing.

### 2.4. RT-qPCR analysis of miRNA expression

Briefly, 2 μg of total RNA was reversely transcribed into cDNA using the Mir-X miRNA First-Stand Synthesis Kit (Takara, Shiga, Japan). Then, 2 μL of cDNA was used for quantitative PCR analysis that was performed on the Bio-Rad CFX96 Detection System (Bio-Rad, Hercules, CA) using the SYBR Premix EX Taq (Takara, Shiga, Japan) and miRNA-specific primers (Ribobio, Guangzhou, China). All reactions were performed in triplicate to remove any outliers. In addition, miR-191-5p and U6 were selected as the reference genes according to our previous study [31]. Any disagreement was resolved by consultation.

Whole blood samples (5 mL) were collected from each participant by venipuncture. Serum was separated within 1 h of blood collection by centrifugation at 1500 × g for 10 min at 4 °C to completely remove cell debris. The supernatant (serum) was then transferred to RNase/DNase-free tubes and stored at -80 °C until further processing.

#### 2.4. RT-qPCR analysis of miRNA expression

Briefly, 2 μg of total RNA was reversely transcribed into cDNA using the Mir-X miRNA First-Stand Synthesis Kit (Takara, Shiga, Japan). Then, 2 μL of cDNA was used for quantitative PCR analysis that was performed on the Bio-Rad CFX96 Detection System (Bio-Rad, Hercules, CA) using the SYBR Premix EX Taq (Takara, Shiga, Japan) and miRNA-specific primers (Ribobio, Guangzhou, China). All reactions were performed in triplicate to remove any outliers. In addition, miR-191-5p and U6 were selected as the reference genes according to our previous study [31]. The miRNA expression was normalized to the mean of these two reference genes.

Whole blood samples (5 mL) were collected from each participant by venipuncture. Serum was separated within 1 h of blood collection by centrifugation at 1500 × g for 10 min at 4 °C to completely remove cell debris. The supernatant (serum) was then transferred to RNase/DNase-free tubes and stored at -80 °C until further processing.

#### 2.4. RT-qPCR analysis of miRNA expression

Briefly, 2 μg of total RNA was reversely transcribed into cDNA using the Mir-X miRNA First-Stand Synthesis Kit (Takara, Shiga, Japan). Then, 2 μL of cDNA was used for quantitative PCR analysis that was performed on the Bio-Rad CFX96 Detection System (Bio-Rad, Hercules, CA) using the SYBR Premix EX Taq (Takara, Shiga, Japan) and miRNA-specific primers (Ribobio, Guangzhou, China). All reactions were performed in triplicate to remove any outliers. In addition, miR-191-5p and U6 were selected as the reference genes according to our previous study [31]. The miRNA expression was normalized to the mean of these two reference genes.

#### 2.5. Functional analysis of miRNAs

HT-29 (NCI-DTP Cat# HT-29, RRID:CVCL_0320) and SW480 (CLS Cat# 300302/p716_SW-480, RRID:CVCL_0546) were purchased from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China) and cultured in a DMEM medium supplemented with 10% fetal bovine serum (Gibco, Carlsbad, CA, USA). Transient transfection of chemically synthesized miRNA mimics (Ribobio, Guangzhou, China) was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) following the manufacturer’s instructions. Cell migration assays were performed as previously described [32].

#### 2.5. Functional analysis of miRNAs

HT-29 (NCI-DTP Cat# HT-29, RRID:CVCL_0320) and SW480 (CLS Cat# 300302/p716_SW-480, RRID:CVCL_0546) were purchased from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China) and cultured in a DMEM medium supplemented with 10% fetal bovine serum (Gibco, Carlsbad, CA, USA). Transient transfection of chemically synthesized miRNA mimics (Ribobio, Guangzhou, China) was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) following the manufacturer’s instructions. Cell migration assays were performed as previously described [32].

#### 2.6. Statistical analysis

All the statistical analysis was performed with SPSS (version 18.0, Chicago, IL, USA), MedCalc 9.3.9.0 and R software (version 3.4.2; http://www.Rproject.org). P-values of 0.05 were considered to be statistically significant. The data distribution of each group was determined by the Kolmogorov-Smirnov test. Continuous variables were presented as the mean ± standard deviation (SD) or median (interquartile range) when the values were normally or non-normally distributed, respectively. Statistical differences between the groups were assessed using the Mann-Whitney U test and Student’s t-test as appropriate. Categorical variables were presented as numbers (%) and analyzed using Pearson’s chi-squared test or Fisher’s exact test as appropriate. KEGG pathway enrichment analysis was conducted using the “clusterProfiler” package [33]. The performance of the nomogram was evaluated using calibration curve and ROC curve (“rms” package). DCA was performed using the “dca.R” (decisioncurveanalysis.org).

#### 3. Results

##### 3.1. Clinical characteristics

Table 1 shows the baseline clinical and pathological characteristics, which were similar between the training and validation cohorts (all P > 0.5). In addition, LN status was found to be significantly associated with preoperative CT-reported LN status and CEA levels in both cohorts.

#### 3.2. Selection of candidate miRNAs to predict LN metastasis

In the discovery stage, we identified 30 miRNAs that are differentially expressed between the serum samples from two groups of primary stage T3 CRC patients, including 10 LN- and 10 LN+ patients, using HTS (Table S1). In the training stage, 22 miRNAs passed the quality test after miRNAs with Ct mean values higher than 35 and detection rates lower than 75% were excluded from the RT-qPCR analysis of 60 serum samples from 30 LN- and 30 LN+ CRC patients. Using RT-qPCR analysis, four differently expressed miRNAs (miR-122-5p, miR-146b-5p, miR-186-5p and miR-193a-5p), showing a consistent trend with...
the sequencing data (all $P < 0.05$, Fig. S1). Then, we analyzed another 158 serum samples from CRC patients (80 LN- and 78 LN+) and confirmed the above phenomenon. The combined 108 LN- and 110 LN+ patients were used as the training set. As shown in Fig. 2, these four miRNAs showed significantly different expression between LN- and LN+ patients (all $P < 0.0001$, with AUC values ranging from 0.681 to 0.811 (Table S2). We further evaluated the expression of these four miRNAs in the validation cohort consisting of 98 LN- and 100 LN+ patients. The alteration patterns of the miRNA expression in the validation set were consistent with those in the training set, with AUCs ranging from 0.722 to 0.796 (Fig. S2).

**Table 1** Characteristics of study participants in training set and validation set.

| Variable, N. (%) | Discovery set | Training Set | Validation Set | P* | P* |
|------------------|---------------|--------------|----------------|----|----|
|                  | LN (-) | LN (+) | Total | LN (-) | LN (+) | Total | LN (-) | LN (+) |
| Sex Male         | 10 (100) | 10 (100) | 126 (57.8) | 59 (54.6) | 67 (60.9) | 116 (58.6) | 55 (56.1) | 61 (61.0) | 0.486 | 0.871 |
| Female           | 0 (0)   | 0 (0)   | 92 (42.2)  | 49 (45.4)  | 43 (39.1)  | 82 (41.4)  | 43 (43.9)  | 39 (39.0)  |        |        |
| Agea            | 5 (50)  | 3 (30)  | 111 (50.9) | 61 (56.5)  | 50 (45.5)  | 101 (51.0) | 54 (55.1)  | 47 (47.0)  | 0.254 | 0.985 |
| Tumor location  | 5 (50)  | 7 (70)  | 107 (49.1) | 47 (43.5)  | 60 (54.5)  | 97 (49.0)  | 44 (44.9)  | 53 (53.0)  |        |        |
| Colon           | 5 (50)  | 5 (50)  | 74 (33.9)  | 35 (32.4)  | 39 (35.5)  | 63 (32.0)  | 33 (33.7)  | 36 (36.0)  | 0.731 | 0.846 |
| Rectum          | 5 (50)  | 5 (50)  | 144 (66.1) | 73 (67.6)  | 71 (64.5)  | 129 (65.2) | 65 (66.3)  | 64 (64.0)  |        |        |
| Well             | 1 (10)  | 0 (0)   | 26 (12.1)  | 16 (14.8)  | 10 (9.3)   | 24 (12.3)  | 14 (14.3)  | 10 (10.3)  | 0.033 | 0.994 |
| Moderate        | 8 (80)  | 7 (70)  | 140 (65.1) | 74 (68.5)  | 66 (61.7)  | 126 (64.6) | 69 (70.4)  | 57 (58.8)  |        |        |
| Poor            | 1 (10)  | 3 (30)  | 49 (22.8)  | 18 (16.7)  | 31 (29.0)  | 45 (23.1)  | 15 (15.3)  | 30 (29.0)  |        |        |
| CT-reported LN status | 6 (60) | 3 (30) | 98 (45.0) | 62 (57.4) | 36 (32.7) | 101 (51.0) | 67 (68.4) | 34 (34.0) | <0.0001 | 0.217 |
| LN-negative      | 4 (40)  | 7 (70)  | 120 (55.0) | 46 (42.6)  | 74 (67.3)  | 97 (49.0)  | 31 (31.6)  | 66 (66.0)  |        |        |
| LN-positive       | 7 (70)  | 6 (60)  | 142 (65.1) | 78 (72.2)  | 64 (58.2)  | 131 (66.2) | 72 (73.5)  | 59 (59.0)  | 0.031 | 0.826 |
| CEA level         | 3 (30)  | 4 (40)  | 76 (34.9)  | 30 (27.8)  | 46 (41.8)  | 67 (33.8)  | 26 (26.5)  | 41 (41.0)  |        |        |
| Abnormal         | 10 (100) | 0 (0)  | 174 (79.8) | 66 (61.1)  | 108 (98.2) | 160 (80.8) | 61 (62.2)  | 99 (99.0)  |        |        |
| Local invasion    | 6 (60)  | 2 (2)   | 104 (47.7) | 57 (52.8)  | 47 (42.7)  | 94 (47.5)  | 52 (53.1)  | 42 (42.0)  | 0.097 | 0.961 |
| T1–T4           | 4 (40)  | 8 (8)   | 92 (42.2)  | 39 (36.1)  | 53 (48.2)  | 84 (42.4)  | 36 (36.7)  | 48 (48.0)  |        |        |
| Tumor sizeb      | 10 (100) | 10 (100) | 174 (79.8) | 66 (61.1)  | 108 (98.2) | 160 (80.8) | 61 (62.2)  | 99 (99.0)  |        |        |
| 𝜷< 4.7 cm        | 5 (50)  | 7 (70)  | 107 (49.1) | 47 (43.5)  | 60 (54.5)  | 97 (49.0)  | 44 (44.9)  | 53 (53.0)  |        |        |
| ≥ 4.7 cm         | 5 (50)  | 5 (50)  | 74 (33.9)  | 35 (32.4)  | 39 (35.5)  | 63 (32.0)  | 33 (33.7)  | 36 (36.0)  | 0.731 | 0.846 |
| Unkown           | 0 (0)   | 0 (0)   | 22 (10.1)  | 12 (11.1)  | 10 (9.1)   | 20 (10.1)  | 10 (10.2)  | 10 (10.0)  |        |        |

P*: the difference between the training set and validation set.

a The average age was 61.
b Tumor size measured in greatest transverse diameter (cm), and the mean was 4.7 cm.
significantly promoted migration ability of HT-29 and SW480 cells, while miR-193a-5p induced the opposite effects (Fig. S3). Pathway analysis for these four miRNAs is shown in Fig. S4.

3.3. Development and validation of a 4-miRNA panel to predict LN status of CRC

Through logistic regression analysis, all four miRNAs were all identified as independent predictive factors for LN status in CRC (Fig. 3). A risk score formula of miRNA-based panel was built to predict LN status as follows: Logit (P = LN metastasis) = −1.916 + miR-122-5p*0.495 + miR-146-5p*0.869 + miR-186-5p*0.899 + miR-193a-5p*(-0.377). Subsequently, we calculated the risk scores for all CRC patients, using this formula. Descriptive analyses for distributions of risk scores and LN status in both the training and validation sets showed that LN+ patients generally had higher risk scores (Fig. 4a and 4d). The 4-miRNA panel yielded an AUC of 0.907 (95% CI: 0.860–0.942, Fig. 4b) in the training cohort and 0.870 (95% CI: 0.815–0.913, Fig. 4e) in the validation set (Table S2), while the AUC for CT-reported LN status was 0.623 (95% CI: 0.549–0.697) in the training set and 0.675 (95% CI: 0.600–0.750) in the validation set. The miRNA-based panel could provide better predictive efficacy than conventional CT-reported LN status (P < .0001 in both training and validation sets). Moreover, the miRNA panel showed a good discriminatory ability in the CT-reported LN negative subgroup, with AUC values of 0.797 (95% CI: 0.701–0.892, Fig. 4c) in the training set and 0.764 (95% CI: 0.660–0.868, Fig. 4f) in the validation sets. When the patients were stratified based on the clinicopathological factors, a significant association between the miRNA-based panel and LN status was found in all subgroups (Fig. S5).

3.4. Development and validation of an individualized prediction nomogram

Logistic regression analysis revealed the 4-miRNA panel and the CT-reported LN status were independent risk predictors of LN status (Table 2). The model that incorporated the above independent predictive factors was developed and presented as the nomogram (Fig. 5a). The calibration plot of our nomogram showed the bias-corrected line lay close to the ideal curve (the 45-degree line), implying a good agreement between prediction and observation in the training set (Fig. 5b). The non-significant Hosmer-Lemeshow test statistic (P = .268) indicated a good fit to the model. The ROC analysis yielded an AUC of 0.913 (95% CI: 0.878–0.948, Fig. 5d) for the training set, which implied the discrimination performance was favorable.

In agreement with the training set, the favorable calibration of the nomogram was confirmed in the validation set (Fig. 5c). The AUC of the validation set was 0.883 (95% CI: 0.835–0.930, Fig. 5e). Moreover, the DCA showed that if the threshold probability of a patient or a clinician is >12%, using the nomogram to predict LN status adds more benefit than the “treat-all” or “treat-none” scheme. It also showed higher net benefit than CT-reported LN status (Fig. 6).

4. Discussion

Non-invasive molecular markers to accurately predict LN status before surgery are urgently needed to optimize individually-tailored therapy in CRC. In the present study, we identified a novel serum 4-miRNA signature that discriminated with high accuracy between the serum miRNA profiles of CRC patients with and without LN metastasis. Furthermore, an inclusive nomogram incorporating the 4-miRNA signature and CT-reported LN status was constructed for the prediction of LN metastasis in CRC patients, which displayed satisfactory predictive accuracy in both the training and validation sets.

Circulating miRNAs are promising non-invasive cancer biomarkers with great translational potential to be used in personalized medicine. Nonetheless, despite some advances having been made, much research has focused on a few preselected miRNAs, leaving the majority of miRNAs unexplored. In this study, differential miRNAs expression profiles were initially determined in pooled serum samples from CRC patients with and without LN metastasis using high-throughput sequencing during the discovery stage. With this foundation, significantly differentially expressed miRNAs were further identified by RT-qPCR validation. To the best of our knowledge, this is the first comprehensive analysis of predictive biomarkers for LN metastasis based on circulating miRNA expression profiles in CRC patients.

miRNAs are known to be secreted by various cell types and can be shuttled between cells, and thus modulate gene expressions and cellular activities. Dysregulation of miRNA expression in different types of cells in tumor microenvironments has been documented to contribute to tumor metastasis [34–36]. Considering circulating miRNAs may arise from heterogeneous sources, we not only validated the expression of candidate miRNAs in CRC serum, but also confirmed their effect on the metastatic capacity of CRC cells. Thereafter, the four most promising miRNAs were selected, on the basis that they exhibited differential expression associated with LN metastasis and could exert regulatory effects on the metastatic behavior of CRC cells. Functional enrichment analysis of the KEGG signaling pathway showed the top 20 pathways involved, such as the PI3K-Akt signaling pathway, suggesting these miRNAs serve a critical role in CRC metastasis. Furthermore, these miRNAs have been previously reported to be involved in metastasis of CRC [37–43]. These findings, together with our observations, suggest

![Fig. 3](image-url) Forest plot summary of analyses of LN status. Univariate and multivariate logistic regression for the four predictive miRNAs of LN metastasis in the training set. The squares on the transverse lines represent the odds ratio (OR), and the transverse lines represent the 95% confidence interval (95% CI).
the potential use of the four circulating miRNAs as predictive biomarkers for LN metastasis in CRC patients. However, we failed to detect consistent changes in serum miR-203 and miR-200c as previously reported in CRC patients [22,23]. This conflict in the result may be due to varying ethnic compositions in the samples.

Nomograms are the visualization of statistical models specifically developed to optimize the predictive accuracy of individuals. Preoperative nomograms estimating LN metastases can aid clinicians in identifying patients who may derive greater clinical benefit from more extensive surgery [44–47]. In the current study, we postulated that an inclusive model incorporating a serum miRNA signature and clinical risk factors might improve the accuracy of node staging. Thus, we built a risk score formula of the four-miRNA panel that could provide better predictive efficacy as compared to conventional CT scans, and then we combined all significant independent predictors, including the miRNA-based panel and CT-reported LN status, to a nomogram. Wu et al. previously also presented a noninvasive nomogram model, with an AUC of 0.788 for prediction of LN status in CRC [30]. Their model incorporates a radiomics signature and clinical risk factors, but lacks sensitive and specific molecular markers. The miRNA-based prediction models for operatively predicting LN status have recently been successfully developed in breast and hepatocellular cancer [28,48]. Our study is the first to investigate the usefulness of a 4-miRNA panel as an effective molecular approach for the preoperative prediction of LN status in CRC. Our data suggest this miRNA-based panel could provide clinicians with relatively accurate LN status assessment, and the nomogram may be a useful tool for preoperative prediction of LN metastasis, aiding in individualized management decisions, and ultimately contributing to improved survival among CRC patients.

The most important argument for the adoption of the nomogram into clinical use is to justify whether nomogram-assisted decisions in management could improve patient outcomes. However, current measures of prediction performance, such as predicted-versus-observed tests for model calibration, AUC or concordance index (often used interchangeably) cannot fulfill this prospect. Therefore, we used decision analysis curves to estimate the clinical utility of our prediction

---

**Table 2**

| Clinical variable     | Subgroups                  | Univariate analysis | Multivariate analysis |
|-----------------------|----------------------------|---------------------|-----------------------|
|                       | OR                         | 95%CI               | P                     | OR                         | 95%CI               | P                     |
| Sex                   | Female vs Male             | 0.770               | 0.451–1.317           | 0.340                     |                       |                       |
| Age                   | ≥61 vs <61                 | 1.550               | 0.911–2.638           | 0.106                     |                       |                       |
| Tumor location        | Rectum vs Colon            | 0.886               | 0.507–1.547           | 0.670                     |                       |                       |
| Differentiation       | Well vs Ref                | 1.360               | 0.590–3.271           | 0.451                     |                       |                       |
|                       | Moderate vs Poor           | 2.067               | 0.998–7.124           | 0.050                     |                       |                       |
| CT-reported LN status | Positive vs Negative       | 2.755               | 1.591–4.771           | 0.0003                    | 2.605                 | 1.272–5.335           | 0.009                 |
| CEA level             | Abnormal vs Normal         | 1.832               | 1.044–3.214           | 0.035                     | 1.727                 | 0.817–3.649           | 0.152                 |
| miRNA-based panel     | High risk vs Low risk      | 22.937              | 11.124–47.295         | <0.0001                   | 22.902                | 10.819–48.478         | <0.0001               |

OR: odds ratio; CI: confidence interval.
Fig. 5. The nomogram for preoperative prediction of LN status and its predictive performance. The nomogram to predict probability of LN metastasis for CRC patients in training set, with the miRNA-based panel and CT-reported LN status incorporated (a). Calibration curves of the nomogram in the training (b) and validation (c) sets (bootstrap 1000 repetitions). Nomogram-predicted probability of LN metastasis is plotted on the x-axis and actual probability is plotted on the y-axis. The dashed 45-degree line represents a perfect prediction by an ideal model, and the solid line represents the performance of our nomogram, of which a closer fit to the dashed line means a better prediction. ROC curves based on the nomogram for the probability of LN metastasis in the training (d, AUC = 0.913, 95%CI: 0.878–0.948) and validation (e, AUC = 0.883, 95%CI: 0.835–0.930) sets.
nomogram based on threshold probability—the probability at which the clinician or patient would proceed with some action [29,47,49,50]. The decision curves showed that if the threshold probability is higher than 12%, using the nomogram to predict LN status adds more net benefit than the “treat-all” or “treat-none” scheme.

Our study has the following strengths: A relatively large number of enrolled participants, genome-wide miRNA profiling followed by training set and validation set, identification of a novel miRNA-based panel, and construction of a clinically useful nomogram, which indicates the clinical practicality and innovation of our research. Despite the relatively high predictive ability of our miRNA-based panel, one limitation should be taken into consideration: The present study is a single-center retrospective analysis, with limited generalizability as all subjects are of the same ethnicity, and the distribution of clinical features might be distinct in other regions, making it unsuitable for other races and areas. Therefore, our results should be further validated by prospective research in a multicenter trial on diverse ethnic populations.

In conclusion, our results suggest that the serum miRNA-based panel obtained by gene expressing profiling can be combined with CT evaluation to improve the accuracy of preoperative prediction of LN status. This predictive nomogram model has great applicability potential in the non-invasive clinical evaluation of patients at risk of LN metastasis, and may be conveniently used to optimize treatment strategies by avoiding unnecessary LN-related procedures in CRC.

Availability of data and material

The data sets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Chuanxin Wang conceived and designed the experiments. Ailin Qu, Yongmei Yang and Xin Zhang performed all the experiments. Ailin Qu, Wenfei Wang, Yingjie Liu, Guixi Zheng, and Lutaou Du analyzed the data. Chuanxin Wang, Ailin Qu and Yongmei Yang wrote the manuscript. All authors read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2018.09.052.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7–30.
[2] Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart AK. Revised TN classification for colon cancer based on national survival outcomes data. J Clin Oncol 2010;28:264–71.
[3] Benson AR, Venook AP, Cederquist L, et al. Colon cancer, version 1.2017, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2017;15:370–98.
[4] van de Velde CJ, Ariest E, Roelens PG, et al. EURECCA colorectal: multidisciplinary mission statement on better care for patients with colon and rectal cancer in Europe. Eur J Cancer 2013;49:2784–90.
[5] Nakayama C, Tanaka C, Kodera Y. Current options for the diagnosis and management of colorectal cancer. Gastrointest Tumors 2013;1:25–32.
[6] Bipat S, Glas AS, Slors FJ, Zwinderman AH, Bossuyt PM, Stoker J. Rectal cancer: local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging—a meta-analysis. Radiology 2004;232:773–83.
[7] Shin SS, Jeong YW, Min JJ, Kim HR, Chung TW, Kang HK. Preoperative staging of colorectal cancer: CT vs. integrated FDG PET/CT. Abdom Imaging 2008;33:270–7.
[8] Glasgow SC, Bleier JJ, Burgart LJ, Finne CO, Lowry AC. Meta-analysis of histopathological features of primary colorectal cancers that predict lymph node metastases. J Gastrointest Surg 2012;16:1019–28.
[9] Romano G, Veneziano D, Acunzo M, Croce CM. Small non-coding RNA and cancer. Carcinogenesis 2017;38:485–91.
[10] Barret DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.
[11] Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–66.
[12] Iorio MV, Croce CM. microRNA involvement in human cancer. Carcinogenesis 2012;33:1126–33.
[13] Davood S. Novel biomarkers of metastatic cancer. Expert Rev Mol Diagn 2010;10:581–90.
Xie X, Geng Y, Gu W, Huang J, Pei H, Jiang J. Prognostic role of tissue and circulating microRNA-200c in malignant tumors: a systematic review and meta-analysis. Cell Physiol Biochem 2015;35:1188–200.

Matullo G, Naccarati A, Pardini B. MicroRNA expression profiling in bladder cancer: the challenge of next-generation sequencing in tissues and biofluids. Int J Cancer 2016;138:2334–45.

Zeng K, Zhang CY. Circulating microRNAs: a novel class of biomarkers to diagnose and monitor human cancers. Med Res Rev 2012;32:326–48.

Lindner K, Haier J, Wang Z, Watson DJ, Hussey DJ, Hummel R. Circulating microRNAs: emerging biomarkers for diagnosis and prognosis in patients with gastrointestinal cancers. Clin Sci (Lond) 2015;128:1–15.

Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical relevance of circulating cell-free microRNAs in cancer. Nat Rev Clin Oncol 2014;11:145–56.

Imaoka H, Toyama Y, Ogikami M, et al. Circulating microRNA-203 predicts metastases, early recurrence, and poor prognosis in human gastric cancer. Gastric Cancer 2016;19:744–53.

Zhuo S, Yao D, Chen J, Ding N. Circulating mir-20a and mir-203 for screening lymph node metastasis in early stage cervical cancer. Genet Test Mol Biomarkers 2013;17:631–4.

Inns J, James V. Circulating microRNAs for the prediction of metastasis in breast cancer patients diagnosed with early stage disease. Breast 2015;24:364–9.

Hur K, Toyama Y, Ogikami Y, et al. Circulating microRNA-203 predicts prognosis and metastasis in human colorectal cancer. Gut 2017;66:654–65.

Toyama Y, Hur K, Tanaka K, et al. Serum miR-200c is a novel prognostic and metastasis-predictive biomarker in patients with colorectal cancer. Ann Surg 2014;259:735–43.

Jiang X, Wang W, Yang Y, et al. Identification of circulating microRNA signatures as potential noninvasive biomarkers for prediction and prognosis of lymph node metastasis in gastric cancer. Oncotarget 2017;8:65132–42.

Azizian A, Kramer F, Jo P, et al. Preoperative prediction of lymph node status by circulating Mir-18b and Mir-20a during chemoradiotherapy in patients with rectal cancer. World J Surg 2015;39:2329–35.

Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol 2015;33:861–8.

Shim JH, Jun MJ, Han S, et al. Prognostic nomograms for prediction of recurrence and survival after curative liver resection for hepatocellular carcinoma. Ann Surg 2015;261:939–46.

Xie X, Tan W, Chen B, et al. Preoperative prediction nomogram based on primary tumor miRNA signature and clinical-related features for axillary lymph node metastasis in early-stage invasive breast cancer. Int J Cancer 2018;142:1901–10.

Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.

Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34:2157–64.

Zheng G, Wang H, Zhang X, et al. Identification and validation of reference genes for qPCR detection of serum microRNAs in colorectal adenocarcinoma patients. PLoS One 2013;8:e63025.

Zheng CX, Qu AL, Yang YM, Zhang X, Zhang SC, Wang CX. miR-422a is an independent prognostic factor and functions as a potential tumor suppressor in colorectal cancer. World J Gastroenterol 2016;22:5589–97.