
Kai Li
Yusheng Lin
Yu Zhou
Xiao Xiong
Lu Wang

See next page for additional authors

Follow this and additional works at: https://digitalcommons.pcom.edu/scholarly_papers

Part of the Medicine and Health Sciences Commons

Recommended Citation
Li, Kai; Lin, Yusheng; Zhou, Yu; Xiong, Xiao; Wang, Lu; Li, Junkuo; Zhou, Fuyou; Guo, Yi; Chen, Shaobin; Chen, Yuping; Tang, Hui; Qiu, Xiaofu; Cai, Songwang; Zhang, Dianzheng; Bremer, Edwin; Jim Yeung, Sai-Ching; and Zhang, Hao, "Salivary Extracellular MicroRNAs for Early Detection and Prognostication of Esophageal Cancer: A Clinical Study," (2023). PCOM Scholarly Papers. 2218. https://digitalcommons.pcom.edu/scholarly_papers/2218

This Article is brought to you for free and open access by DigitalCommons@PCOM. It has been accepted for inclusion in PCOM Scholarly Papers by an authorized administrator of DigitalCommons@PCOM. For more information, please contact jaclynwe@pcom.edu.
Salivary Extracellular MicroRNAs for Early Detection and Prognostication of Esophageal Cancer: A Clinical Study

Kai Li,1,2 Yusheng Lin,2,3,4 Yu Zhou,2 Xiao Xiong,1 Lu Wang,2 Junkuo Li,5 Fuyou Zhou,5 Yi Guo,6 Shaobin Chen,7 Yuping Chen,7 Hui Tang,8,9 Xiaofu Qiu,10 Songwang Cai,11 Dianzheng Zhang,12 Edwin Bremer,3 Sai-Ching Jim Yeung,13 and Hao Zhang10,14,15

1Department of Urology, Guangdong Second Provincial General Hospital, Faculty of Medical Science and Integrated Chinese and Western Medicine Postdoctoral Research Station, Jinan University, Guangzhou, Guangdong, China; 2Institute of Precision Cancer Medicine and Pathology, Jinan University Medical College, Guangzhou, Guangdong, China; 3Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 4Graduate School, Shantou University Medical College, Shantou, Guangdong, China; 5Department of Thoracic Surgery, Anyang Tumor Hospital, The Fourth Affiliated Hospital of Henan University of Science and Technology, Anyang, Henan, China; 6Endoscopy Center, Affiliated Cancer Hospital of Shantou University Medical College, Shantou, Guangdong, China; 7Department of Thoracic Surgery, Affiliated Cancer Hospital of Shantou University Medical College, Shantou, Guangdong, China; 8Department of Central Laboratory, The First Affiliated Hospital of Jinan University, Guangzhou, China; 9Department of Clinical Laboratory, The Fifth Affiliated Hospital of Jinan University (Heyuan Shenhe People’s Hospital), Heyuan, China; 10Department of Urology, Guangdong Second Provincial General Hospital, Guangzhou, Guangdong, China; 11Department of Thoracic Surgery, The First Affiliated Hospital of Jinan University, Jinan University, Guangzhou, China; 12Department of Bio-Medical Sciences, Philadelphia College of Osteopathic Medicine, Philadelphia, Pennsylvania; 13Department of Emergency Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas; 14Department of General Surgery, The First Affiliated Hospital of Jinan University, Jinan University, Guangzhou, China; and 15Institute of Precision Cancer Medicine and Pathology, School of Medicine, Minister of Education Key Laboratory of Tumor Molecular Biology, Jinan University, Guangzhou, China

**BACKGROUND & AIMS:** Early detection of esophageal squamous cell carcinoma (ESCC) will facilitate curative treatment. We aimed to establish a microRNA (miRNA) signature derived from salivary extracellular vesicles and particles (EVPs) for early ESCC detection and prognostication. **METHODS:** Salivary EVP miRNA expression was profiled in a pilot cohort (n = 54) using microarray. Area under the receiver operator characteristic curve (AUROC) and least absolute shrinkage and selector operation regression analyses were used to prioritize miRNAs that discriminated patients with ESCC from controls. Using quantitative reverse transcription polymerase chain reaction, the candidates were measured in a discovery cohort (n = 72) and cell lines. The prediction models for the biomarkers were derived from a training cohort (n = 342) and validated in an internal cohort (n = 207) and an external cohort (n = 226).

**RESULTS:** The microarray analysis identified 7 miRNAs for distinguishing patients with ESCC from control subjects. Because 1 was not always detectable in the discovery cohort and cell lines, the other 6 miRNAs formed a panel. A signature of this panel accurately identified patients with all-stage ESCC in the training cohort (AUROC = 0.968) and was successfully validated in 2 independent cohorts. Importantly, this signature could distinguish patients with early-stage (stage I/II) ESCC from control subjects in the training cohort (AUROC = 0.968, sensitivity = 92.00%, specificity = 89.17%) and internal (sensitivity = 90.32%, specificity = 91.04%) and external (sensitivity = 90.32%, specificity = 91.04%) validation cohorts.
Early diagnosis is of paramount importance for cancer management, as survival rates increase significantly when cancers are diagnosed at an early stage.\textsuperscript{1,2} Serum-based biomarkers, imaging techniques, and tissue biopsies play indispensable roles in tumor diagnosis.\textsuperscript{3} However, the use of these techniques for early diagnosis has proved challenging. For instance, traditional serum-based biomarkers do not have the requisite sensitivity and specificity.\textsuperscript{1} With molecular imaging, microscopic lesions are often missed and, therefore, many solid tumors are not effectively diagnosed by imaging until later stages.\textsuperscript{5} Furthermore, although imaging is a mainstay in cancer diagnosis, it is labor-intensive and too expensive for screening a large number of asymptomatic individuals. Finally, tissue biopsies, usually involving a large-core needle, an endoscope, or open surgery, are invasive, risky, costly, painful, and sometimes not feasible due to the inaccessibility of tumors.\textsuperscript{6,7}

An alternative diagnostic tool for early detection of cancer that has gained prominence in recent years is liquid biopsy. Commonly used liquid biopsy techniques include the capture of circulating tumor cells (CTCs), circulating tumor DNA (ctDNA), and extracellular vesicles (EVs).\textsuperscript{8–13} Although CTCs and ctDNA have been well studied,\textsuperscript{12,13} sampling CTCs or ctDNA from peripheral blood remains challenging due to the cell-to-cell heterogeneity, low quantities in the bloodstream, and the complex isolation procedures.\textsuperscript{14–16} These drawbacks, therefore, restrict the use of CTCs and ctDNA for comprehensive characterization of tumors and their application in liquid biopsies.

EVs have attracted the most interest in liquid biopsy due to the ease of sampling, low cost, and enriched biological information. EVs are released by virtually all cell types and are involved in cancer initiation and progression.\textsuperscript{17–19} EVs exist in various types of bodily fluids in adequate amounts\textsuperscript{17,18} and contain various molecular components of their originating cells, including nucleic acids (DNA and various types of RNA), proteins, lipids, and metabolites that reflect the status of the cells.\textsuperscript{20} Analysis of microRNAs (miRNAs) in EVs is of particular interest as a noninvasive cancer biomarker due to the relatively high miRNA levels in EVs and the sensitivity of techniques available.\textsuperscript{21–23} Importantly, miRNAs are selectively sorted into EVs through binding to AGO2 or other RNA-binding proteins and are directly implicated in the pathologic process of cancer.\textsuperscript{24,25} Therefore, they may reflect the cancer status. Novel and high-performance biomarkers are needed in personalized clinical management for patients with cancer. The EV-based liquid biopsy that enables serial sampling in a convenient and noninvasive manner could be a preferred choice.

Esophageal squamous cell carcinoma (ESCC) is an aggressive malignancy of the gastrointestinal tract.\textsuperscript{26,27} ESCC is ranked the sixth most common malignancy in China, and it is the fourth most common cause of cancer-related death worldwide.\textsuperscript{26} Despite significant progress in diagnosis and treatment, dealing with ESCC remains a great challenge because of a lack of specific symptoms at the early stage, considerable metastatic and recurrence potential, as well as resistance to conventional treatment. The 5-year survival rate is only 18%, which translates to a large number of deaths annually due to the lack of effective diagnostic and therapeutic methods.

**WHAT YOU NEED TO KNOW**

**BACKGROUND AND CONTEXT**

Extracellular vesicle–based liquid biopsy is not only noninvasive and effective, but also can provide reliable biomarkers for early detection and prognosis of esophageal carcinoma.

**NEW FINDINGS**

Our newly developed salivary extracellular vesicle and particle microRNA signature can accurately detect early-stage esophageal squamous cell carcinoma and efficiently predict high-risk cases with poor outcome.

**LIMITATIONS**

This study involved Chinese subjects only, recruited at 2 study centers.

**CLINICAL RESEARCH RELEVANCE**

Our salivary 6-microRNA signature holds high potential to impact clinical practice of cancer screening by enabling noninvasive and timely detection of early-stage esophageal squamous cell carcinoma in high-risk populations. In addition, it facilitates risk stratification for personalized treatment strategies, offering a valuable tool for improving patient outcomes.

**BASIC RESEARCH RELEVANCE**

The 6 microRNAs selectively targeting messenger RNAs encoding proteins pivotal in gene regulation, ion binding, and nucleic acid interactions indicate their potential as pivotal players in tumorigenesis and disease progression. This discovery suggests that these microRNAs hold great promise as potential biomarkers for esophageal cancer management.

**Abbreviations used in this paper:** ATH, Anyang Tumor Hospital; AUROC, area under receiver operating characteristics curve; CHSWMG, Cancer Hospital of Shantou University Medical College; CTC, circulating tumor cell; ctDNA, circulating tumor DNA; DS, diagnostic score; ESCC, esophageal squamous cell carcinoma; EV, extracellular vesicle; EVP, extracellular vesicle and particle; HR, hazard ratio; mRNA, messenger RNA; NPV, negative predictive value; OS, overall survival; PFS, progression-free survival; PPV, positive predictive value; RS, risk score.

© 2023 The Author(s). Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). 0016-5085 https://doi.org/10.1053/j.gastro.2023.06.021
of salivary EVP RNA, as shown in Supplementary Figure 5. The CHSUMC cohort was used for constructing the diagnostic and prognostic models and internal validation and the ATH cohort was used for external validation. Computer-generated random numbers were used to assign samples from CHSUMC to the training cohort, consisting of 222 patient samples and 120 control samples, and the internal validation cohort, consisting of 140 patient samples and 67 control samples.

Statistical Analysis

Comparisons of miRNA expression or diagnostic score (DS) between cancer and control groups were performed with the Mann-Whitney U test. Comparisons of the miRNA expression or DS among different groups were evaluated using Kruskal-Wallis test with Dunn’s multiple comparisons test. For comparison of miRNA expression between paired groups of tumor tissue and adjacent normal tissue, a rank-sum test (Wilcoxon matched-pairs signed-rank test) was used.

The sample size for the discovery cohort was determined a priori. Based on the effect size of 0.8, a error probability of .007 for each miRNA, and power of 0.8, there would be 36 cases and 36 controls needed to reject the null hypothesis.

The differences of proportions in clinicopathologic characteristics were analyzed with the Fisher exact test. Binary logistic regression was employed to derive a formula to predict the risk score (RS) of each subject. Area under the receiver operating characteristics curve (AUROC) was used to assess the predictive performance of a 6-member EVP miRNA (6-EVP miRNA) signature. The optimal cutoff value for classification using the 6-EVP miRNA signature was based on the Youden index.

The incidence rates of ESCC in endemic areas of the Chaoshan region and Lin Xian (within Anyang) in China were obtained from previous reports. Both cigarette smoking and alcohol consumption were taken into account to calculate the estimated incidence rates of ESCC. Positive predictive value (PPV) and negative predictive value (NPV) were estimated according to the model specificity, sensitivity, and incidence rates of ESCC:

\[
PPV = \frac{\text{incidence} \times \text{sensitivity}}{\text{incidence} \times \text{sensitivity} + (1 - \text{incidence})(1 - \text{specificity})} \times 100
\]

\[
NPV = \frac{(1 - \text{incidence}) \times \text{specificity} \times (1 - \text{sensitivity})}{(1 - \text{incidence}) \times \text{specificity} + \text{incidence} \times (1 - \text{sensitivity})} \times 100
\]

Survival was analyzed using the Kaplan-Meier method with the log-rank test as well as univariate and multivariate Cox proportional hazards modeling. A stepwise backward approach was applied in the discovery phase to identify the highly predictive miRNAs. Final Cox proportional hazards models were constructed using a 6-EVP miRNA signature. Age, gender, histologic differentiation, tumor length, and stage were used as covariates, and the models were evaluated for validity by calculating Martingale score and Schoenfeld residuals using R package “ggaiddiagnostics.”

We used G*power (https://www.psychologie.hu/d/arbetsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html) for a priori estimation of sample size. All other statistical analyses were conducted using R, version 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org/). P < .05 was considered significant and all tests were 2-sided.
Details for sample collection and the experimental process are included in the Supplementary Materials and Methods.

Results

Identification of Salivary Extracellular Vesicle and Particle MicroRNA Markers in Esophageal Squamous Cell Carcinoma

To identify potential miRNA biomarkers in ESCC non-invasively, a pilot study was conducted comparing the differentially expressed salivary EVP miRNAs in patients with ESCC and control subjects. The morphology and particle size of EVPs were verified using transmission electron microscopy and NanoSight analyses, confirming the presence of oval- or bowl-shaped particles, with a mean diameter of 105 nm (Figure 1A and B). This result suggests that most of the isolated products are EVs. Furthermore, typical EV proteins, such as ALIX, TSG101, CD63, and CD9, were found upon immunoblotting and the non-EV protein calnexin was absent (Figure 1C). The miRNA microarray was conducted to analyze salivary EVP RNA from 25 patients
with ESCC and 29 control subjects (Figure 1D); the pathophysiological characteristics of these 54 participants are shown in Supplementary Table 1. The ROC analysis was performed to investigate which miRNAs can discriminate patients from controls. Fifty-six candidate miRNAs with AUROC > 0.65 were selected for further analyses (Supplementary Table 2 and Supplementary Figure 1A). Seven differentially expressed miRNAs with high potential for discriminating patients with ESCC from the controls were identified (3 up-regulated [ie, miR-4505, miR-142-3p, and miR-1268a] and 4 down-regulated miRNAs [ie, miR-6126, miR-1972, miR-4701-3p, and miR-4274] in patients

![Figure 2. Detection of candidate miRNA biomarkers in patients with ESCC and control subjects. (A) Unsupervised hierarchical clustering of 7 miRNAs selected by least absolute shrinkage and selector operation for discriminating patients with ESCC (n = 25) from control subjects (n = 29). (B) Expression of 7 miRNAs (ie, miR-1972, miR-4274, miR-4701-3p, miR-6126, miR-1268a, miR-4505, and miR-142-3p) were measured by quantitative reverse transcription polymerase chain reaction (RT-qPCR) in a discovery cohort of patients with ESCC (n = 36) and controls (n = 36). Data for miR-142-3p were not shown because it was undetectable by means of RT-qPCR in many samples. Error bars: SEM; **P < .01; ***P < .001 by Mann-Whitney U test.](image)
with ESCC) (Figure 2A, Supplementary Figure 1B, and Supplementary Table 2). In a subsequent discovery cohort of 72 subjects (comprising 36 patients and 36 control subjects; Supplementary Figure 2), 6 miRNAs were reliably measurable and their expression pattern was consistent with findings from the initial microarray assay (Figure 2B). Expression of miR-142-3p proved to be undetectable in most samples (GT values > 35 or undetermined, Supplementary Figure 3A). Furthermore, samples from 9 patients with benign epithelial hyperplasia and 15 patients with gastroesophageal reflux were collected from a biobank to evaluate the expression levels of the 6 detectable miRNAs compared with the samples from the discovery cohort. The levels of all miRNAs were similar between control and patients without cancer, but significant differences were observed in the expression of all miRNAs between control and patients with ESCC, as well as patients without cancer and those with ESCC (Supplementary Figure 3B–G). To compare the levels of these miRNAs in patient tumor tissues and adjacent normal tissues, 10 patients with ESCC from the discovery cohort were randomly selected and quantitative reverse transcription polymerase chain reaction was used for measurement. Significantly higher levels of miR-1268a and miR-4505 and lower levels of miR-1972, miR-4274, miR-4701-3p, and miR-6126 were observed in the tumor tissues of patients with ESCC compared with the adjacent normal tissues (Supplementary Figure 3H). In addition, quantitative reverse transcription polymerase chain reaction was conducted on the EVP miRNAs purified from the culture media of either ESCC or the immortalized human esophageal epithelial cell lines. Consistently, higher levels of EVP miR-1268a and miR-4505 and lower levels of EVP miR-1972, miR-4274, miR-4701-3p, and miR-6126 were observed in the ESCC cells (Supplementary Figure 4). Therefore, these 6 miRNAs were used to further construct the prediction model.

**Construction and Validation of a Diagnostic Model With 6-Extracellular Vesicle and Particle MicroRNA Signature**

To construct and validate a potential diagnostic model, a total of 521 patients with ESCC and 254 control subjects from 2 hospitals in China were enrolled (Supplementary Figure 5). Supplementary Table 3 describes the demographic and clinicopathologic characteristics of the subjects in the CHSUMC training cohort (222 patients with ESCC and 120 control subjects), the CHSUMC internal validation cohort (140 patients with ESCC and 67 control subjects), and the ATH external validation cohort (159 patients with ESCC and 67 control subjects). The demographic and clinicopathologic characteristics were comparable among the control subjects and patients in 3 cohorts (Supplementary Table 3). In the training cohort, the levels of miR-1268a (0.065 ± 0.002) and miR-4505 (0.054 ± 0.002) in patients with ESCC were significantly higher than those of the controls (0.036 ± 0.002 and 0.030 ± 0.002, respectively; both, P < .001, Mann-Whitney U test; Figure 3A). In contrast, miR-1972, miR-4274, miR-4701-3p, and miR-6126 were significantly lower in patients with ESCC (0.048 ± 0.002, 0.021 ± 0.001, 0.022 ± 0.001, 0.029 ± 0.001, respectively) than those of the controls (0.089 ± 0.004, 0.045 ± 0.003, 0.047 ± 0.002, 0.055 ± 0.003, respectively; P < .001 for all, Mann-Whitney U test; Figure 3A). These findings were consistent in both the internal and external validation cohorts (Figure 3B and C).

The diagnostic performance for all 6 miRNAs was evaluated using ROC analysis. Based on the ROC analyses and a stepwise logistic regression model, all 6 miRNAs can be considered significant independent predictors (Supplementary Table 4). Models constructed with all possible combinations of the 6-EVP miRNAs were examined, and the model that included all 6 miRNAs significantly outperformed all other combinations (DeLong test, P < .001; Figure 3D). The AUROC for this 6-EVP miRNA signature was 0.968 (95% CI, 0.953–0.983). A DS was calculated for each subject using a formula based on these 6 miRNAs, weighted by their regression coefficient: DS = –29.826 × miR-1972-45.915 × miR-4274-44.776 × miR-4701-3p-42.413 × miR-6126 + 41.745 × miR-1268a + 63.143 × miR-4505 – 2.491. The values of DS in patients with ESCC were significantly higher than those of the controls (all, P < .001, Mann-Whitney U test; Figure 3E). At the optimal cutoff value (ie, 1.020), the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) obtained for the training cohort were 87.39%, 91.67%, 95.10%, and 79.71%, respectively (Table 1). Using the optimal cutoff value for DS derived from the training cohort (ie, 1.020), the diagnostic performance of DS was confirmed in both the internal and external validation cohorts (sensitivity = 86.43% and 88.05%, respectively; specificity = 91.05% and 88.06%, respectively; PPV = 95.28% and 94.59%, respectively; and NPV = 76.25% and 75.64%, respectively; Table 1).

The PPV and NPV, when applied in screening, will be influenced by the pretest probability (see formulas in Materials and Methods). If the salivary EVP DS score was to be used to screen a population with risk factors of both alcohol and cigarette consumption in high ESCC incidence regions like Shanxi and Anhui,34–36 PPV would be expected from 33.75% to 90.65% and NPV from 86.98% to 99.22% at different ages based on the results from the internal validation cohort and PPV would be expected from 28.00% to 88.10% and NPV from 88.00% to 99.29% at different ages based on the results from the external validation cohort (Supplementary Table 5). These calculations suggested that the DS derived from salivary EVPS would be potentially useful for screening a high-risk population with risk factors for ESCC.

**Application of the 6-Extracellular Vesicle and Particle MicroRNA Signature in Early-Stage Esophageal Squamous Cell Carcinoma**

The diagnostic performance of the 6-EVP miRNA signature in early-stage ESCC was further evaluated. In the training cohort CHSUMC, both miR-1268a and miR-4505 were significantly higher in early-stage patients than in control subjects (both, P < .001, Kruskal-Wallis test with Dunn’s multiple comparisons test; Figure 4A). The levels of miR-1972, miR-4274, miR-4701-3p, and miR-6126 were...
Figure 3. Construction and validation of a salivary diagnostic model to detect ESCC in 3 cohorts by using 6-EVP miRNAs. (A) Expression levels of 6 miRNAs were measured by means of quantitative reverse transcription polymerase chain reaction (RT-qPCR) in a training cohort of patients with ESCC (n = 222) and controls (n = 120). (B) Expression levels of 6 miRNAs were measured by means of RT-qPCR in an internal validation cohort of patients with ESCC (n = 140) and controls (n = 67). (C) Expression of 6 miRNAs were measured by means of RT-qPCR in an external validation cohort of patients with ESCC (n = 159) and controls (n = 67). (D) Comparisons of the best AUROCs in 6 different categories of combinations (from 1 miRNA up to 6 miRNAs) of the 6 miRNAs in the training cohort. The best AUROC in each category was shown in the plot. a, miR-1972; b, miR-4274; c, miR-4701-3p; d, miR-6126; e, miR-1268a; f, miR-4505. (E) The levels of DS were compared between controls and patients with ESCC in 3 independent cohorts. For all the panels in this figure, error bars: SEM; ***P < .001 by Mann-Whitney U test.
significantly higher in controls than in early-stage patients (all, \( P < .001 \), Kruskal-Wallis test with Dunn’s multiple comparisons test; Figure 4A). Similarly, all 6 miRNAs in early-stage patients had significantly different levels from controls in both validation cohorts (Figure 4B and C).

A DS for early-stage ESCC (DSearly) was calculated for each subject using a formula based on these 6 miRNAs weighted by their regression coefficient. DSearly = –28.826 × miR-1972-51.788 × miR-4274-56.366 × miR-4701-3p-44.351 × miR-6126 + 38.766 × miR-1268a + 68.03 × miR-4505 + 2.129. Based on ROC analyses of the training cohort, the 6-EVP miRNA signature had a best AUROC of 0.969 with an optimal cutoff value (–0.137) as a binary classifier chosen by the Youden index to discriminate early-stage patients from controls (\( P < .001 \), DeLong test; Figure 4D). The levels of DSearly in early patients with ESCC were significantly higher than those of the controls (\( P < .001 \) for all, Mann-Whitney U test; Figure 4E). Using the cutoff value of –0.137, sensitivity for identifying early-stage ESCC was 92.00%, specificity was 89.17%, PPV was 87.62%, and NPV was 93.04% (Table 2). The performance of the cutoff value of –0.137 for DSearly was then tested in the internal and external validation cohorts, where sensitivities for identifying early-stage ESCC were 90.32% and 91.07%, respectively; specificities were 91.04% and 88.06%, respectively; PPVs were 90.32% and 86.44%, respectively; and NPVs were 91.04% and 92.19%, respectively (Table 2). Taken together, these data demonstrated that the 6-EVP miRNA signature was capable of distinguishing patients with early-stage ESCC from the controls.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Cohort</th>
<th>Cancer</th>
<th>Test positive, n</th>
<th>Test negative, n</th>
<th>Total, n</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>Training</td>
<td>Absent</td>
<td>10</td>
<td>110</td>
<td>120</td>
<td>87.39</td>
<td>91.67</td>
<td>95.10</td>
<td>79.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>194</td>
<td>28</td>
<td>222</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>204</td>
<td>138</td>
<td>342</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Internal validation</td>
<td>Absent</td>
<td>6</td>
<td>61</td>
<td>67</td>
<td>127</td>
<td>86.43</td>
<td>91.05</td>
<td>95.28</td>
<td>76.25</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>121</td>
<td>19</td>
<td>140</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>127</td>
<td>80</td>
<td>207</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>External validation</td>
<td>Absent</td>
<td>8</td>
<td>59</td>
<td>67</td>
<td>–</td>
<td>88.05</td>
<td>88.06</td>
<td>94.59</td>
<td>75.64</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>140</td>
<td>19</td>
<td>159</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>148</td>
<td>78</td>
<td>226</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

NOTE. The cutoff value calculated in the training cohort was applied to the internal validation and validation cohorts. Test positive in this analysis is based on a miRNA signature score higher than cutoff value (ie, 1.02); the remaining individuals were classified as test negative.

Establishment and Validation of a Risk-Stratification Model for Early-Stage Esophageal Squamous Cell Carcinoma Using the 6-Extracellular Vesicle and Particle MicroRNA Panel

For further investigation and potential clinical application, a prognostic RS based on the 6-EVP miRNA panel was generated to stratify patients with favorable clinical outcomes. The risk-stratification model with the coefficients weighted by the Cox regression model was established in the training cohort: RS = –14.138 × miR-1972-17.253 × miR-4274-6.546 × miR-4701-3p-15.913 × miR-6126 + 10.685 × miR-1268a + 7.753 × miR-4505. A median cutoff value for the RS was chosen to categorize patients into a high-risk or low-risk group (Supplementary Figure 6A and Supplementary Table 6). High risk is associated with greater tumor depth, lymph node metastasis, and poor histologic differentiation (\( P = .001 \), \( P < .001 \), respectively; Fisher exact test; Supplementary Table 6) and with a higher probability of earlier death than those with low RS in all 3 cohorts (Supplementary Figure 6B–D). Kaplan-Meier analysis identified that a high RS was associated with both shorter overall survival (OS) and progression-free survival (PFS) (both, \( P < .001 \), log-rank test; Figure 5A and 5D) in the training cohort. Similarly, patients with high RS had worse OS and PFS than those with low RS in the internal (both, \( P = .005 \), log-rank test; Figure 5B and 5E) and external cohorts (both, \( P < .001 \), log-rank test; Figure 5C and 5F). The univariate Cox regression analysis with the clinicopathologic factors and the 6-EVP miRNA signature showed a significantly higher risk of progression in patients with larger tumor dimensions, higher TNM stage, and higher RS in the training cohort (Supplementary Tables 7 and 8). Multivariable Cox regression analysis revealed that the 6-EVP miRNA signature-based RS was an independent predictor of OS (as a continuous variable: hazard ratio [HR], 2.74; 95% CI, 1.97–3.82; \( P < .001 \), Supplementary Table 7) and PFS (as a continuous variable: HR, 2.45; 95% CI, 1.66–3.61; \( P < .001 \), Supplementary Table 9) as a categorical variable: HR, 2.73, 95% CI, 1.67–4.45; \( P < .001 \), Supplementary Table 10). In addition, a median cutoff value for RS was chosen to categorize patients...
with different stages into a high-risk or a low-risk group; and patients with high risk had a higher probability of death earlier than those with low risk in the training cohort (Supplementary Figure 7A–D). The OS for early-stage patients with a higher RS is significantly shorter than that for those with a lower RS, albeit they were at the same stages.
Table 2. Performance of the Early Diagnostic Score to Differentiate Patients With Early-Stage Esophageal Squamous Cell Carcinoma From Control Subjects in Multiple Cohorts

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Cohorts</th>
<th>Cancer</th>
<th>Test positive, n</th>
<th>Test negative, n</th>
<th>Total, n</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early DS</td>
<td>Training</td>
<td>Absent</td>
<td>13</td>
<td>107</td>
<td>120</td>
<td>92.00</td>
<td>89.17</td>
<td>87.62</td>
<td>93.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>92</td>
<td>8</td>
<td>100</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>105</td>
<td>115</td>
<td>220</td>
<td>90.32</td>
<td>91.04</td>
<td>90.32</td>
<td>91.04</td>
</tr>
<tr>
<td></td>
<td>Internal validation</td>
<td>Absent</td>
<td>6</td>
<td>61</td>
<td>67</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>56</td>
<td>6</td>
<td>62</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>62</td>
<td>67</td>
<td>129</td>
<td>91.07</td>
<td>88.06</td>
<td>86.44</td>
<td>92.19</td>
</tr>
<tr>
<td></td>
<td>External validation</td>
<td>Absent</td>
<td>8</td>
<td>59</td>
<td>67</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>51</td>
<td>5</td>
<td>56</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>59</td>
<td>64</td>
<td>123</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

NOTE. The cutoff value calculated in the training cohort was applied to the internal validation and validation cohorts. Test positive in this analysis is based on a miRNA signature score higher than cutoff value (ie, ~0.137); the remaining individuals were classified as test negative.

(P = 0.006, P < .001, and P = 0.002 for stage I, stage II, and stage III patients, respectively, in the log-rank test; Supplementary Figure 8). These findings indicated that the 6-EVP miRNA signature is a valid prognosticator for OS and PFS in patients with ESCC.

To have an overview of the biological functions associated with this 6-EVP miRNA signature, the target messenger RNAs (mRNAs) of the 6 miRNAs were used to perform gene ontology enrichment analyses. The target mRNAs were significantly enriched in different aspects of biological processes, including biosynthesis and metabolism (Supplementary Table 11). In addition, the target mRNAs were involved in ion binding, nucleic acid binding transcription factor activity, and enzyme binding molecular functions, which were important for gene regulation (Supplementary Table 11). Therefore, these miRNAs that we have identified in this report have the potential to play significant roles in cancer progression.

Discussion

In this multicenter and prospective cohort study, we performed a comprehensive biomarker discovery program and identified a preoperative, saliva-based, EVP miRNA panel for the diagnosis of patients with early-stage ESCC. Specifically, we developed and validated novel diagnostic and prognostic tools based on the expression of 6 EVP miRNAs in saliva. This 6-EVP miRNA signature had high accuracy for the diagnosis of ESCC, especially for patients with early-stage ESCC. Furthermore, a risk-stratification model based on these 6 EVP miRNAs effectively categorized patients with ESCC into high-risk and low-risk groups with significantly different OS and PFS.

RNAs are selectively encapsulated into EVs and accurately reflect the state of the originating cells. Compared with tissue sampling, EV-based strategy minimized the impact by tumor heterogeneity and may therefore be highly valuable for biomarker discovery. A few gene-expression profiling studies on ESCC tissues or cell lines have been performed to discover diagnostic biomarkers, which were subsequently validated in circulating EVs of subjects. However, the RNA profile of EVs differs from that of parental cellular RNA, suggesting that differentially expressed genes identified in tumor tissues or cell lines are not the same as those in EVs. In contrast, our study applied

Figure 4. The assessment of salivary diagnostic model to detect early-stage ESCC in 3 cohorts. (A) Expression of 6 miRNAs were measured by means of quantitative reverse transcription polymerase chain reaction (RT-qPCR) in a training cohort of patients with early-stage ESCC (n = 100), patients with advanced-stage ESCC (n = 122), and controls (n = 120). (B) Expression of 6 miRNAs were measured by means of RT-qPCR in an internal validation cohort of patients with early-stage ESCC (n = 62), patients with advanced-stage ESCC (n = 78), and controls (n = 67). (C) Expression of 6 miRNAs were measured by means of RT-qPCR in an external validation cohort of patients with early-stage ESCC (n = 56), patients with advanced-stage ESCC (n = 103), and controls (n = 67). (D) Comparisons of the best AUROCs in 6 different categories of combinations (from 1 miRNA up to 6 miRNAs) of the 6 miRNAs in the training cohort. The best AUROC in each category was shown in the plot. a, miR-1972; b, miR-4274; c, miR-4701-3p; d, miR-6126; e, miR-1268a; f, miR-4505. E, The levels of DS for early-stage ESCC (DSearly) were compared between controls and patients with early-stage ESCC in 3 independent cohorts. For all the panels in this figure, error bars: SEM; N.S, nonsignificant; ***P < .001 by Kruskal-Wallis test with inter-group comparisons using Dunn’s test (A–C), P < .001 by Mann-Whitney U test (E).
miRNA microarrays to the salivary EVPs of the participants to identify differentially expressed salivary EVP miRNA to begin with, and then subjected the candidate miRNA biomarkers to large-scale, independent validation with salivary EVP samples from 775 participants from 3 independent cohorts. Directly generating the miRNA profiles in salivary EVPs enabled the discovery of relevant biomarkers. The consistency in the samples for profiling and validation ensured reproducibility, which is crucial for further validation in larger cohorts and future incorporation into clinical practice.

Among the panel of 6 miRNAs that we identified, miR-1268a and miR-1972 are the most widely studied. A previous study showed that postoperative adjuvant transarterial chemoembolization treatment had no effects on miR-1268a expression.44 Besides, miR-1268a was shown to mediate temozolomide resistance in glioblastoma45 and a high level of EVP miR-1268a was found in colorectal adenoma organoids.46 The tumor suppressor role of miR-1972 has been investigated in chronic myeloid leukemia,47 osteosarcoma,48 ovarian cancer,49 prostate cancer,50 and papillary thyroid carcinoma.51 Up-regulation of miR-4505 is reported to be associated with lymph node metastasis in intramucosal gastric cancer.52 With regard to miR-4274, its regulating role on oncogene LAMA4 was reported in basal-like breast cancer.53 By directly targeting integrin β1, miR-6126 acts as a tumor suppressor.54 Almost all of these miRNAs except miR-4701-3p have already been shown to play significant roles in cancer, as either oncogenes or tumor suppressors. Together with previous findings, our 6-EVP miRNA signature may implicate in multiple cancer progression and have the great potential in identifying multiple cancer types. To further elucidate the role of these 6-EVP miRNAs in ESCC, gene ontology enrichment analyses were conducted and found that the miRNAs targeted by these miRNAs were significantly enriched and their encoded proteins can bind different ions, nucleic acids, transcription factors, and enzymes involving in gene regulations.

Diagnosis of ESCC in early stages can permit curative treatment.25 However, early detection of ESCC currently involves nonsensitive (imaging), invasive (eg, endoscopy or biopsy), or minimally invasive (eg, esophageal string test, Cytosponge [Medtronic], and transnasal endoscopy) approaches.55,56 Compared with the existing options, the test reported here is more favorable, with greater comfort, convenience, and acceptability. Importantly, the EV-based test allows repeated sampling and offers more comprehensive cancer information, and is impacted less by tumor cell heterogeneity compared with evaluating pieces of tumor samples. Regarding noninvasive biomarker discovery, a limited number of studies have proposed that circulating miRNAs or EVP miRNAs may serve as a diagnostic
biomarker for ESCC, and most of these studies reported AUROCs of approximately 0.8.\textsuperscript{41,42,57} Considering nearly all of those studies included patients with ESCC of different stages, the performance of our 6-EVP miRNA signature in identifying patients with early-stage ESCC with an AUROC of 0.969 is excellent and promising. Thus, our results highlight the diagnostic potential of salivary EVP miRNA as a novel type of ESCC biomarker.

Recognizing the benefits of salivary EV-based liquid biopsy, we have focused on discovering the potential roles of salivary EVs in cancer diagnosis for years. In a previous study, we found that salivary EV GOLM1-NAA35 chimeric RNA could be used successfully to detect early-stage ESCC,\textsuperscript{48} indicating that salivary EVs can serve as a cost-effective diagnostic tool for patients with cancer. Indeed, many advantages have been attributed to the use of saliva over other bodily fluids, including easy and inexpensive sampling and minimal discomfort, as well as reduced risk of infection.\textsuperscript{59–62} Yet, this first-time utility of salivary EVs for early ESCC diagnosis will require further investigations in this emerging field.

Clinically effective management of cancer needs to combine early diagnosis with risk stratification, and this study showed that salivary EVP miRNAs can be used for both early diagnosis and stratification. Like most solid tumors, the TNM staging system has remained central to prognostication and treatment guidance for ESCC. However, it has been realized that the TNM system, based on limited anatomic factors, does not provide adequate and accurate information for personalized treatment. For early-stage ESCC, esophagectomy without any adjuvant treatment is widely considered the treatment of choice.\textsuperscript{62} However, the occurrence and development of ESCC is complex, and prognosis in some cases of postoperative patients with early-stage ESCC remains poor.\textsuperscript{63} Given that even higher tumor heterogeneity exists in advanced ESCC, the clinical outcomes and prognosis of patients still differ a lot, even if they are at the same stage and receive similar treatment.\textsuperscript{64} Compared with TNM, our RS could better stratify patients into prognostic groups and improve the accuracy of survival prediction. Especially for the early-stage, low-risk patients stratified by the RS, excessive and expensive treatments can potentially be avoided (Supplementary Figure 9).

Cigarette smoking, alcohol consumption, and age are well-established risk factors for ESCC.\textsuperscript{65} The incorporation of higher-risk factors resulted in a substantial rise in the PPV of the signature. The PPV of our DS reached 91.33% when considering the pretest probability of the most high-risk population. Individuals older than 55 years of age exhibit a PPV of approximately 80% or higher. Thus, the DS has excellent potential for ESCC screening among people who use alcohol and cigarettes, especially in high-risk regions of China. The efficacy of screening tests depends on the pretest probability, suggesting that the DS may be a valuable screening tool for ESCC in other countries with high incidence rates of the disease.

The levels of circulating biomarkers are affected by a variety of individual characteristics, including gender, age, ethnicity, genetic background, lifestyle, and disease history. Therefore, including more participants with the different aforementioned factors and more study centers will be needed to translate our results to clinical practice.

In conclusion, our results showed the potential of a 6-member salivary EVP miRNA panel as noninvasive markers for identifying the patients with early-stage ESCC and predicting individuals with high risk for poor clinical outcomes. DS and RS based on this panel may be useful in ESCC screening in high-risk populations, as well as risk stratification to guide individualized treatment strategies.

Supplementary Material
Note: To access the supplementary material accompanying this article, visit the online version of Gastroenterology at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2023.06.021.

References


46. Handa T, Kuroha M, Nagai H, et al. Liquid biopsy for colorectal adenoma: is the exosomal miRNA derived


Author names in bold designate shared co-first authorship.

Received October 12, 2022. Accepted June 27, 2023.

Correspondence
Address correspondence to: Hao Zhang, MD, PhD, Institute of Precision Cancer Medicine and Pathology, Jinan University Medical College, Guangzhou, 601 Huangpu Avenue West, Guangzhou, Guangdong 510632, China. e-mail: haolabcancercenter@163.com.

Acknowledgments
Kai Li and Yusheng Lin contributed equally to this work.

CreditAuthorship Contributions
Kai Li, PhD (Data curation: Lead; Investigation: Lead; Methodology: Lead; Visualization: Lead; Writing – original draft: Lead).
Yusheng Lin, PhD (Formal analysis: Equal; Methodology: Equal; Validation: Equal; Writing – original draft: Lead).
Ziyi Li, PhD (Data curation: Equal; Formal Analysis: Equal; Methodology: Equal; Validation: Equal).
Yongjun Cui, PhD (Data curation: Equal; Formal Analysis: Equal; Methodology: Equal; Validation: Equal).
Kai Li and Yusheng Lin contributed equally to this work.

Funding
The work was supported in part by grants from the National Natural Science Foundation of China (82072683, 81777308, 81071736, 81572976, and 30973058 to Hao Zhang); the National Natural Science Foundation of Guangdong Province of China (2021A1515012022 and 2018A030317024 to Hao Zhang); the Science and Technology Planning Project of Guangdong Province of China (2019A030317024 to Hao Zhang); and Special Project on the Integration of Industry, Education and Research of Guangdong Province of China (2019001000024 to Hao Zhang).

Data Availability Statement
All data obtained and/or analyzed in this study are available from the corresponding author upon reasonable request.
Supplementary Figure 1. Discovery of candidate miRNA biomarkers to distinguish patients with ESCC and control subjects. (A) Unsupervised hierarchical clustering of 56 miRNAs with AUROC > 0.65 for discriminating patients with ESCC (n = 25) from control subjects (n = 29). (B) Least absolute shrinkage and selector operation (LASSO) accuracy profiles of the 56 candidate miRNAs. The accuracy (y-axis) was plotted against variable numbers, and the combination of 7 miRNAs with highest accuracy were selected to build the miRNA signature.
Supplementary Figure 2. Flow diagrams accounting for patient numbers in discovery cohorts.
Supplementary Figure 3. Expression of miRNAs in controls, benign esophageal disease, and cancer. (A) The scatter plot showed the various Ct values of patients with ESCC (red) and control subjects (blue). The expression of miR-1972 (B), miR-4274 (C), miR-4701-3p (D), miR-6126 (E), miR-1268a (F), and miR-4505 (G) in control subjects (blue), patients with gastroesophageal reflux disease (GORD, green), patients with benign epithelial hyperplasia (BH, orange), and patients with ESCC (red) was plotted. Error bars: SEM. N.S, nonsignificant. *P < .05; **P < .01; ***P < .001 by Kruskal-Wallis test with inter-group comparisons using Dunn’s test. (H) The expression of miR-1972, miR-4274, miR-4701-3p, miR-6126, miR-1268a, and miR-4505 in tumor tissue (red) and adjacent normal tissue (blue) was plotted. *P < .05; **P < .01 by Wilcoxon matched-pairs signed-rank test.
Supplementary Figure 4. Detection of candidate miRNA biomarkers in EVPs derived from ESCC cell lines. Quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of 6 miRNAs in a panel of ESCC cell lines (filled bars) and immortal normal esophageal epithelial cell lines (open bar). The expression of miR-1972 (A), miR-4274 (B), miR-4701-3p (C), miR-6126 (D), miR-1268a (E), and miR-4505 (F) in 13 ESCC cell lines (filled bar) and 3 immortal normal esophageal epithelial cell lines (open bar) was plotted.
Supplementary Figure 5. Flow diagrams accounting for patient numbers in the 2 patient cohorts. (A) The CHSUMC cohort (training and internal validation cohorts). (B) The ATH cohort (external validation cohort).
Supplementary Figure 6. Prognostic analysis of the RS in multicohorts. (A) The distribution and median value of the RS in the training cohort. (B) The distributions of OS status, OS, and RS (left panel), as well as PFS status, PFS, and RS (right panel) in the training cohort. (C) The distributions of OS status, OS, and RS (left panel), as well as PFS status, PFS, and RS (right panel) in the internal validation cohort. (D) The distributions of OS status, OS, and RS (left panel), as well as PFS status, PFS, and RS (right panel) in the external validation cohort.
Supplementary Figure 7. Prognostic analysis of the RS in patients with ESCC with different stages in the training cohort. The distribution and median value of the RS (left panel), and the distributions of OS status, OS, and RS (right panel) in patients with stage I ESCC (A), stage II ESCC (B), stage III ESCC (C), and stage IV ESCC (D).
Supplementary Figure 8. Kaplan-Meier curves of OS according to the RS in training cohort. (A) Patients with stage I ESCC, (B) patients with stage II ESCC, (C) patients with stage III ESCC, and (D) patients with stage IV ESCC. P values were calculated using the log-rank test.
Supplementary Figure 9. Potential clinical management algorithm for ESCC incorporating the measurement of 6 miRNAs in salivary EVPs as a decision point.