The efficacy of Raman spectroscopy in the diagnosis of esophageal cancer: a systematic review and meta-analysis

Jianqi Hao1,2, Cong Chen1,2, Hongyu Jin2, Nan Chen1,2, Jian Zhou1,2, Yuzhou Zhu2, Kayi Chung2, Qiang Pu1,3

1Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu, China; 2West China School of Medicine, Sichuan University, Chengdu, China; 3Western China Collaborative Innovation Center for Early Diagnosis and Multidisciplinary Therapy of Lung Cancer, Sichuan University, Chengdu, China

Contributions: (I) Conception and design: Q Pu, J Hao; (II) Administrative support: Q Pu; (III) Provision of study materials or patients: H Jin, K Chung; (IV) Collection and assembly of data: C Chen, Y Zhu; (V) Data analysis and interpretation: J Hao, J Zhou, N Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Qiang Pu. Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu, China. Email: puqiang100@163.com.

Background: Esophageal cancer is characterized by its high occurrence rate and difficulty to treat successfully, therefore early diagnosis is extremely important. Our article is aimed to comprehensively analyze the relative effectiveness of Raman spectroscopy (RS) in the diagnosis of suspected esophageal cancer.

Methods: We performed a systemic search in PubMed, EMBASE, CNKI and Web of science from 2007 to 2020. We used the diagnostic data of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) to estimate the pooled sensitivity, specificity, positive and negative likelihood ratios (LR), and diagnostic odds ratio (DOR), with 95% confidence intervals (CI), based on random effects models. The SROC curve analysis and AUC were performed to estimate the overall performance of Raman spectroscopy. QUADAS-2 guidelines were used to evaluate the quality of each study.

Results: Nine articles were included according to our inclusion and exclusion criteria. General pooled diagnostic sensitivity and specificity of RS to esophageal cancer were 0.91 (95% CI, 0.89–0.93) and 0.92 (95% CI, 0.91–0.94). The pooled PLR and NLR were 18.98 (95% CI, 6.61–54.49) and 0.09 (95% CI, 0.05–0.16), respectively. The DOR was 217.21 (95% CI, 68.32–690.53) indicating high accuracy. And the AUC of SROC curve was 0.9779.

Conclusions: We found a promisingly high sensitivity and specificity of RS in the diagnosis of suspected esophageal mass.

Keywords: Raman spectroscopy; diagnostic efficacy; esophageal cancer

Submitted Feb 05, 2020. Accepted for publication Jul 08, 2020.
doi: 10.21037/tcr-20-854
View this article at: http://dx.doi.org/10.21037/tcr-20-854

Introduction

Esophageal cancer is a malignant lesion formed by exceed proliferation of esophageal squamous epithelium or glandular epithelium. Esophageal cancer is the eighth most common malignant tumors in the world (1,2), which is characterized by its high occurrence rate and difficulty to treat successfully following the thoracic surgery. Especially in China, the incidence of esophageal cancer ranks fifth with an average annual death toll of about 150,000 (3,4). The treatment of early esophageal cancer is comparatively simple, and the postoperative survival rate is also high. The 5-year survival rate of early esophageal squamous cell carcinoma is 62.9–92.6%, and the advanced stage is only...
about 30% (5,6). However, early esophageal cancer often has no symptoms. When symptoms such as swallowing obstruction appearing, it is always advanced. It suggests that early diagnosis of esophageal cancer from routine physical examination will undoubtedly have important significance for the treatment of patients with esophageal cancer (7).

In order to enhance early diagnosis and treatment, more frequent use is imaging techniques, especially thoracic computed tomography (CT). For patients with suspected esophageal masses, who expect long-term surveillance or non-surgical interventions, an esophageal mass biopsy (EMB) is usually recommended to observe pathological features to guide subsequent treatment options (8). However, a number of factors may have an impact on EMB diagnostic accuracy, including the cut-off number of specimens and the interpretive degrees of pathologists (9). Accordingly, the diagnostic accuracy of EMB varies from 79% to 100%, important clinical diseases easy to be incorrectly identified including fibrosis and necrosis (10). Therefore, a more trustworthy diagnostic method is urgently necessary.

Recently, it has reported that Raman spectroscopy (RS) has been applied clinically to determine the benign and malignant essence of tumor in surgeries for its ability to optically characterize the internal compositional properties (11,12). What's more, RS examination can be carried out ex or in vivo and it also can be a real-time, label-free and nondestructive (13). Theoretically, RS detects a variation of wave-length or Raman shift resulted from the inelastic light scattering from certain molecules (14). Different molecules have distinct combinations of Raman shifts which can produce unique spectral signatures (15). Therefore, Raman spectra are markedly related to the internal compositional features of tissues of different properties. Importantly, the availability to be performed in vivo, label-free, real-time and non-destructively perfectly addresses the deficiencies of traditional EMB. In the past decade, many studies concerned on the accuracy, sensitivity, and specificity of RS in determining the quality of unknown esophageal mass varied widely from one another, and some studies failed to recruit a sufficient number of patient samples, which could lead to potential bias and inaccuracy (16). Therefore, in order to comprehensively analyze the accurate diagnostic efficiency of RS in determining the benign and malignant features of esophageal tumors, mainly on the parameters of sensitivity, specificity and accuracy, we conducted meta-analysis and systematic evaluation to determine the value of clinical RS. We present the article in accordance with the PRISMA reporting checklist (available at http://dx.doi.org/10.21037/tcr-20-854).

**Methods**

**Literature research**

Two independent reviewers conducted a systemic search through PubMed, EMBASE, CNKI and Web of science. The combinations of the relevant medical subject heading (MeSH) terms, key words, and word variants for “esophagus”, “esophageal”, “Raman spectra” were performed for the primary search. All available studies were published up to January 2020 with no other special limitation, if any discordance happened, we resolved it by consensus. In addition to searching online and in order to identify potentially eligible articles, we also hand-searched the bibliographies of review articles.

**Selection criteria and exclusion criteria**

As a meta-analysis, articles like single sample experiment, comments, case report, letters review articles and editorials were eliminated from the study. Finally, remaining studies were carefully selected when meeting the significant criteria as follows: (I) without animal tissues in the experiments; (II) reported the use of RS in esophageal cancer; (III) recruiting time from January 2007; (IV) used histopathology to confirm the diagnosis; (V) reported the true positive (TP), false positive (FP), true negative (TN) and false negative (FN), based on which the sensitivity and specificity values can be calculated.

Exclusion criteria: (I) studies that involved nonhuman subjects, (II) studies with no relevant data on diagnostic performance; (III) studies like case reports and case series were excluded. If a potential discrepancy was detected, a blinded third reviewer was assigned to adjudicate the conflict.

**Data extraction**

The parameters were extracted by two independent experienced investigators with a standard extraction table, if we had any confliction, we would launch a discussion to reach a consensus. The listed information of the essays was extracted basic information like title, author, nationality and enrolled year. Then the primary parameters, which indicated the diagnostic value, including TP, FP, TN and
FN were extracted from all the included studies.

**Statistical analysis**

The diagnostic value of RS for esophageal cancer was evaluated by using the primary data of TP, FP, TN and FN. Then pooled sensitivity, specificity, positive and negative likelihood ratios (LR), and diagnostic odds ratio (DOR) were integrated by random effects models with their P values and 95% CIs. Sensitivity and specificity were shown in forest plots using Meta-Disc version 1.4 statistical software.

Meanwhile, combination of sensitivity and specificity were assessed by summary receiver operator characteristics (SROC). The overall diagnostic accuracy of RS was estimated by the area under the curve (AUC) of SROC. An AUC value of greater than 0.9 defined a diagnostic tool as excellent, while between 0.8 and 0.9 as good, between 0.7 and 0.8 as fair and less than 0.7 as poor. The SROC curves were also made through Meta-Disc version 1.4. At the same time, according to the detection method *in vivo* and *in vitro*, we conducted a subgroup analysis.

**Quality assessments and publication bias**

We evaluated the quality of included studies through the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) guidelines by Review Manager 5.3. We evaluated publication bias through Deeks Funnel Plot Asymmetry Test (consider the existence of publication bias when P<0.05) by Stata 14.2 (StataCorp, USA).

**Results**

**Search results**

The initial process of study screening was shown in Figure 1. A total of 69 potential correlative articles were searched in PubMed and other databases in January 10th 2020, finally 9 articles (17-24) were eligible on the basis of our criterion.

**Characteristics of the included studies**

The general characteristics of the 9 eligible articles were summarized in Table 1. After searching the authenticated databases, article screening and quality assessment process, 9 articles with high quality and reliability, and comprehensible design with full texts and accessible data were considered for this systematic review. All included researches were published in English except two which was published in Chinese. The total number of patients incorporated was 695 with one study didn’t report its number of patients, and the total number of spectra incorporated was 3,834. And specimens were collected from patients from January 2007 to January 2020. Diagnostic algorithm includes leave-one patient-out, cross validation (LOPCV) in one article, principal components analysis (PCA) in one article, linear discriminate analysis (LDA) in three articles and partial least squares-discriminant analysis (PLS-DA) in two articles. In terms of the nationalities, 4 studies were from China (n=4), 3 the studies were from the same team of Singapore (n=3) and others were performed in Japan (n=1) and England (n=1). Of all the studies, 3 studies detected the esophageal mass by *in vivo* tissues, and 4 studies detected the esophageal mass by *ex vivo* tissues, whereas the other two studies were measured from urine and hemoglobin. Histopathology was used as the golden standard to confirm the diagnosis.

**Pooled results**

Among the nine studies, three studies used RS to screen suspected Barrett’s esophagus. However, only one of the three studies reported the date of TN, FN, TP, FP. So, we cannot conduct a subgroup analysis about the diagnosis of Barrett’s esophagus. Additionally, all the nine studies used RS to tell apart the benign and malignant feature of a particular tissue during an operation. And since some studies were conducted *ex vivo* and other studies were performed *in vitro*. Thus, we did the subgroup evaluation, which was divided into *ex vivo* and *in vitro* groups.

**General pooled data**

The sensitivity of the seven included articles which used RS to screen esophageal cancer with a particular tissue during an operation, ranged from 0.81 (95% CI, 0.66–0.91) in a study with 91 samples to 0.97 (95% CI, 0.94–0.99) which recruited 48 patients (1,172 spectra). The pooled sensitivity was 0.91 (95% CI, 0.89–0.93), which indicated a relatively low incidence rate of missed diagnosis. Particularly, among the seven included studies, except for one with sensitivity of 81%, the other six studies all maintained a sensitivity more than 85%. The forest plot of pooled sensitivity of all the seven studies was shown in Figure 2A.

The specificity of the seven studies ranged from 0.85 (95% CI, 0.82–0.87) in a study with 373 patients (200
spectra) to 1.00 (95% CI, 0.94–1.00) by studies with 64 patients (128 spectra). The general pooled specificity was 0.92 (95% CI, 0.91–0.94), which was also a satisfactory parameter indicating a comparatively low rate of incorrect diagnosis. The forest plot of pooled specificity of all the seven studies was shown in Figure 2B.

The pooled PLR and NLR were 18.98 (95% CI, 6.61–54.49) and 0.09 (95% CI, 0.05–0.16), respectively. The DOR was 217.21 (95% CI, 68.32–690.53) indicating high accuracy. The overall diagnostic accuracy was evaluated through the SROC curve analysis. And the AUC of the SROC curve was 0.9779. The plots were shown in Figure 2C.

Subgroup analysis

Ex vivo group

Of all the included studies, five studied conducted an
<table>
<thead>
<tr>
<th>Reference</th>
<th>Country</th>
<th>N1</th>
<th>N2</th>
<th>N3</th>
<th>Age</th>
<th>Diagnostic algorithm</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Vivo or vitro</th>
<th>Spectra</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almond, 2013</td>
<td>UK</td>
<td>62</td>
<td>673</td>
<td>798</td>
<td>NA</td>
<td>LDA</td>
<td>172</td>
<td>10</td>
<td>27</td>
<td>241</td>
<td>86%</td>
<td>88%</td>
<td>Vitro</td>
<td>830 nm</td>
<td>U</td>
</tr>
<tr>
<td>Bergholt, 2011</td>
<td>Singapore</td>
<td>107</td>
<td>207</td>
<td>1189</td>
<td>66</td>
<td>LDA</td>
<td>30</td>
<td>14</td>
<td>3</td>
<td>216</td>
<td>91%</td>
<td>93.9%</td>
<td>Vivo</td>
<td>785 nm</td>
<td>94.7%</td>
</tr>
<tr>
<td>Bergholt, 2014</td>
<td>Singapore</td>
<td>373</td>
<td>NA</td>
<td>200</td>
<td>NA</td>
<td>U</td>
<td>67</td>
<td>110</td>
<td>10</td>
<td>610</td>
<td>87%</td>
<td>85%</td>
<td>Vivo</td>
<td>785 nm</td>
<td>90%</td>
</tr>
<tr>
<td>Feng, 2016</td>
<td>China</td>
<td>169</td>
<td>NA</td>
<td>149</td>
<td>NA</td>
<td>PLS-DA</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>100%</td>
<td>100%</td>
<td>Vitro</td>
<td>785 nm</td>
<td>100%</td>
</tr>
<tr>
<td>Ishigaki, 2015</td>
<td>Japan</td>
<td>NA</td>
<td>91</td>
<td>NA</td>
<td>NA</td>
<td>LDA</td>
<td>34</td>
<td>3</td>
<td>8</td>
<td>47</td>
<td>81%</td>
<td>94%</td>
<td>Vitro</td>
<td>785 nm</td>
<td>U</td>
</tr>
<tr>
<td>Wang, 2015</td>
<td>China</td>
<td>48</td>
<td>U</td>
<td>1172</td>
<td>U</td>
<td>PLS-DA/LOPCV</td>
<td>196</td>
<td>19</td>
<td>6</td>
<td>717</td>
<td>97%</td>
<td>97.4%</td>
<td>Vivo</td>
<td>785 nm</td>
<td>U</td>
</tr>
<tr>
<td>Bergholt, 2011</td>
<td>Singapore</td>
<td>27</td>
<td>75</td>
<td>75</td>
<td>65</td>
<td>U</td>
<td>32</td>
<td>2</td>
<td>1</td>
<td>40</td>
<td>97%</td>
<td>95.2%</td>
<td>Vivo</td>
<td>785 nm</td>
<td>96%</td>
</tr>
<tr>
<td>Jiang, 2007</td>
<td>China</td>
<td>64</td>
<td>128</td>
<td>128</td>
<td>56</td>
<td>U</td>
<td>61</td>
<td>0</td>
<td>3</td>
<td>64</td>
<td>95.31%</td>
<td>100%</td>
<td>Vitro</td>
<td>785 nm</td>
<td>U</td>
</tr>
<tr>
<td>Zhou, 2013</td>
<td>China</td>
<td>41</td>
<td>41</td>
<td>123</td>
<td>U</td>
<td>PCA</td>
<td>19</td>
<td>1</td>
<td>2</td>
<td>19</td>
<td>90.48%</td>
<td>95%</td>
<td>Vitro</td>
<td>785 nm</td>
<td>U</td>
</tr>
</tbody>
</table>

U, unknow; N1, number of patients; N2, number of samples; N3, number of spectra; PCA, principal component analysis; LDA, linear discriminate analysis; PLS-DA, partial least squares-discriminant analysis; LOPCV, leave-one patient-out, cross validation.
The pooled diagnostic efficacy of Raman spectroscopy in esophageal cancer. (A) The forest plot of pooled sensitivity and their 95% confidence intervals (CIs) of Raman spectroscopy to diagnose esophageal cancer of all the nine studies. (B) The forest plot of pooled specificity and their 95% confidence intervals (CIs) of Raman spectroscopy to diagnose esophageal cancer of all the nine studies. (C) The pooled PLR, NLR, DOR and SROC curve of Raman spectroscopy in diagnosis of esophageal cancer. PLR, positive likelihood ratios; NLR, negative likelihood ratios; DOR, diagnostic odds ratios; SROC, summary receiver operator characteristics.

The examination of esophageal specimens in vivo. The sensitivity of RS for esophageal specimens ranged from 0.81 (95% CI, 0.66–0.91) to 1.00 (0.94–1.00) and the pooled sensitivity was 0.94 (95% CI, 0.91–0.96). The specificity of RS for esophageal specimens ranged from 0.94 (95% CI, 0.83–0.99) to 1.00 (95% CI, 0.89–1.00) and the pooled specificity was 0.96 (95% CI, 0.94–0.98). The overall PLR and NLR were 20.16 (95% CI, 12.56–32.38) and 0.12 (95% CI, 0.06–0.22), respectively. The DOR was 186.77 (95% CI, 69.63–500.93). The SROC curve analysis was also used to evaluate the overall diagnostic accuracy. And the AUC was 0.9912. All of the plots of ex vivo group were shown in Figure 3.

In vivo group
Of all the included studies, four studied conducted an examination of esophageal specimens in vivo. The pooled
sensitivity was 0.94 (95% CI, 0.91–0.96) ranged between 0.87 (95% CI, 0.77–0.94) and 0.97 (95% CI, 0.94–0.99). The pooled specificity was 0.92 (95% CI, 0.90–0.93) ranged between 0.85 (95% CI, 0.82–0.87) and 0.97 (95% CI, 0.96–0.98). The overall PLR and NLR were 15.70 (95% CI, 3.90–63.23) and 0.07 (95% CI, 0.03–0.18), respectively. The DOR was 214.67 (95% CI, 33.71–1,732.40). The SROC curve analysis was also used to evaluate the general diagnostic accuracy. And the AUC was 0.9812. All of the plots of in vivo group were shown in Figure 4.

Quality assessment and publication bias

Standard quality evaluation of the 9 included studies were performed based on QUADAS-2 by Review Manager 5.3. Also, publication bias was evaluated through Deeks Funnel Plot Asymmetry Test (consider the existence of publication bias when P<0.05) which was conducted by Stata 14.2 (StataCorp, USA). The QUADAS-2 graphical display of the evaluation of the risk of bias was shown in Figure 5. The Deeks’ funnel plot asymmetry test was used to assess publication bias and it demonstrated that no significant publication bias was found (P=0.52), which was shown in Figure 6.

Discussion

This meta-analysis was performed on the basis of the standard protocol for a systematic review with nine articles were taken into account. Two independent reviewers were assigned in study screening, data extraction and quality assessment process. The SROC curve analysis was simultaneously applied. We expect to assess the value of Raman spectroscopy in diagnosing esophageal cancer and its value in clinical application.

This meta-analysis gave an evidence that RS had a good diagnostic accuracy in esophageal cancer, with the general pooled diagnostic sensitivity and specificity being 0.91 (95% CI, 0.89–0.93) and 0.92 (95% CI, 0.91–0.94). According to the general pooled data, we found that the overall sensitivity and specificity were over 90%, a high efficacy of early diagnosis of esophageal cancer was reconfirmed.
which suggested that RS had a low missed diagnosis rate in distinguishing between benign and malignant aspects of esophageal cancer. Moreover, we also found that the general pooled DOR was (95% CI, 68.32–690.53) by random effect model, indicating a high accuracy. In SROC analysis, AUC was 0.9779. Generally, the value of AUC is lying between 0.5 and 1.0, and the larger AUC indicates better performance. In the study, the AUC value was very close to 1, which suggested an excellent diagnostic efficiency. In addition, four of the nine included studies reported the accuracy of the RS in diagnosis of esophageal cancer range from 90% to 100%, also confirmed our conclusion.

In the subgroup analysis, the overall sensitivity and specificity of the diagnosis exceeded 90%, whether Raman spectroscopy analyze esophageal cancer tissues in vivo or in vitro both showed high diagnostic accuracy. Compared to intraoperative biopsy, the diagnosis of Raman spectroscopy might have the same diagnostic performance, and with an irreplaceable function that frozen biopsy cannot reach. Furthermore, in the subgroup analysis, we found that two studies did not use intraoperative masses to verify the benign and malignant esophageal cancer. They used hemoglobin and urinary modified nucleotides to diagnose esophageal cancer in vitro, so we didn’t put them into general analysis.

In the study of Zhou (24), by detecting the hemoglobin solution of 21 patients with esophageal cancer and 20 healthy people, they found that the RS of hemoglobin in patients with esophageal cancer and healthy people were significantly different. According to the analysis of characteristic peaks, they reported that the contents of tyrosine, phenylalanine and the number of vibration of pyrrole ring in hemoglobin of patients with esophageal cancer were slightly lower than those of healthy people. Meanwhile, compared with healthy people, esophageal cancer patients had the increased iron ions in low spin state in hemoglobin, while the iron ions in high spin state decreased, which indicated that the iron ions in high spin state in hemoglobin of esophageal cancer patients were transferred to low spin state, which was consistent with the phenomenon that the blood samples of cancer patients were more easily hemolyzed. Hence, they believe that Raman
Figure 5 The graphical display of the evaluation of the risk of bias and concerns regarding applicability of the selected studies. (A) Risk of bias and applicability concerns evaluation of included studies in pool. (B) Risk of bias and applicability concerns evaluation of included studies individually.
In our meta-analysis, we reconfirmed that RS had an excellent diagnostic efficiency in esophageal cancer. Simultaneously, we acknowledged several limitations in this study. Firstly, RS had not been widely admitted as a normal clinical diagnostic tool, therefore inadequate number of clinical researches were published, which absolutely lowered the number of articles we could include. Secondly, since most of the included researches were conducted in Asia, it might cause selection bias. Thirdly, due to limitations of current research, our study did not involve tumor subtypes. Further comprehensive study is needed so that it can target the subtypes in order to provide more precise clues for clinical practice.

**Conclusions**

Through our meta-analysis, we found a promisingly high sensitivity and specificity of RS in the diagnosis of suspected esophageal mass.

**Acknowledgments**

Funding: Supported by the Major Scientific and Technological Projects of the New Generation of Artificial Intelligence in Sichuan Province in 2018 under Grant 2018GZDZX0035.

**Footnote**

Peer Review File: Available at http://dx.doi.org/10.21037/tcr-20-854

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at http://dx.doi.org/10.21037/tcr-20-854

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-20-854). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are

**Figure 6** The Deeks’ funnel plot asymmetry test.
appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial republication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


