Systematic review

BJS Open Open Access

Liquid biopsy for cancer diagnosis using vibrational spectroscopy: systematic review

D. J. Anderson[1,3], R. G. Anderson[3], S. J. Moug[3] and M. J. Baker[1,2]

[1]WestCHEM, Department of Pure and Applied Chemistry, and [2]ClinSpec Diagnostics, Technology and Innovation Centre, University of Strathclyde, Glasgow, and [3]Clinical Research Unit, Royal Alexandra Hospital, NHS Greater Glasgow and Clyde, Paisley, UK

Correspondence to: Mr D. J. Anderson, Department of Pure and Applied Chemistry (Level 6), Technology and Innovation Centre, University of Strathclyde, 99 George Street, Glasgow G1 1RD, UK (e-mail: david.anderson13@nhs.net)

Background: **Vibrational spectroscopy (VS) is a minimally invasive tool for analysing biological material to detect disease. This study aimed to review its application to human blood for cancer diagnosis.**

Methods: **A systematic review was undertaken using a keyword electronic database search (MEDLINE, Embase, PubMed, TRIP and Cochrane Library), with all original English-language manuscripts examining the use of vibrational spectral analysis of human blood for cancer detection. Studies involving fewer than 75 patients in the cancer or control group, animal studies, or where the primary analyte was not blood were excluded.**

Results: **From 1446 results, six studies (published in 2010–2018) examining brain, bladder, oral, breast, oesophageal and hepatic cancer met the criteria for inclusion, with a total population of 2392 (1316 cancer, 1076 control; 1476 men, 916 women). For cancer detection, reported mean sensitivities in each included study ranged from 79·3 to 98 per cent, with specificities of 82·8–95 per cent and accuracies between 81·1 and 97·1 per cent. Heterogeneity in reporting strategies, methods and outcome measures made meta-analysis inappropriate.**

Conclusion: **VS shows high potential for cancer diagnosis, but until there is agreement on uniform standard reporting methods and studies with adequate sample size for valid classification models have been performed, its value in clinical practice will remain uncertain.**

Funding information
No funding

Paper accepted 23 March 2020
Published online 19 May 2020 in Wiley Online Library (www.bjsopen.com). **DOI:** 10.1002/bjs5.50289

Introduction

The detection and treatment of cancers at earlier stages will improve overall patient survival[1–3]. Optimal treatment at an early stage may contribute to faster recovery and better health-related quality of life.

Current cancer diagnostic modalities are generally invasive, requiring histopathological analysis of biopsied tissue. This can result in complications[4], requires significant sample processing[5], is time consuming[6] and is subject to inter-operator variability in interpretation[7,8].

An alternative is blood analysis for cancer detection. This 'liquid biopsy' has been reported widely as a means of achieving early diagnosis[9,10]. The technique usually focuses on the detection of a single element such as circulating tumour DNA, microRNA, or the presence of circulating tumour cells within the sample[11], rather than the phenotypic response of the body as a whole. It is expensive (estimated $1750 (€1610, exchange rate 8 April 2020) per patient) and requires skill acquisition and specialized equipment[12]. Other techniques are also evolving, including spectroscopy liquid biopsy. Rather than focusing on detection of a single element reflecting the presence of cancer, vibrational spectroscopy (VS) evaluates the entire host phenotypic response, thereby detecting the presence of cancer cells and metabolites (including tDNA/RNA, base pairs and cancer proteins), along with tissue changes (including protein/lipids) and the host immune response. It therefore encompasses a multimarker approach to cancer diagnostics. Proof-of-concept studies[13–18] have demonstrated that it is minimally invasive, valid, reproducible, easily learned and accurate for cancer diagnosis.

Spectroscopy can be defined as the 'study of the interaction between electromagnetic waves and matter'[19]. This

interaction can be used to determine the unique molecular composition or 'chemical fingerprint'[20] of the sample under analysis. VS typically refers to the techniques of infrared (IR) and Raman spectroscopy. Both techniques can be used on tiny volumes of blood[13,14], require little sample preparation, and are simple to perform[20]. VS provides a unique 'biological fingerprint' of the entire sample under analysis, and detectable variations within that fingerprint can be used to diagnose different disease pathologies[15,20]. It can be employed as a label-free, non-destructive and non-invasive approach to specimen analysis, including the analysis of fluids and tissues, allowing identification of specific 'spectral biomarkers'[16] of disease, including cancers[17,18].

Despite growth in the literature and increased public awareness, no systematic review has been done. This systematic review was undertaken to evaluate the use of VS in the analysis of human blood for cancer diagnosis, paying specific attention to confidence in and effectiveness of the technique, along with an assessment of the evidence for its use in clinical practice.

Methods

This review was performed in accordance with the standards described by the PRISMA statement[21]. The review was registered with the International Prospective Register of Systematic Reviews (PROSPERO) (registration number CRD42018115187), where the protocol and search strategy are available[22,23].

Table 1 Search methods

Method	Detail
Electronic database search	MEDLINE
	Embase
	PubMed
	TRIP (Turning Research Into Practice)
	Cochrane Library
	Web of Science
Other methods for identifying relevant research	Augmentation by review of reference list or bibliography of identified studies
Journals hand-searched	Not carried out for any specific journal

Table 2 Criteria for inclusion of studies in the systematic review

Item	Criteria
Population, or participants and conditions of interest	Human population with histopathologically confirmed cancer diagnosis
Intervention/exposure/ investigation	Application of vibrational spectroscopy to the analysis of human blood with the specific aim of cancer diagnosis
Comparisons/control groups	Non-cancer population as control group
Outcomes of interest	Descriptors of diagnostic potential (sensitivity, specificity, accuracy, positive predictive value, negative predictive value, κ values, receiver operating characteristic (ROC) curve)
Setting	Laboratory analysis of human blood with vibrational spectroscopy
Study design	Any study design fitting the above criteria

Search strategy

An online electronic database search was undertaken using the platforms of MEDLINE, Embase, PubMed, TRIP (Turning Research Into Practice) and Cochrane Library. The date of last search was 2 October 2018. Any publication to this date was considered for inclusion. The search strategy outlined below was employed in the MEDLINE online electronic database, and adapted for use with other databases as needed for their search system. Databases were searched from inception. Search terms were identified by: reviewing Medical Subject Headings (MeSH) for relevant and appropriate terms; and extracting key terminology/keywords from reviews and a sample of potentially relevant primary data studies. A test set of potentially relevant studies was used to ensure the search terms retrieved 100 per cent of the test set. A summary of the search strategy is shown in *Table 1*. The results of the literature search were downloaded into EndNote™ X8 software (Clarivate Analytics, London, UK). Exact article duplicates were removed using the appropriate tool in Endnote™ X8 software.

Eligibility criteria

A summary of eligibility criteria for this review, following the PICOS framework (Population, Intervention, Comparison, Outcome (PICO) process with added qualitative search terms), is detailed in *Table 2*. Exclusion criteria are shown in *Table 3*.

Only studies with both a cancer and control population group, each containing more than 75 participants, were included, based on a literature review[24] indicating that to generate a diagnostic model for sample classification more than this number of participants should be recruited to each group. Studies of non-blood, non-human, animals or pooled cells and with no controls were excluded, as were review articles, opinion papers and commentaries. Studies for identification of existing tumour markers were

Table 3 Criteria for exclusion of studies from the systematic review

Study with fewer than 75 participants in either arm (cancer and control)*
Not an original paper examining the use of VS in the analysis of human blood for cancer detection/diagnosis (review article, opinion or commentary)
Method other than VS used as primary method of analysis
Non-human study
Analysis primarily of non-blood-based analyte (tissue, cell)
English language restriction will apply

*Existing literature[24] identifies that, for an effective model to be constructed for classification and diagnosis, a data set of 75 participants or more is required, justifying the lower limit for participant inclusion in the present systematic review. VS, vibrational spectroscopy.

excluded. A review of grey literature/unpublished work was not conducted.

Studies of humans with *ex vivo* vibrational spectral analysis of blood and components (serum/plasma) for the detection of cancer were included. There were no demographic restrictions. Blood samples must have undergone laboratory analysis by VS for inclusion. No discrimination was made for subtypes of VS or for the constituent of blood analysed.

Publications had to report an indicator of the diagnostic capability of VS (sensitivity, specificity, area under the receiver operating characteristic (ROC) curve (AUC), accuracy). Current standard diagnostic testing for the cancer under investigation (histopathology) must have been used as the reference standard. Only peer-reviewed articles were considered. An English-language restriction was applied. No date range restrictions were used.

Data processing and interpretation

Many of the studies used chemometric or multivariable analysis techniques, the details of which are beyond the scope of this review, which simply states the techniques used in each study. For those interested, the authors would direct readers to the review by Biancolillo and Marini[25], which summarizes the main techniques used in the included studies.

Screening for eligibility/inclusion

Articles identified from the literature search were reviewed independently at the title and abstract level by two members of the review team. Disagreements were resolved by discussion, and arbitration with a third independent reviewer if required. A copy of articles meeting the inclusion criteria based on the title and abstract review was obtained for full-text review, unless the article was not available after an attempt had been made to obtain it. Copies of articles that were assessed indeterminately for relevance on title and abstract review were also obtained to determine eligibility based on full-text review. Studies were no longer considered once the title and abstract review clearly indicated that they did not fit the inclusion criteria. Full-text review was performed independently by two members of the review team. Valid studies were assessed for quality before any retrieval of data was performed.

Quality assessment of eligible studies

Two reviewers independently assessed each eligible study. All eligible studies were assessed using the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool (University of Bristol)[26]. Any areas of conflict between the two reviewers were resolved with arbitration with a third reviewer if required.

Outcomes and data extraction

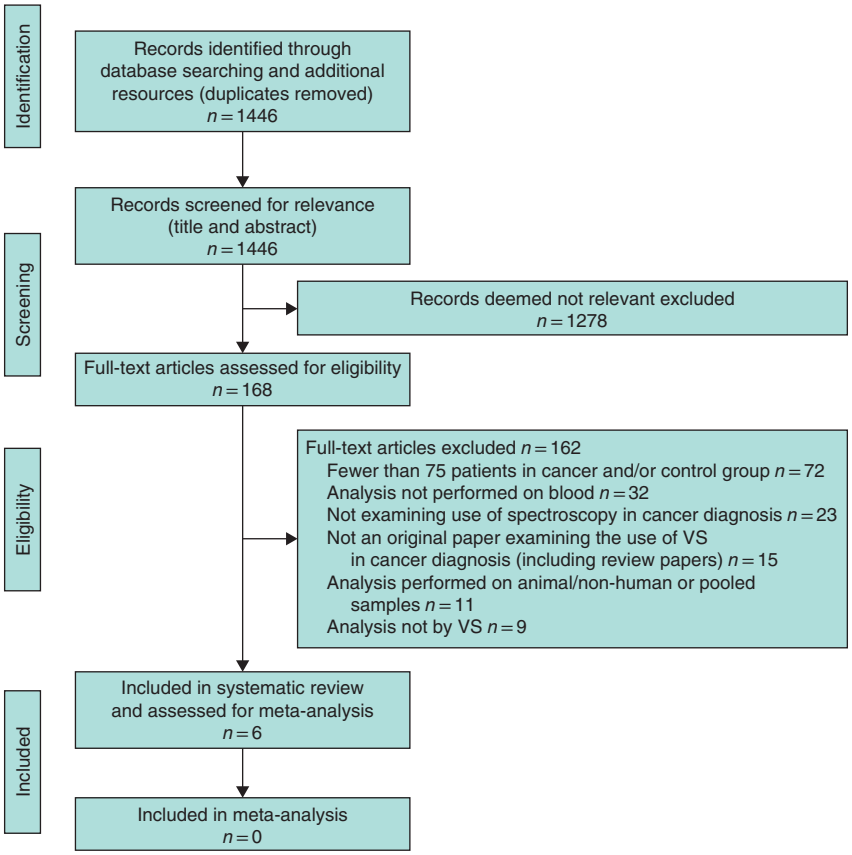
Included outcomes related to ability of VS to diagnose cancer by analysis of human blood including: true and false rates of positivity and negativity, sensitivity, specificity, accuracy, positive and negative predictive values, and AUC values.

Data were extracted from each article and stored in Excel® (Microsoft, Redmond, Washington, USA) format. Data extraction was cross-checked independently. Where multiple spectral techniques or data analysis techniques were evaluated within a study, data are described based on the most effective technique used. Where data were presented for both a training set and test/cross-validation set, data from the test set are presented as these reflect most closely the performance of the test in clinical practice. An initial intention to report a meta-analysis of the data from the included studies was not undertaken owing to a variety of factors that indicated that this exercise was not scientifically valid. A detailed explanation appears in the discussion.

Results

The literature search produced 1446 records of which 1278 were excluded on review of title and abstract, and a further 162 on full-text review. After full-text review, six studies[6,27–31], published between 2010 and 2018, were selected for inclusion (*Fig. 1*). A summary of the included studies is given in *Table 4*. *Fig. 2* presents the QUADAS quality assessment for the included studies.

Fig. 1 PRISMA diagram for the selection of studies



VS, vibrational spectroscopy.

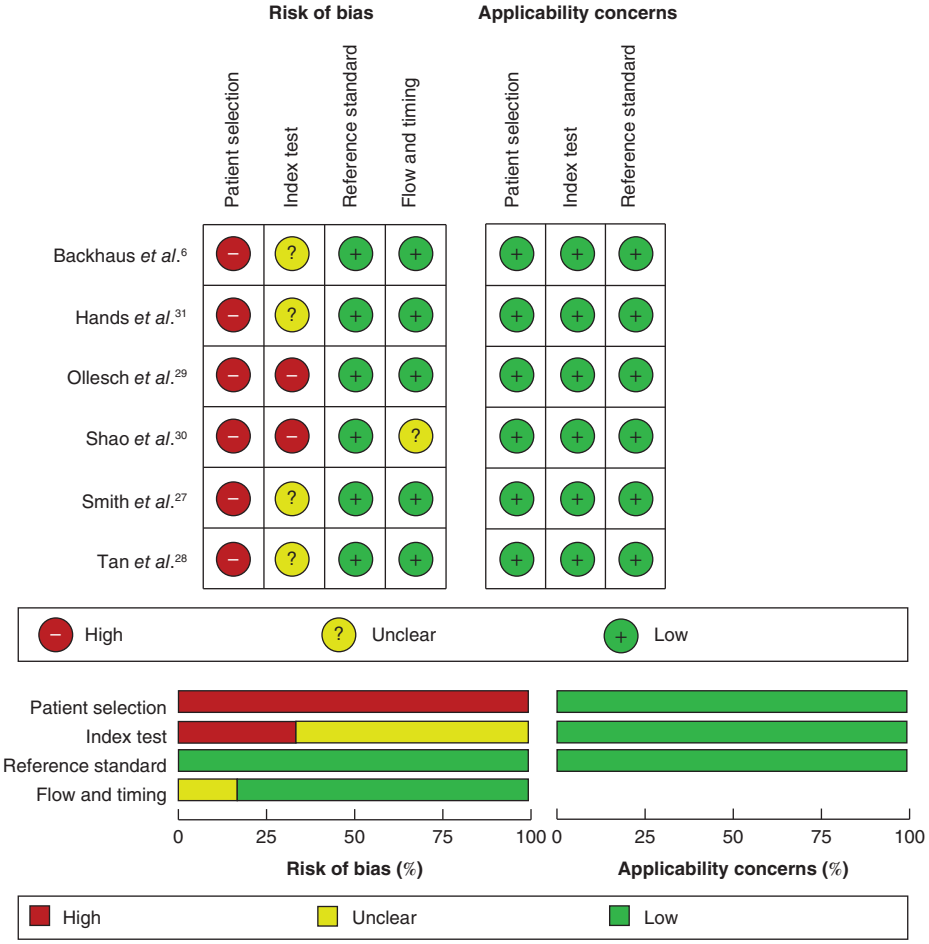
Table 4 Summary of included study characteristics and reported outcome measures

Reference	No. of patients	Men		Women		Age (years)*	Sensitivity (%)*	Specificity (%)*	Accuracy (%)
		Cancer	Control	Cancer	Control				
Hands *et al.*[31]	433†	133	64	178	58	51·27	91·5	83·0	n.s.
Smith *et al.*[27]	433†	133	64	178	58	51·27	92·8	91·5	92·4
Tan *et al.*[28]	370	125	93	100	52	n.s.	79·3	82·8	81·1
Ollesch *et al.*[29]	652	305	198	83	66	71	93	92	92
Shao *et al.*[30]	744	245	313	50	136	53·2	n.s.	n.s.	97·1
Backhaus *et al.*[6]	193	0	0	97	96	55	98	95	n.s.

*Values are means. †Studies performed from the same patient population but with different results owing to different methods of data analysis. n.s., Not stated.

Five studies[6,27–29,31] reported sensitivity and specificity of spectroscopy against a confirmed histopathological result, whereas the other study[30] provided only overall diagnostic accuracy. Three studies[6,27,31] used IR spectroscopy of human serum, one study[29] used IR spectroscopy of human serum and plasma, and two[28,30] used Raman spectroscopy of human serum. The total number of participants was 2392, with 1316 cancer (311 brain, 225 oral, 388 bladder, 196 liver, 99 oesophageal and 97 breast), and 1076 non-cancer participants. Patient location was stated as China in two papers[28,30], as Germany in one[6], and as the UK in two[27,31]. The total population

Fig. 2 Quality Assessment of Diagnostic Accuracy Studies (QUADAS) analysis of the six included studies



comprised 1476 men and 916 women. Mean age was 57·6 years. Breakdown of subgroups by age and sex within the studies was generally not available.

All studies examining by IR spectroscopy were conducted in the mid-IR range. Raman spectra were also all collected in the mid-IR range. All studies were performed on dried blood samples. There was variance between the studies regarding fasting status at sampling, with one study[30] reporting fasted samples, two studies[27,31] reporting unfasted samples, and the other three studies[6,28,29] not reporting on the fasting status of participants. Volume of analysed blood was reported in two studies[27,31] (1 µl), but not reported in the others, and sampled volume to the blood collection tube was also reported by only two studies: Tan *et al.*[28] (5 ml) and Shao *et al.*[30] (3 ml). Storage temperature for samples was also varied, with four studies[27,29–31] storing at −80°C and two[6,28] reporting storage at only −20°C. Methods for blood fractionation were generally slightly different between the studies, but there are many accepted laboratory protocols that can fractionate blood effectively without issue.

There was wide variation between the included studies with regard to data analysis techniques. Two studies[27,29] performed random forest analysis, two[28,29] used linear discriminant analysis, and the other analysis methods were each performed by only one study: support vector machines[31]; partial least squares[30]; principal component analysis[28]; cluster analysis and artificial neural networks[6]. Of note, no study reported the blood storage system used to collect the blood samples from participants.

Demographic details of each included study are shown in *Table 4*, and a detailed summary of each study is presented below.

Hands and colleagues[31] presented a study using IR spectroscopy (attenuated total reflection Fourier-transform infrared (ATR-FTIR) with a diamond crystal plate) in the analysis of blood serum samples examining the discrimination of glioblastoma multiforme and metastatic brain cancer among UK patients. A total of 3897 spectra were collected. Serum separation was by centrifugation at 1200g for 10 min, followed by storage at −80°C before analysis. Fasting status of the participants at time of sampling was not documented, but on contacting the author it was possible to confirm that samples were not fasted. Analysis was performed on 1 μl serum. Preprocessing was with vector normalization using the Savitsky–Golay method[32]. Analysis was through a feature-fed support vector machine (SVM) technique; *n*-fold ($n = 5$) cross-validation was performed in Matlab™ (https://www.mathworks.com/products/matlab.htm) and with in-house written protocols. Best results were achieved with feature-fed SVM with 130 spectral features, reporting optimal sensitivities and specificities (defined as those that best describe the sample set based on disease grouping) respectively of 98·1 (range 82–98·1) and 97·6 (66–97·6) per cent in discrimination of cancer *versus* non-cancer. Mean sensitivities and specificities were 91·5 and 83·0 per cent respectively. Accuracy, positive and negative predictive values, and ROC curves were not described.

Smith and co-workers[27] performed a study using the same participants and spectral technique as described above for Hands *et al*.[31], evaluating discrimination of glioblastoma multiforme from healthy samples. Preprocessing was also similar, using vector normalization and the Savitsky–Golay method. A different approach to data processing and analysis was used with random forest machine-learning mechanisms (RF-MLM) and two-dimensional correlation analysis performed in Matlab™ and with in-house written protocols. With this approach, they achieved best results with RF-MLM for normalized second derivative data, achieving a sensitivity of 92·8 (range 82·5–92·8) per cent and a specificity of 91·5 (76·6–91·5) per cent in cancer *versus* non-cancer discrimination, and an overall accuracy of 92·4 (81·3–92·4) per cent. Positive and negative predictive values were 93·7 (91·6–93·7) and 81·2 (51·1–82·1) per cent respectively. Other descriptors of diagnostic accuracy were not described.

Tan *et al*.[28] performed a study using surface-enhanced Raman spectroscopy (SERS), evaluating serum for discrimination of oral squamous cell carcinoma, comparing against both a disease-free control and a positive control group (with mucoepidermoid carcinoma), recruiting patients from the First Affiliated Hospital of Xiamen University, China. A total of 370 spectra were collected. Some 5 ml of whole blood was sampled from each patient. Participants were fasted at the time of sample collection. Blood component separation was by centrifugation at 3400 r.p.m. for 10 min. Samples were then stored at −20°C. Preprocessing was with vector normalization and the Savitsky–Golay method. With data processing and analysis by principal component analysis and linear discriminant analysis, performed in LabSpec™ 2.0 (HORIBA Scientific, Stanmore, UK), they achieved a sensitivity and specificity of 79·3 and 82·8 per cent respectively for discrimination of cancer against control (no ranges described). The overall accuracy was 81·1 per cent. Other descriptors of diagnostic accuracy were not given.

Ollesch and colleagues[29] carried out a study using ATR-FTIR spectroscopy with a diamond crystal plate, examining the discrimination of urinary bladder cancer *versus* control (urocystitis or urethral infection). There was no description of the total number of spectra collected, the site of enrolment, or patient fasting status. Plasma was collected from both EDTA and sodium citrate tubes but collection of serum was not described. Samples were stored at −80°C. Participants were divided randomly into four different balanced or unbalanced groups, which created a level of complexity within the interpretation of the study itself (these subsets were attempting to analyse the ability to discriminate between healthy control, disease control and cancer samples). Data processing and analysis was by random forest linear discriminant analysis and minimum redundancy maximum relevance feature selection. The authors achieved a sensitivity and specificity of 93 and 92 per cent respectively. Overall accuracy was 92 per cent. Other descriptors of diagnostic accuracy were not given.

Shao and co-workers[30] also used SERS on human serum in an examination of liver cancer *versus* a positive control (oesophageal cancer) and non-cancer controls (other hepatic disease), with participants recruited from Beijing YouAn Hospital, China. All participants were fasted at the time of sampling 3 ml whole blood. Blood component separation was with centrifugation at 3000 r.p.m. for 10 min. Samples were stored at −80°C. Preprocessing was with baseline subtraction, smoothing and area normalization, with data processing and analysis using orthogonal partial least squares–discriminant analysis and *n*-fold ($n = 10$) cross-validation, as well as ROC curves for validation. Individual sensitivities and specificities were not described, but overall diagnostic accuracy for hepatocellular carcinoma (HCC) reached 97·1 per cent. The AUC value was 0·998. For oesophageal cancer, overall diagnostic accuracy reached 95·3 per cent, with an AUC value of 0·977. Positive and negative predictive values were not described. Unlike

other included studies, sensitivity and specificity values for different HCC tumour grades were not described. AUC values for HCC grades 0, 1, 2 and 3 were 0·996, 0·998, 0·989 and 0·992 respectively. Unfortunately, the classification system used to determine tumour grade was not stated. No further analysis based on tumour grade was included.

Backhaus *et al*.[6] used IR spectroscopy of human serum in an examination of breast cancer. Participants were recruited from St Vencentius Krankenhaus, Karlsruhe, Germany, and a total of 579 spectra were collected. Fasting status and the volume of blood collected at the time of sampling were not described. Samples were stored at −20°C. Preprocessing was with vector normalization and the Savitsky–Golay method. Data processing and analysis was by cluster analysis and artificial neural networks, achieving a sensitivity of 98 per cent and a specificity of 95 per cent in the discrimination of cancer *versus* non-cancer. Other descriptors of diagnostic accuracy were not given.

Discussion

VS is a mature technology, yet its use in medical diagnostics remains in its infancy. In this review, reported diagnostic outcome measures from all included studies were high, with four of five studies reporting sensitivities at or above 90 per cent, and the other 79·3 per cent (range for all studies 79·3–98 per cent). In addition, all five studies[6,27–29,31] reported high specificities of between 83·0 and 95 per cent. Where reported, other diagnostic measures (accuracy, predictive values) were also consistently high.

The levels of diagnostic test accuracy described compare favourably with current reference standards. Examples include the current blood-based marker for prostate cancer, prostate-specific antigen, which has a sensitivity for prostate cancer of 20·5 per cent and a specificity of 93·6 per cent[33]. The test used in the current UK Breast Screening Programme has a sensitivity of 81–91 per cent[34]. Breast core biopsy for histopathological examination has reported sensitivity values of 90·1–93 per cent[35], and is more operator-dependent than the techniques described in this review. Oesophagogastroduodenoscopy has a sensitivity in upper gastrointestinal cancer screening of 88 per cent and a specificity of 85 per cent[36]. The equivalent lower gastrointestinal investigation of colonoscopy has a sensitivity of 92·5 per cent[37], and the newly instigated bowel screening tool, the quantitative faecal immunochemical test (qFIT) has a reported sensitivity of 79 per cent[38]. MRI of brain cancers has been shown to have a sensitivity of 72 per cent and a specificity of 65 per cent for evaluation of high-grade glioma[39]. Overall, the reported diagnostic measures for non-invasive VS compare favourably with existing diagnostic modalities.

Given the high levels of diagnostic accuracy achieved with VS, it has a potential role as a minimally invasive tool for cancer investigation, although the present review has identified several limitations indicating that significant further work is required to allow progression towards clinical use.

As demonstrated by the small number of studies suitable for inclusion, the majority of the existing literature involves small studies (significantly fewer than 75 participants), using widely differing methodologies and methods for the reporting of analytical outcomes, all of which make direct comparison between studies difficult. Publication of robust larger studies is paramount to evaluate the technique further.

Reporting methods in the included studies were inconsistent and at times confusing, with multiple reported values for outcome measures described over several pages or tables. Comparison between results was difficult. Only one study[6] provided complete data for true positive, true negative, false positive and false negative results used to determine sensitivity and specificity, and only one[30] provided ROC curves. With regard to predictive values, again only one study[27] reported these. Most importantly, none of the studies provided complete data for all of the key areas in analysis of diagnostic accuracy tests: true and false positivity and negativity, sensitivity, specificity, and positive and negative predictive values.

For meta-analysis, bivariable and hierarchical summary receiver operating characteristic (HSROC) models[40] were considered to deal with all of this variability in the analysis[41,42]. On performing initial meta-analysis, it was noted that the score for heterogeneity was low, with an I^2 value of 23·61 per cent, reflecting consistently high levels of sensitivity and specificity in individual studies. On review of the full texts, however, it was clear that there was considerable heterogeneity between the studies, as outlined above regarding methods, populations and statistical analyses. Quality assessment of the included papers clearly demonstrated a lack of standardization across a variety of reporting measures. As a result, it was considered that any results of a meta-analysis would carry low confidence in terms of validity, with the potential to be misleading to the reader.

All of the included studies were of case–control design with reasonably sized patient cohorts, although it is accepted that this design is prone to bias and can produce inflated estimates of diagnostic test accuracy. The use of the term ‘optimal’ sensitivity in some studies suggests the

possibility that authors may have been cherry-picking the best results and introducing reporting bias. Much of the literature on VS, however, relates to small comparative or proof-of-concept studies with no control groups. Few studies were therefore identified as suitable for inclusion in the present review, so the pool from which conclusions have been drawn is small.

One important methodological issue merits specific attention. None of the included studies reported on variations in blood collection tubes and methods used for blood sampling. To date, no analysis of these variables has been published. It has been shown[43] that storage at low temperatures (−80°C) and storage in plastic tubes has no effect on generated spectra, whereas variations in sample drying can have an effect.

Although promising, considerable work is required before the adoption of VS into routine clinical use can be considered. Efforts are being made to standardize VS as a diagnostic test for human disease[15,44,45]. In addition, reporting of sensitivity, specificity and accuracy along with ROC curves in the presentation of future work would seem to be at least a minimum requirement.

Acknowledgements

The authors thank S. Barry (Chancellor's Fellow and Lecturer, Department of Mathematics and Statistics, Strathclyde University) for her input regarding bivariable and HSROC analysis, and for reviewing the statistical methodologies of the included studies.

Disclosure: The authors declare no conflict of interest.

References

1 Cancer Research UK. *Bowel Cancer Survival Statistics*. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#heading-Two [accessed 26 February 2018].
2 Cancer Research UK. *Breast Cancer Survival Statistics*. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Two [accessed 26 February 2018].
3 Cancer Research UK. *Prostate Cancer Survival Statistics*. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/survival#heading-Two [accessed 26 February 2018].
4 Nazir B. Pain during transrectal ultrasound-guided prostate biopsy and the role of periprostatic nerve block: what radiologists should know. *Korean J Radiol* 2014; **15**: 543–553.
5 Kumar R, Srivastava R, Srivastava S. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. *J Med Eng* 2015; **2015**: 457906.
6 Backhaus J, Mueller R, Formanski N, Szlama N, Meerpohl H-G, Eidt M *et al.* Diagnosis of breast cancer with infrared spectroscopy from serum samples. *Vib Spectrosc* 2010; **52**: 173–177.
7 Kendall C, Stone N, Shepherd N, Geboes K, Warren B, Bennett R *et al.* Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. *J Pathol* 2003; **200**: 602–609.
8 Stamey TA, Caldwell M, McNeal JE, Nolley R, Hemenez M, Downs J. The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? *J Urol* 2004; **172**: 1297–1301.
9 McArdle H. Lung cancer blood test can detect disease before symptoms show. *The Herald*, 9 September 2019.
10 Park A. Promising blood test could help to predict breast cancer recurrence. *TIME Magazine*, 9 September 2019.
11 Lianidou E, Pantel K. Liquid biopsies. *Genes Chromosom Cancer* 2019; **58**: 219–232.
12 Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R *et al.* Phylogenetic ctDNA analysis depicts early stage lung cancer evolution. *Nature* 2017; **545**: 446–451.
13 Kong K, Kendall C, Stone N, Notingher I. Raman spectroscopy for medical diagnostics – from *in-vitro* biofluid assays to *in-vivo* cancer detection. *Adv Drug Del Rev* 2015; **89**: 121–134.
14 Lacombe C, Untereiner V, Gobinet C, Zater M, Sockalingum GD, Garnotel R. Rapid screening of classic galactosemia patients: a proof-of-concept study using high-throughput FTIR analysis of plasma. *Analyst* 2015; **140**: 2280–2286.
15 Byrne HJ, Baranska M, Puppels GJ, Stone N, Wood B, Gough KM *et al.* Spectropathology for the next generation: *quo vadis*? *Analyst* 2015; **140**: 2066–2073.
16 Mitchell AL, Gajjar KB, Theophilou G, Martin FL, Martin-Hirsch PL. Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting. *J Biophoton* 2014; **7**: 153–165.
17 Elmi F, Movaghar AF, Elmi MM, Alinezhad H, Nikbakhsh N. Application of FT-IR spectroscopy on breast cancer serum analysis. *Spectrochim Acta A* 2017; **187**: 87–91.
18 Li Q, Hao C, Kang X, Zhang J, Sun X, Wang W *et al.* Colorectal cancer and colitis diagnosis using Fourier transform infrared spectroscopy and an improved K-nearest-neighbour classifier. *Sensors (Basel)* 2017; **17**: E2739.
19 Banwell CN, McCash EM. *Fundamentals of Molecular Spectroscopy* (3rd edn). McGraw-Hill**:** London, 1983.
20 Baker MJ, Hussain SR, Lovergne L, Untereiner V, Hughes C, Lukaszewski RA *et al.* Developing and understanding biofluid vibrational spectroscopy: a critical review. *Chem Soc Rev* 2016; **45**: 1803–1818.
21 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.

22 Centre for Reviews and Dissemination, University of York. Protocol for Systematic Review: 'A systematic review of vibrational spectroscopy in the analysis of human blood for cancer diagnosis', as registered with PROSPERO. https://www.crd.york.ac.uk/PROSPEROFILES/115187_PROTOCOL_20181112.pdf [accessed 5 October 2018].
23 Centre for Reviews and Dissemination, University of York. Search Strategy for Systematic Review: 'A systematic review of vibrational spectroscopy in the analysis of human blood for cancer diagnosis', as registered with PROSPERO. https://www.crd.york.ac.uk/PROSPEROFILES/115187_STRATEGY_20181112.pdf [accessed 5 October 2018].
24 Beleites C, Neugebauer U, Bocklitz T, Krafft C, Popp J. Sample size planning for classification models. *Anal Chim Acta* 2013; **760**: 25–33.
25 Biancolillo A, Marini F. Chemometric methods for spectroscopy-based pharmaceutical analysis. *Front Chem* 2018; **6**: 576.
26 Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB *et al.*; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–536.
27 Smith BR, Ashton KM, Brodbelt A, Dawson T, Jenkinson MD, Hunt NT *et al.* Combining random forest and 2D correlation analysis to identify serum spectral signatures for neuro-oncology. *Analyst* 2016; **141**: 3668–3678.
28 Tan Y, Yan B, Xue L, Li Y, Luo X, Ji P. Surface-enhanced Raman spectroscopy of blood serum based on gold nanoparticles for the diagnosis of the oral squamous cell carcinoma. *Lip Health Dis* 2017; **16**: 73.
29 Ollesch J, Heinze M, Heise HM, Behrens T, Bruning T, Gerwert K. It's in your blood: spectral biomarker candidates for urinary bladder cancer from automated FTIR spectroscopy. *J Biophotonics* 2014; **7**: 210–221.
30 Shao L, Zhang A, Rong Z, Wang C, Jia X, Zhang K *et al.* Fast and non-invasive serum detection technology based on surface-enhanced Raman spectroscopy and multivariate statistical analysis for liver disease. *Nanomedicine* 2018; **14**: 451–459.
31 Hands JR, Clemens G, Stables R, Ashton K, Brodbelt A, Davis C *et al.* Brain tumour differentiation: rapid stratified serum diagnostics via attenuated total reflection Fourier-transform infrared spectroscopy. *J Neurooncol* 2016; **127**: 463–472.
32 Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 1964; **36**: 1627–1639.
33 Ankerst DP, Thompson IM. Sensitivity and specificity of prostate-specific antigen for prostate cancer detection with high rates of biopsy verification. *Arch Ital Urol Androl* 2006; **78**: 125–129.
34 Hofvind S, Geller BM, Skelly J, Vacek PM. Sensitivity and specificity of mammographic screening as practised in Vermont and Norway. *Br J Radiol* 2012; **85**: e1226–e1232.
35 Santos IP, Barroso EM, Bakker Schut TC, Caspers PJ, van Lanschot CGF, Choi DH *et al.* Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics. *Analyst* 2017; **142**: 3025–3047.
36 Ro TH, Mathew MA, Misra S. Value of screening endoscopy in evaluation of esophageal, gastric and colon cancers. *World J Gastroenterol* 2015; **21**: 9693–9706.
37 Martín-López JE, Beltrán-Calvo C, Rodríguez-López R, Molina-López T. Comparison of the accuracy of CT colonography and colonoscopy in the diagnosis of colorectal cancer. *Colorectal Dis* 2014; **16**: O82–O89.
38 Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. *Ann Intern Med* 2014; **160**: 171.
39 Law M, Yang S, Wang H, Babb JS, Johnson G, Cha S *et al.* Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *Am J Neuroradiol* 2003; **24**: 1989.
40 Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers – part II. Statistical methods of meta-analysis. *Korean J Radiol* 2015; **16**: 1188–1196.
41 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, 2018.
42 The Cochrane Collaboration. *Review Manager (RevMan)*. The Nordic Cochrane Centre: Copenhagen, 2014.
43 Lovergne L, Lovergne J, Bouzy P, Untereiner V, Offroy M, Garnotel R *et al.* Investigating pre-analytical requirements for serum and plasma based infrared spectro-diagnostic. *J Biophotonics* 2019; **12**: e201900177.
44 Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM *et al.* Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* 2014; **9**: 1771–1791.
45 Lovergne L, Bouzy P, Untereiner V, Garnotel R, Baker MJ, Thiéfin G *et al.* Biofluid infrared spectro-diagnostics: pre-analytical considerations for clinical applications. *Faraday Discuss* 2016; **187**: 521–537.