Accuracy of Ex-vivo Fluorescence Confocal Microscopy in Margin Assessment of Solid Tumors: A Systematic Review


Department of Targeted Intervention, University College London, London, United Kingdom (MA, RA-M, GS), and Department of Urology (TA-H, GS) and Department of Pathology (AH), University College London Hospitals, London, United Kingdom

Summary
Fluorescence confocal microscopy (FCM) is a novel technology that enables rapid high-resolution digital imaging of non-formalin-fixed tissue specimens and offers real-time positive surgical margin identification. In this systematic review, we evaluated the accuracy metrics of ex vivo FCM for intraoperative margin assessment of different tumor types. A systematic search of MEDLINE via PubMed, Embase, Cochrane Central Register of Controlled Trials, Web of Science, and Scopus was performed for relevant papers (PROSPERO ID: CRD42022372558). We included 14 studies evaluating four types of microscopes in six different tumor types, including breast, prostate, central nervous system, kidney, bladder, and conjunctival tumors. Using the Quality Assessment of Diagnostic Accuracy Studies tool, we identified a high risk of bias in patient selection (21%) and index test (36%) of the included studies. Overall, we found that FCM has good accuracy metrics in all tumor types, with high sensitivity and specificity (>80%) and almost perfect concordance (>90%) against final pathology results. Despite these promising findings, the quality of the available evidence and bias concerns highlight the need for adequately designed studies to further define the role of ex vivo FCM in replacing the frozen section as the tool of choice for intraoperative margin assessment: (J Histochem Cytochem XX, XXX–XXX, XXXX)

Keywords
brain neoplasms, breast neoplasms, confocal microscopy, margins of excision, prostatic neoplasms

Introduction
A positive surgical margin (PSM) denotes the presence of residual cancer cells at the margin of resected specimens, and has been shown to increase the risk of local recurrence and decrease overall survival in several cancer types.1–4 Numerous studies have further stratified recurrence risk levels based on margin status, with the recommendation of additional therapies if a PSM is present.5–8 Accurate identification of PSMs during surgery is, therefore, crucial to guide secondary resection and avoid the need for reoperations or adjuvant treatment.

Frozen section (FS) is currently the standard method for intraoperative diagnosis and margin assessment in surgical oncology. However, it has limited sensitivity because only sections that are most likely to be
representative are analyzed. If the entire surface of the specimen was analyzed using FS, the procedure would be too labor-, staff-, and time-intensive. In addition, incomplete cutting, tissue folding, distorted architecture, and freezing artifacts may occur. If the slides are not interpretable, resampling may be requested, further delaying the results. These drawbacks have restricted the adoption of routine FS in many centers. Moreover, analyzing small samples may result in tissue loss, compromising the accuracy of the formalin-fixed paraffin-embedded analysis (FFPE).

Fluorescence confocal microscopy (FCM), also known as confocal laser microscopy, is a novel imaging technique that produces high-resolution images of fresh specimens with cellular-level detail using photo-reactive dyes. Compared with FS, FCM has the advantage of generating digital scans in minutes, requiring fewer consumables and minimal tissue processing. In addition, FCM can be placed within the operating room, avoiding the need to transport samples across sites. Although FCM is now routinely used in dermatology, interest in using FCM to guide surgical resection has increased in recent years. However, we realized a systematic review on this topic is lacking. The primary objective of this article is to summarize the current knowledge on ex-vivo FCM accuracy metrics in assessing the margin status of surgical specimens of non-skin solid tumors.

Methods
This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocol (PRISMA-P) and the Cochrane Handbook for Systematic Reviews on Diagnostic Test Accuracy. Before conducting the review, a study protocol was developed and registered on the PROSPERO database (ID: CRD42022372558), outlining the search strategy, inclusion criteria, and risk of bias measures.

Search Strategy
An electronic search was performed on MEDLINE via PubMed, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), Web of Science, and Scopus. The search strategy is available in the Appendix.

Eligibility
Study screening, data extraction, and risk of bias assessment were performed using the automated screening Covidence developed by the Cochrane Collaboration. Two independent reviewers (M.A. and R.A.-M.) screened the titles and abstracts, as well as the full texts of potentially eligible studies. Both randomized and non-randomized studies were included. Abstracts presented at conferences were included, whereas case reports (with two or fewer patients) and non-English papers were excluded. Non-peer-reviewed articles, editorial features, and news stories were also excluded. No time restrictions on the date of publication were used, and ongoing and unpublished studies were not included.

Study Population and Target Condition
Studies including all types of cancer, except skin cancer, were eligible. Studies performed in animal models were excluded.

Intervention
Patients who underwent surgical resection were included. Only analysis of the resected specimen was included, whereas evaluation of biopsies was excluded.

Index Test
All types of ex vivo FCMs were included in this study, whereas studies evaluating in vivo handheld probe-based devices designed for endomicroscopic use were excluded unless the tissue was evaluated in an ex vivo setting. Other optical laser microscopies such as optical coherence tomography (OCT), Raman spectroscopy, near-infrared spectroscopy, photoacoustic microscopy, and narrow band imaging were all excluded.

Reference Standard
Definitive histopathology report of the tumor status or the clinical routine (FFPE and hematoxylin and eosin-stained permanent section).

Outcomes
Studies reporting the accuracy of ex-vivo FCM were included. Only studies reporting a binary outcome were considered, detecting the presence of PSMs or the presence of a tumor detected by the evaluator. Studies with only diagnostic or qualitative outcomes were excluded.

Data Extraction
Data extraction was performed by four reviewers (M.A., R.A.-M., T.A.-H., and G.S.), who extracted the following information: study title, author, year of publication,
country, full-text availability, study design, type of confocal microscope evaluated, type of control, blinding, profession of evaluator(s), number of evaluator(s), type of tumor, number of patients, number of images, number of PSMs identified by FCM and histopathology, sensitivity, specificity, and the rate of concordance between the index test and the reference standard (Cohen’s kappa $\kappa$-agreement).

**Quality Assessment of Included Studies**

As recommended by the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, we used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool\(^{[16]}\) to assess the risk of bias of included studies. Items were individually tailored for use in this review. The quality assessment involved four domains: patient selection, index test, reference standard, and flow and timing. Each domain included three questions, with two signaling questions and one risk of bias assessment. Signaling questions could be answered as “Yes,” “No,” or “Unclear,” whereas the risk of bias and concerns for applicability could be rated as “High,” “Low,” or “Unclear.” The results of the risk of bias assessment and applicability concerns were used to inform the strength of the evidence. The template for the risk of bias assessment is available in the Appendix.

**Statistical Analysis**

As per the protocol, we planned a meta-analysis of accuracy results to be performed if more than three eligible studies in a particular tumor type were retrieved and the same outcome measures were reported. Forest plots of sensitivity and specificity, pooled sensitivity and specificity, and inconsistency would be calculated from the data retrieved, using a binary covariate model. In addition, a funnel plot would be constructed to test for publication bias.

However, due to the limited number of valid and applicable studies retrieved, the authors reached a consensus not to perform a meta-analysis. Only two studies on breast and two on prostate had low applicability concerns and risk of bias, so summary estimates were not presented, in accordance with the recommendations from the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.\(^{[17]}\)

**Results**

**Study Selection**

We conducted a comprehensive database search on December 23, 2022. Figure 1 shows the PRISMA flow diagram. Initially, 2559 studies were identified, and after removing duplicates, 2120 titles were retained for screening. Of these, 50 full-text studies were assessed for eligibility, with 14 eventually meeting our inclusion criteria for data extraction. We were unable to retrieve the full text for one study.

**Excluded Studies**

Of the 36 studies that underwent full-text assessment, we excluded 15 that focused on diagnostic accuracy for cancer subtypes rather than margin assessment. Nine studies did not meet our inclusion criteria for the index test, and seven evaluated margin status but lacked retrievable sensitivity and specificity data. Of these, four were abstracts from medical conferences. We also excluded two studies that used animal models, and three others that were unrelated to cancer or focused on ablation margins or medical device engineering.

**Quality of the Studies**

Our risk of bias and applicability assessment is presented in Tables 1 and 2 and Fig. 2. We found a high risk of bias in patient selection (21%), index test assessment (36%), and flow and timing (14%). We also identified applicability concerns in patient selection (29%) and index test (29%) because some studies focused on diagnostic accuracy rather than margin evaluation, despite including evaluations of tumor present at the margin. Although we included these studies in our analysis for data completeness, we acknowledge that the lack of healthy tissue as a comparator in some studies remains a concern. In addition, some of the studies included in our analysis were retrospective case series and feasibility studies, which raises concerns about blinding strategies and patient selection criteria.

**Results of the Included Studies**

Fourteen studies were included in this systematic review, evaluating breast cancer,\(^{18–31}\) prostate cancer,\(^{20,27}\) brain tumors,\(^{26–28}\) kidney,\(^{29}\) bladder,\(^{30}\) and conjunctival tumors.\(^{31}\) Of these, six studies used the VivaScope FCM (München, Germany), four used the Histolog FCM (Lausanne, Switzerland), three used the Carl Zeiss FCM (Jena, Germany), and one used a multimodal microscope that combined fluorescence microscopy, OCT, and FCM. Tables 3 and 4 summarizes the accuracy measurements of the included studies. However, 95% confidence intervals are not presented, as they were not reported in all but two studies.\(^{20,27}\)
Breast Cancer

The four included studies evaluated a total of 123 breast specimens.\textsuperscript{18–21} Two studies reported high accuracy levels (>90%) for margin assessment.\textsuperscript{18,21} Golatta et al.\textsuperscript{18} published an abstract on 50 patients, reporting a sensitivity of 45.8% for the Histolog scanner compared with 39.6% with routine clinical practice involving ultrasound (US), with more accurate information anticipated when the full paper is available. Scimone et al.\textsuperscript{21} studied a multimodal FCM in 20 specimens, reporting no false negatives and two false positives.

Sandor et al.\textsuperscript{20} reported a sensitivity of 54% and specificity of 85% for the detection of PSM using the Histolog scanner with an en-face technique. Seven PSMs were identified, suggesting a potential reduction in re-excision rates of 58.3% if surgery was guided using this methodology (15). In contrast, Nackenhorst et al.\textsuperscript{19} were unable to identify any true PSM due to the absence of ink on cross-sectional images. However, they report that adipose tissue architecture in FCM images was comparable.
to FFPE sections and superior to FS, making it an attractive alternative to process tissue rich in adipose cells.

**Prostate Cancer**

A total of 146 prostate specimens were evaluated across four studies.\textsuperscript{22–25} The majority of studies demonstrated an almost perfect $\kappa$-agreement (>80%) between FCM and FFPE for assessing tumor margins, and no PSMs detected in one study.\textsuperscript{23}

The group led by Rocco et al. contributed three of the included studies. In their initial study, they utilized the VivaScope microscope to analyze 11 specimens, all of which had negative margins and were concordant with FFPE.\textsuperscript{23} In their second study, they described a Mohs-like shaving technique and identified a PSM in one out of eight patients.\textsuperscript{24} In the last study, they analyzed 24 specimens, detecting four PSMs intraoperatively. After secondary resection, all 24 patients had negative final surgical margin status at the site adjacent to the neurovascular bundles.\textsuperscript{24}

However, one patient was found to have a PSM outside the Mohs section imaged by FCM. Baas et al.\textsuperscript{22} evaluated the performance of the Histolog scanner using an en-face technique, and found a high sensitivity (86%) and specificity (96%) against FFPE, with a near perfect $\kappa$-agreement (79.7%). They also compared the technique against FS analysis (NeuroSAFE) and reported a $\kappa$-agreement of 80%.

### Central Nervous System Neoplasms

A total of 243 specimens were evaluated across three studies for brain tumors. One study used the desktop-based Zeiss LSM 710/5Live DUO system (Carl Zeiss Meditec, Oberkochen Germany), and two studies used the CONVIVO confocal laser endomicroscopy (CLE) system (Carl Zeiss Meditec, Oberkochen Germany). FCM exhibited a high specificity (ranging from 90% to 100%) for diagnosing neurological tumors, whereas sensitivity varied from 36% to 95%.\textsuperscript{26–28} Cohen's $\kappa$-agreement on the presence of tumor was 77.8%-100%\textsuperscript{27,28} and not reported nor retrievable in one study.\textsuperscript{26}

Martirosyan et al.\textsuperscript{28} classified any images showing any distinguishable tumor irrespective of subtyping as positive, whereas non-diagnostic images were classified as negative, resulting in a 94.9% sensitivity and 100.0% specificity. Acerbi et al.\textsuperscript{26} evaluated only glioblastoma specimens and because sensitivity and specificity were not reported, 66.7% of the biopsies taken at the margin were concordant with the FFPE. Belykh et al.\textsuperscript{27} took a different approach, and interpreted CLE images as lesional and non-lesional, with non-diagnostic images included in non-lesional. The prevalence of positive tumor diagnosed by FFPE was 84%, and the sensitivity and specificity achieved by a trained neuropathologist were 72% and 90%, respectively. Interestingly, the accuracy of surgeons to interpret CLE images was

---

**Table 1. Results of QUADAS-2 Risk of Bias Assessment of the Included Studies.**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Tumor Type</th>
<th>Patient Selection</th>
<th>Index Test</th>
<th>Reference Standard</th>
<th>Flow and Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golatta et al.\textsuperscript{18}</td>
<td>Breast</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Nackenhorst et al.\textsuperscript{19}</td>
<td>Breast</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Sandor et al.\textsuperscript{20}</td>
<td>Breast</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Scimone et al.\textsuperscript{21}</td>
<td>Breast</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Baas et al.\textsuperscript{22}</td>
<td>Prostate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Rocco et al.\textsuperscript{23}</td>
<td>Prostate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Rocco et al.\textsuperscript{24}</td>
<td>Prostate</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Rocco et al.\textsuperscript{25}</td>
<td>Prostate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Acerbi et al.\textsuperscript{26}</td>
<td>Brain</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Belykh et al.\textsuperscript{27}</td>
<td>Brain</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Martirosyan et al.\textsuperscript{28}</td>
<td>Brain</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Esperto et al.\textsuperscript{29}</td>
<td>Bladder</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Vreuls et al.\textsuperscript{30}</td>
<td>Kidney</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Iovieno et al.\textsuperscript{31}</td>
<td>Eye</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Count</th>
<th>Patient Selection</th>
<th>Index Test</th>
<th>Reference Standard</th>
<th>Flow and Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>10</td>
<td>9</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>High</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Unclear</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

also assessed, and the experienced neurosurgeon performed slightly better with a sensitivity and specificity of 74% and 92%.

**Other Cancers**

The studies included in our review evaluated surgical margins of bladder, kidney, and conjunctival tumors. Esperto et al.\(^{29}\) reported perfect concordance between FCM and FS, with high sensitivity (88%) and specificity (81%) for urethral margins when compared with the FFPE. Vreuls et al.\(^{30}\) investigated the use of FCM in partial nephrectomies and demonstrated perfect concordance (100%) on the presence of tumor, with both FCM and FFPE identifying two PSMs in six tumor specimens with sensitivity and specificity of 100%. Iovieno et al.\(^{31}\) evaluated 16 consecutive patients with conjunctival tumors correctly identifying three PSMs using FCM.
Is FCM Accurate in Assessing Surgical Margins?

Discussion

To our knowledge, this is the first systematic review to aggregate knowledge on the accuracy of ex-vivo FCM for assessing surgical margins in non-cutaneous solid tumors. Because our findings suggest that FCM is a promising tool with a high sensitivity, specificity, and concordance across most cancer types, it should be noted that most studies have small sample sizes, and caution is warranted due to the high risk of bias and absence of a priori power calculation. Moreover, given the diverse morphological features among tumor subtypes and differences in margin assessment across surgical specialties, we will discuss our findings separately on each tumor type.

Breast Cancer

Clinical guidelines recommend secondary resection of the tumor bed when a PSM is found during breast conserving. A meta-analysis reported that cytology had the highest pooled sensitivity (91%) and specificity (95%), followed by FS (86% sensitivity, 96% specificity) and optical imaging (85% sensitivity, 87% specificity). Intraoperative US and radiography demonstrated low sensitivity (59% and 53%, respectively) and specificity (81% and 84%, respectively). The poor uptake of FS and cytology in routine practice is limited by slow turnaround time and logistical constraints. FCM has theoretical advantages that may overcome these limitations. Three of the included studies showed improved accuracy for diagnosing the presence of a PSM compared with FFPE. The multimodal bespoke microscope used in the study by Scimone et al. had the highest reported accuracy, although the interpretation of the results should be taken with care owing to the study being sponsored by the industry and high risk of bias in patient selection. Unfortunately, no other studies evaluating this platform were found.

Nevertheless, the use of FCM in evaluating margins is more advanced in breast cancer than in other tumor types, and several studies have reported that FCM can distinguish clear morphological features and potentially reduce re-operation rates. Future research should evaluate the impact of FCM on oncological outcomes and assess whether it can be incorporated as a new standard of care.

Prostate Cancer

In radical prostatectomy, achieving oncological control while minimizing the risk of post-surgical complications such as incontinence and erectile dysfunction remains a challenge for surgeons. A significant barrier to performing nerve-sparing surgical dissection near the prostate is the concern regarding the extent of cancer involvement in the adjacent nerve bundles. NeuroSAFE, a standardized intraoperative FS technique, can address this risk. However, like FS in other tumor types, limitations of NeuroSAFE could be addressed by FCM. Previous studies have proven that pathologists can use FCM to reliably identify prostate cancer in...
biopsy cores. Our review confirms that FCM also has high accuracy when applied to margin evaluation of the posterolateral margins of the prostate.

Interestingly, Baas et al. identified that in patients undergoing secondary resection due to a PSM identified on NeuroSAFE, all PSMs of ≤5 mm were negative on the secondary resected specimens. This provides further evidence to strengthen the concept that small (<3 mm) PSMs may be irrelevant and have the same risk of biochemical recurrence as those with negative margins.

It is important to highlight that we found several potential bias concerns. Blinding was not performed in one study, and interpretation was performed by a single pathologist in all studies. Furthermore, most studies had a small sample size and low prevalence of PSMs. In addition, the majority of the available studies were conducted by the same research team, which may limit the generalizability of the results.

Central Nervous System Neoplasms

Neurosurgical treatment for central nervous system (CNS) neoplasms aims to achieve complete resection while preserving functional safety. However, due to the infiltrative nature of histological subtypes such as gliomas and the high risk of damaging vital brain structures, intraoperative margin assessment in neurosurgery differs from other tumor types. First, due to the associated side effects of removing neural tissue, the detection of PSMs relied on biopsies instead of evaluating the margin of resected specimens. Second, two of the studies used a handheld probe device with a small field of view designed for in-vivo scanning. Third, a high proportion of tumor under FFPE was found in these studies, likely due to sampling at the tumor margin rather than the resection margin; therefore, reported sensitivity and specificity need to be interpreted with caution. Fourth, intraoperative assessment during brain surgery has two components, providing differential diagnosis and identifying tumor at the margin; the definition of concordance may vary in neurosurgery where margins are defined as presence of tumor, infiltrating tumor, or healthy tissue.

Surprisingly, in the study by Belykh et al., an experienced neurosurgeon trained in FCM image interpretation was able to diagnose Gliomas with high accuracy compared to an inexperienced neurosurgeon supported by a non-specialist pathologist. This finding supports the idea that the interpretation of confocal images may have a shorter learning curve and could be a reliable tool if

### Table 4. Diagnostic Performance of FCM in Margin Assessment Across Included Studies.

<table>
<thead>
<tr>
<th>Reference</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV</th>
<th>NPV</th>
<th>κ-Agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golatta et al.</td>
<td>48</td>
<td>252</td>
<td>0</td>
<td>0</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100.0a</td>
</tr>
<tr>
<td>Nackenhorst et al.</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>0a</td>
<td>91a</td>
<td>0a</td>
<td>83a</td>
<td>87.9a</td>
</tr>
<tr>
<td>Sandor et al.</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>54</td>
<td>85</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Scimone et al.</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>100a</td>
<td>83a</td>
<td>80a</td>
<td>100a</td>
<td>90.0a</td>
</tr>
<tr>
<td>Prostate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baas et al.</td>
<td>12</td>
<td>79</td>
<td>3</td>
<td>2</td>
<td>86</td>
<td>96</td>
<td>80</td>
<td>98</td>
<td>79.7</td>
</tr>
<tr>
<td>Rocco et al.</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>100a</td>
<td>NA</td>
<td>100a</td>
<td>100.0a</td>
</tr>
<tr>
<td>Rocco et al.</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100.0a</td>
</tr>
<tr>
<td>Rocco et al.</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100.0a</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acerbi et al.</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>36a</td>
<td>100a</td>
<td>100a</td>
<td>10a</td>
<td>NR</td>
</tr>
<tr>
<td>Belykh et al.</td>
<td>73</td>
<td>18</td>
<td>2</td>
<td>30</td>
<td>72</td>
<td>90</td>
<td>97</td>
<td>38</td>
<td>77.8a</td>
</tr>
<tr>
<td>Martirosyan et al.</td>
<td>99</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>95a</td>
<td>100a</td>
<td>100a</td>
<td>58a</td>
<td>90.6a</td>
</tr>
<tr>
<td>Bladder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esperto et al.</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0.88</td>
<td>0.81</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Kidney</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vreuls et al.</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100.0a</td>
</tr>
<tr>
<td>Eye</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iovieno et al.</td>
<td>3</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100a</td>
<td>100.0a</td>
</tr>
</tbody>
</table>

Abbreviations: TP, true positive; TN, true negative; FP, false positive; FN, false negative; PPV, positive predictive value; NPV, negative predictive value; κ, Cohen’s kappa coefficient; NR, not retrievable; NA, not applicable; FCM, fluorescence confocal microscopy.

*aExtrapolated from the raw data provided from the article.

*aAlthough concordance between tumor versus non-tumor is 100%, the value presented is the concordance between if margin is defined as tumor, infiltrative, or healthy tissues.
the evaluator is adequately trained. Overall, FCM accuracy for guiding resections of CNS tumors appears promising. However, the available evidence is limited by the quality of the studies.

**Other Cancers**

Esperto et al.\(^\text{29}\) and Vreuls et al.\(^\text{30}\) both presented their findings in bladder and renal cancer, respectively, as abstracts. However, we anticipate that more accurate and comprehensive results will be published upon the release of their full reports. Of note, Vreuls et al.\(^\text{30}\) mentioned that FCM could diagnose angiomyolipoma, but differentiating between types of renal cell carcinoma was not feasible.

Iovieno et al.\(^\text{31}\) were the first to conduct a prospective trial to investigate margin assessment using FCM. The study provided additional details about FCM's morphological characteristics in different tissue samples. Unfortunately, FCM analysis in this study was not blinded.

**Different Types of Confocal Microscopes**

Most breast cancer articles employed Histolog FCM, whereas VivaScope FCM was predominantly used in urology studies. Both microscopes are desktop-based systems. The VivaScope system employs a dual laser system, with a reflectance mode at a wavelength of 785 nm and fluorescence at 488 nm, enabling dual contrast and high-resolution images in various depths. In contrast, the Histolog FCM excites tissue fluorescence with a laser at 488 nm, and fluorescence emission with a wavelength of >500 nm is collected, producing a toluidine-blue monostain-like image. The scanner tray of Histolog FCM allows a larger scan area (48 × 36 mm), whereas VivaScope FCM provides a maximum total scan area of 25 × 25 mm and a higher vertical resolution of at least 4 μM with a magnification of ×550. However, VivaScope FCM takes up to 4 or 5 min per sample, sacrificing contrast and depth capabilities compared with the Histolog FCM which requires approximately 50 sec per image.

The CONVIVO system operates as a miniaturized confocal microscope that operates using a 488-nm wavelength blue laser light and provides a real-time field of view of 475 × 275 μM. The desktop Carl Zeiss FCM is the previous version of the CONVIVO system. Both microscopes have the advantage of fast image acquisition that can be transformed into a video, allowing both ex-vivo analysis and in-vivo analysis when applied as a handheld probe. However, the field of view restricts its use in larger samples.

Scimone et al. (Scimone, Krishnamurthy, Maguluri, Preda, Park, Grimble, Song, Ban and Iftimia, 2021)\(^\text{21}\) used a multimodal microscope that combines fluorescence microscopy, FCM, and OCT. The images acquired were processed by a semi-automated algorithm generating a series of en-face and cross-sectional images. The proof-of-concept study showed excellent sensitivity (100%) and specificity (83.3%), but no other studies have confirmed these findings. Further evaluation directly comparing the performance of each FCM is needed to determine the preferred brand specific to the tissue type.

**Limitations**

Several limitations of this review must be acknowledged. First, due to the novelty of FCM as a technique, the literature lacks clinical trials with adequate power, and most of the studies identified aimed to demonstrate proof of concept. As a result, despite being planned in the protocol, conducting a meta-analysis was not feasible. In addition, some studies focused primarily on diagnostic accuracy rather than margin assessment, but they were included as we were able to retrieve or extrapolate FCM's sensitivity and specificity from available results. Last, cancer subtypes were not further classified in our review, and because there are morphological differences between histological subtypes, the accuracy of FCM may vary, and the ability to detect the presence of tumor may be affected.

**Future Perspective**

As interest in this technology grows, subsequent studies should be planned with adequate power calculations to provide narrow confidence interval estimates on the accuracy of this technique. To reduce the risk of evaluator bias, blinded assessment should be incorporated. In addition, studies should aim to explore FCM potential application in more tumor types, including those that are difficult to evaluate using FS. Furthermore, studies should also focus on exploring the cost-effectiveness of FCM and its potential impact on patient outcomes. To ensure consistency and comparability across studies, we encourage authors to consider following the (Standard for Reporting of Diagnostic Accuracy Studies) (STARD) guidelines for diagnostic accuracy in their reports.\(^\text{40}\)

In conclusion, currently available evidence shows FCM has potential to facilitate margin assessment of multiple tumor types, given its apparent high diagnostic accuracy. FCM may enable more centers to perform real-time evaluation of PSMs and transform the pathology workflow. However, the current limitations in the evidence emphasize the need for more rigorously designed studies to confirm FCM diagnostic accuracy before it is incorporated in clinical practice.
Appendix

Search Strategy

The electronic databases and search strategy that will be included are as follows:

- MEDLINE (via PubMed)
  - (“Microscopy, Confocal”[Mesh] OR Fluorescence OR Laser OR Histolog OR VivaScope) AND “Margins of Excision”[Mesh]

- Embase
  - (((Histolog or VivaScope or Confocal) and (Surgical margin or margin of excision or resection margin)).af.

- CENTRAL
  - (Microscopy, Confocal):ti,ab,kw OR (VivaScope):ti,ab,kw OR (Histolog):ti,ab,kw AND (Margins of Excision):ti,ab,kw

- Web of Science
  - (((TS=(confocal) OR TS=(con-focal) OR TS=(laser) OR TS=(laser-scanning)) AND (TS=(microscopies) OR TS=(microscopy)) OR TS=(histolog) OR TS=(vivascope)) AND (TS=(excision) OR TS=(resection) OR TS=(surgical) OR TS=(positive) OR TS=(negative) OR TS=(tumour-free) OR TS=(tumor-free) OR TS=(tumour) OR TS=(tumor) OR TS=(free) OR TS=(assessment)) AND (TS=(margin) OR TS=(margins)))

- Scopus
  - TITLE-ABS-KEY (((confocal OR con-focal OR laser OR laser-scanning) AND (microscopies OR microscopy)) OR (histolog OR vivascope)) AND ((excision OR resection OR surgical OR positive OR negative OR tumour-free OR tumour OR free OR assessment) AND (margin OR margins)) AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “cr”))

Quality Assessment

QUADAS Template

1. Patient selection 1
   Could the selection of patients have introduced bias?
   - Yes
   - No
   - Unclear

2. Patient selection 2
   Did the study avoid inappropriate exclusion?
   - Yes
   - No
   - Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

3. Patient selection 3
   Are there any concerns that the included patients do not match the review question? Could the selection of patients have introduced bias (high = high risk of bias; low = low risk of bias)?
   - High
   - Low
   - Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

4. Index test 1
   Were the FCM results interpreted without the knowledge of the results of the control?
   - Yes
   - No
   - Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

5. Index test 2
   Could the interpretation of the FCM results have introduced bias?
   - Yes
   - No
   - Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

6. Index test 3
   Are there concerns that the index test or its interpretation differs from the review question? Could the conduct or interpretation of the FCM results have introduced bias (high = high risk of bias; low = low risk of bias)?
   - Yes
   - No
   - Unclear

(Extractors will also be able to add supporting text to justify their judgments.)
Is FCM Accurate in Assessing Surgical Margins?

- High
- Low
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

7. Reference standard 1

Is the reference standard likely to correctly classify the target condition? (Is this reference standard performed by an adequately trained pathologist?)

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

8. Reference standard 2

Were the reference standard results interpreted without the knowledge of the results of the index test (FCM)?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

9. Reference standard 3

Could the reference standard, its conduct, or its interpretation have introduced bias (high = high risk of bias; low = low risk of bias)?

- High
- Low
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

10. Flow and timing 1

Did all patients receive a reference standard?
Did all patients receive the same reference standard?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

11. Flow and timing 2

Were all patients included in the analysis?
Are the proportions and reasons for missing data similar across index tests?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

12. Flow and timing 3

Could the patient flow have introduced bias (high = high risk of bias; low = low risk of bias)?

- High
- Low
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

13. Other sources of bias

State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

Applicability Concerns

14. Patient selection

Are there any concerns in regard to the applicability to our research question?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)
15. Index test

Are there any concerns in regard to the applicability to our research question?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

16. Reference standard

Are there any concerns in regard to the applicability to our research question?

- Yes
- No
- Unclear

(Extractors will also be able to add supporting text to justify their judgments.)

Competing Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Author Contributions
All authors have contributed to this article as follows: design and plan of the review (MA, RA, GS) and literature search (MA, RA). All authors reviewed the included studies and drafted and revised the manuscript. All authors have read and approved the final manuscript.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs
Matthew Au https://orcid.org/0000-0003-4977-531X
Ricardo Almeida-Magana https://orcid.org/0000-0002-3437-1649
Tarek Al-Hammouri https://orcid.org/0000-0003-2377-0545

Literature Cited


