Systematic review of diagnostic accuracy of reflectance confocal microscopy for melanoma diagnosis in patients with clinically equivocal skin lesions
Stevenson, Alexander D; Mickan, Sharon; Mallett, Susan; Ayya, Mekhala

Published in:
Dermatology practical & conceptual

DOI:
10.5826/dpc.0304a05

Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of record

Link to publication in Bond University research repository.

Recommended citation(APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.
Systematic review of diagnostic accuracy of reflectance confocal microscopy for melanoma diagnosis in patients with clinically equivocal skin lesions

Alexander D. Stevenson¹, Sharon Mickan², Susan Mallett³, Mekhala Ayya¹

¹ University of Oxford, Oxford, England, United Kingdom
² Centre for Evidence-Based Medicine, Department of Primary Care Health Sciences, University of Oxford, Oxford, England, United Kingdom
³ Department of Primary Care Health Sciences, University of Oxford, Oxford, England, United Kingdom

Key words: reflectance confocal microscopy, melanoma, diagnosis, systematic review, meta-analysis, dermoscopy


Received: June 5, 2013; Accepted: August 25, 2013; Published: October 31, 2013

Copyright: ©2013 Stevenson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: None.

Competing interests: The authors have no conflicts of interest to disclose.

All authors have contributed significantly to this publication.

Corresponding author: Alexander Stevenson, FRACGP, Isabella Plains Medical Centre, Canberra ACT Australia. Email address: dradstevenson@gmail.com.

Background: Melanoma is a cancer of the skin and is increasing in incidence in the UK and Europe. Melanoma is a condition that is often curable if detected at an early stage, which makes accurate diagnosis vital. Reflectance confocal microscopy (RCM) is a tool used to image the skin. It gives high magnification images of the skin, which may provide more accurate diagnosis of lesions that are equivocal on clinical examination and dermoscopy.

Objective: To determine the diagnostic accuracy of reflectance confocal microscopy (RCM), for melanoma diagnosis, as an add-on test to clinical examination and dermoscopy in the diagnosis of equivocal pigmented skin lesions using histopathology as the reference standard.

Methods: A search was conducted of MEDLINE, EMBASE and six other electronic databases from inception to present. Forward citation searching and hand searching of reference lists were also conducted. Diagnostic accuracy studies that assess RCM in the diagnosis of melanoma were included in the review. Two contributors conducted the search, data extraction and assessment of methodological quality using QUADAS-2. Statistical analysis was performed using hierarchical bivariate random effects meta-analysis.

Results: 951 titles and abstracts were screened. Five studies comprising 909 lesions were eligible for meta-analysis. Meta-analysis returned a per lesion sensitivity of 93% [95% CI 89-96] and a specificity of 76% [95% CI 68-83].

Conclusions: The utility of reflectance confocal microscopy (RCM) as an add-on test for the diagnosis of melanoma depends on the trade off between over-excising benign lesions and misdiagnosing melanoma as benign. This becomes important when considering lesions on surgically difficult or cosmetically important areas of the body.
Introduction

Melanoma is a cancer of the skin which is increasing in frequency both in the UK and Europe1. Cancer research UK (CRUK) have calculated that in the 35 years from 1975-2010 the age standardized incidence rate in the UK rose from 3.2 per 100000 to 17.2 per 1000002. The biggest risk factor for developing melanoma is exposure to ultraviolet light [3].

Prognosis for melanoma is very much dependent on the stage of the disease when it is diagnosed so early accurate diagnosis of melanoma is crucial. The five-year survival for stage 1A melanoma is 97%. The five-year survival drops rapidly to 10-15% for stage 4 metastatic disease [4]. This rapid decline in survival with higher stage is because the only potentially curative treatment is surgical excision [5]. Adjuvant therapy for non-metastatic melanoma has not yet been demonstrated to provide a survival benefit [6] and no therapy has proven to extend survival for metastatic melanoma [7,8].

The currently accepted best diagnostic method for melanoma is dermoscopy [9].

A recent meta-analysis of dermoscopy in the diagnosis of melanoma pooled the sensitivities and specificities and found a sensitivity of 91% and a specificity of 86% [10]. Most dermoscopy research has been conducted in white skinned populations however there is some evidence of the ability of dermoscopy to work equally well in non-white populations [11].

Reflectance confocal microscopy (RCM) also known as confocal laser scanning microscopy (CLSM) of the skin was first described in the early 1990s [12]. This technology uses a near infrared laser to obtain images of the top layers of the skin. These images are magnified such that they are “quasi-histological.” From the images, information can be obtained regarding cell structure and the architecture of the surrounding tissues. The images are analyzed and combinations of features are assessed to give a positive or negative diagnosis of melanoma. Several criteria have been developed to analyze images of RCM [13]. The test itself takes about ten minutes for imaging and evaluation of a skin lesion.

The goal of diagnosing melanoma is to correctly identify melanomas, while at the same time, excising as few benign lesions as possible. The most appropriate first line examination for this is dermoscopy, which has been shown to be a more accurate diagnostic tool than unaided eye examination [9]. Given the time needed to use RCM, it is most appropriate as a secondary examination add test to dermoscopy for lesions where dermoscopy does not give a confident diagnosis. This role has been suggested previously [14,15].

There have been many narrative reviews on the use of RCM in the diagnosis of melanoma. These articles have focused mainly on describing the technology and discussing its potential role in melanoma diagnosis. RCM technology has advanced since the first instruments were introduced in the 1990's. The devices are now small and ergonomically able to image most areas of the skin. The diagnostic features are easy to learn and reproducible [13,16]. The devices themselves are quite, they are in limited use in clinical practice [17] and combined with the time to assess each lesion, may restrict the use to specialist clinics.

A comprehensive search found no systematic reviews or meta-analysis. Systematic reviews are important as they allow for a more transparent and objective appraisal of the evidence. Meta-analysis where appropriate can enhance the precision of the estimates of individual studies [18]. The objective for this review is to examine the diagnostic accuracy of RCM in the diagnosis of melanoma as an add-on test for lesions that are clinically and/or dermoscopically equivocal/suspicious for melanoma in cohort studies that have used a predefined threshold. This must be a pre-defined scoring system or system of diagnosis but there is no restriction on the system. Meta-analysis will be conducted if there is sufficient consistency between studies in the way the thresholds are applied.

Methods

Search strategy

Electronic searches were conducted of Medline, Embase, CINAHL, the Cochrane Register of Diagnostic Test Accuracy Studies, DARE (Database of Abstracts of Reviews of Effects), Health Technology Assessment (HTA) database (The Cochrane Library), MEDION (The Medion database) and NHS Economic Evaluation Database (NHSEED). The detailed search strategy for Medline (PubMed) used the following terms, CSLM, Laser microscope*, Confocal microscope*, Confocal scanning microscopy*, Microscopy, confocal (MeSH), Melanoma (MeSH), Melanoma*, Hutchinson* freckle, Nevus, Nevi, Mole*, Skin cancer*, Cutaneous neoplasm*, Skin neoplasms (MeSH), Skin neoplasms, Nevus[MeSH], Melanocytes[MeSH], Skin tumour*, Skin tumor*, Skin lesion*, Melanocytic. The terms were adapted to image most areas of the skin. The diagnostic features are easy to learn and reproducible [13,16]. The devices themselves are quite, they are in limited use in clinical practice [17] and combined with the time to assess each lesion, may restrict the use to specialist clinics.

A comprehensive search found no systematic reviews or meta-analysis. Systematic reviews are important as they allow for a more transparent and objective appraisal of the evidence. Meta-analysis where appropriate can enhance the precision of the estimates of individual studies [18]. The objective for this review is to examine the diagnostic accuracy of RCM in the diagnosis of melanoma as an add-on test for lesions that are clinically and/or dermoscopically equivocal/suspicious for melanoma in cohort studies that have used a predefined threshold. This must be a pre-defined scoring system or system of diagnosis but there is no restriction on the system. Meta-analysis will be conducted if there is sufficient consistency between studies in the way the thresholds are applied.

The detailed search strategy for Medline (PubMed) used the following terms, CSLM, Laser microscope*, Confocal microscope*, Confocal scanning microscopy*, Microscopy, confocal (MeSH), Melanoma (MeSH), Melanoma*, Hutchinson* freckle, Nevus, Nevi, Mole*, Skin cancer*, Cutaneous neoplasm*, Skin neoplasms (MeSH), Skin neoplasms, Nevus[MeSH], Melanocytes[MeSH], Skin tumour*, Skin tumor*, Skin lesion*, Melanocytic. The terms were adapted for the other databases as appropriate. The searches were performed from database inception. The search was conducted by two independent reviewers.

Manual searches were conducted of the reference lists of the review articles and studies included in the final analysis. Forward citation searching was performed on all relevant retrieved articles via SCOPUS and Science Citation Index. The ‘related articles’ function of PubMed was used to look at the first 20 articles retrieved. No language restriction was applied to the electronic searches. The search was restricted to studies on humans. All the major journals were indexed. The searching authors felt that there was sufficient information provided in the articles and therefore correspondence with authors was not performed. The studies were included if they met the following criteria:
Assessment of methodological quality
Two authors independently assessed methodological quality of the studies using the QUADAS-2 tool [23]. Any disagreements were resolved by discussion. The results of the quality assessment are presented with a textual methodological quality summary and graphical representation.

Results
Search
The search of the databases was conducted on February 8, 2012. After screening for duplicates 951 studies were examined. A flow diagram of the search can be found in Figure 1.

After examining titles and abstracts the full text of 39 articles were retrieved. There were five articles that met the inclusion criteria. These are shown in Table 1.

Excluded studies
There were five studies, which were derivation studies, or studies that did not validate on a new set of patients. There were 15 descriptive correlation studies, which only described which RCM features were associated with melanoma. There were four case reports or small case series, two narrative review articles, one editorial and one study looking at observer agreement of the RCM features associated with melanoma.

Methodological quality assessment
The exclusion criteria for studies in the review included two major methodological quality criteria. The studies could not be case control studies nor could they be studies that set a diagnostic threshold i.e.: studies that developed a scoring system. Case control studies have been demonstrated to overestimate diagnostic accuracy when compared to cohort studies that use an appropriate spectrum [24]. Studies that derive/set a threshold use multivariable analysis to derive a score. These scores are derived on a certain population. It is very often the case that these scoring systems perform worse when they are validated in another population, however similar [25].

If a study presented several sensitivity and specificity estimates on a receiver operator characteristic curve (ROC) plot then the point estimate used for meta-analysis was the point chosen by the author of the article.

The results are presented graphically using RevMan5 [21]. The studies were combined in a statistical meta-analysis using the METANDI function in STATA [22].

Subgroup analyses was intended for investigation of operator experience and algorithm method however there was an insufficient number of studies.

Findings
Five studies were identified comprising 909 lesions. The average prevalence of melanoma was 36.2% with a range from 29-39. Three studies used the RCM diagnostic scor-
Per lesion sensitivity and specificity are shown on a forest plot in Figure 2. There appears to be minimal heterogeneity per lesion in sensitivity across the studies, with more heterogeneity in the specificity.

Based upon this a hierarchical summary receiver operator characteristic (HSROC) curve was obtained using the bivariate method. The diagnostic accuracy results are quite similar in all studies and there is no evidence of a threshold effect or apparent threshold effect. The plot of this is shown in Figure 3. From meta-analysis the operating point had a sensitivity of 93.3% [95% CI 88.5-96.2, range 91% to 97%] and a specificity of 75.9% [95% CI 67.9-82.5, range 68% to 86%].
patient outcomes compared to the existing diagnostic pathway. It is not enough just to measure the sensitivity and specificity [28].

Duff et al. [29] and Rampen et al. [30] followed patients after melanoma screening, searching for missed melanomas. Duff found no missed melanomas from 1961 patients and Rampen found seven invasive melanomas and two lentigo maligna (a type of melanoma in situ) from 9968 patients seen in the clinic. This data suggests that, in real clinical contexts using current diagnostic technology, few melanomas are missed.

The purpose of this review was to evaluate RCM as an add-on test to existing diagnostic pathways, not to evaluate it as a replacement test. It has been suggested that RCM is more sensitive than dermoscopy [13]. If all lesions that were suspicious to the unaided eye examination were examined given the low number of studies included in the review, statistical subgroup analysis and covariate hierarchical modeling for investigation of heterogeneity were not performed due to low statistical power.

**Discussion**

When examining the use of a new diagnostic test it is important to consider whether its introduction will improve patient outcomes compared to the existing diagnostic pathway. It is not enough just to measure the sensitivity and specificity [28].

Duff et al. [29] and Rampen et al. [30] followed patients after melanoma screening, searching for missed melanomas. Duff found no missed melanomas from 1961 patients and Rampen found seven invasive melanomas and two lentigo maligna (a type of melanoma in situ) from 9968 patients seen in the clinic. This data suggests that, in real clinical contexts using current diagnostic technology, few melanomas are missed.

The purpose of this review was to evaluate RCM as an add-on test to existing diagnostic pathways, not to evaluate it as a replacement test. It has been suggested that RCM is more sensitive than dermoscopy [13]. If all lesions that were suspicious to the unaided eye examination were examined given the low number of studies included in the review, statistical subgroup analysis and covariate hierarchical modeling for investigation of heterogeneity were not performed due to low statistical power.

**TABLE 1.** Characteristics of included studies

<table>
<thead>
<tr>
<th>First author and year</th>
<th>Country of study</th>
<th>Algorithm</th>
<th>Patient Characteristics</th>
<th>Number of lesions</th>
<th>Reference standard</th>
<th>Number of patients</th>
<th>Melanoma frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curchin 2011</td>
<td>Australia</td>
<td>Pellacani 2005, Guitera 2010</td>
<td>Data not available</td>
<td>35</td>
<td>Histology</td>
<td>Not available</td>
<td>37%</td>
</tr>
<tr>
<td>Guitera 2009</td>
<td>Italy and Australia</td>
<td>Pellacani 2005</td>
<td>Median age and IQR 47(36-60), F:M 149:177</td>
<td>326</td>
<td>Histology</td>
<td>326</td>
<td>37%</td>
</tr>
<tr>
<td>Guitera 2010</td>
<td>Australia and Italy</td>
<td>Guitera 2010</td>
<td>Data not available</td>
<td>73</td>
<td>Histology/long term follow up</td>
<td>Not available</td>
<td>37%</td>
</tr>
<tr>
<td>Pellacani 2007</td>
<td>Italy and Australia</td>
<td>Pellacani 2005</td>
<td>Median age and IQR 47.7(35.9-60.4), F:M 158:174</td>
<td>351</td>
<td>Histology</td>
<td>332</td>
<td>37%</td>
</tr>
<tr>
<td>Langley 2007</td>
<td>Canada</td>
<td>No named diagnostic algorithm</td>
<td>Average age: 44 range: 16-84</td>
<td>125</td>
<td>Histology</td>
<td>125</td>
<td>29%</td>
</tr>
</tbody>
</table>

**TABLE 2. QUADAS-2 Risk of bias assessment.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk Of Bias</th>
<th>Applicability Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient Selection</td>
<td>Index Test</td>
</tr>
</tbody>
</table>

© Low Risk  © High Risk  ? Unclear Risk

[Copyright: ©2013 Stevenson et al.]

**Figure 2.** Forest plot of the studies. [Copyright: ©2013 Stevenson et al.]
If the role of RCM is as an add-on test, all lesions that are examined with RCM have already been declared as positive by the existing pathway. Using RCM in this way will not increase the detection rate of the few melanomas that are currently being misdiagnosed as benign as they will have already left the diagnostic pathway. It is possible that the availability of RCM may change clinician confidence and diagnostic threshold. Instead of the clinical/dermoscopic diagnosis being the final step before a management decision is made, RCM would exist as an add-on test. The individual clinician might change his or her threshold to be more sensitive and less specific in order to capture more disease. This has not been addressed in these studies however it may change the sensitivity and specificity of RCM in actual clinical practice. Another area where it may be helpful is if the clinician is suspicious of a lesion, especially featureless pink lesions, and are considering monitoring it the clinician may well use RCM and find that it is positive and proceed with excision. The risk here is if it is RCM negative and the clinician does not follow it up with some monitoring procedure they may miss a melanoma.

To gauge the trade off between the reduction in unnecessary biopsies and the missed melanoma diagnoses the sensitivity and specificity can be applied to an estimated prevalence of melanoma in the spectrum of patients that would be selected for RCM examination.

The average prevalence of melanoma in the studies included in the review was 36%. In a 2002 systematic review of dermoscopy the mean frequency of melanoma was 28%. Previous research has suggested a malignant to benign ration of 1:4 with the expert use of dermoscopy [32]. This translates to a frequency melanoma in dermoscopy positive lesions of 20%. If we assume that in real clinical practice the clearly positive melanomas would not be examined with RCM, we can gauge an estimated frequency of disease in dermoscopy positive lesions would be slightly lower.

If RCM were to be used as an add-on test in clinical practice, the population examined with RCM would be the narrow pre-selected group of those in whom dermoscopy was not clearly positive or not clearly negative. The population of lesions being examined with RCM in these studies was not clear and reproducible. The terms “clinically suspicious” and “equivocal” do not give the reader sufficient information. It is not certain that the lesions examined by RCM in the studies were the same that would be examined by RCM in clinical practice. If the lesions examined in these studies included those that were clearly melanoma then the spectrum of disease would be different to that in clinical use and this could bias the result leading to an over estimate of sensitivity and specificity.

These factors combined with the concept that diagnostic accuracy determined from laboratory condition studies may be different from the diagnostic accuracy in the real life clinical setting [31], mean that the external validity of these results has to be taken cautiously.

If the role of RCM is as an add-on test, all lesions that are examined with RCM have already been declared as positive by the existing pathway. Using RCM in this way will not increase the detection rate of the few melanomas that are currently being misdiagnosed as benign as they will have already left the diagnostic pathway. It is possible that the availability of RCM may change clinician confidence and diagnostic threshold. Instead of the clinical/dermoscopic diagnosis being the final step before a management decision is made, RCM would exist as an add-on test. The individual clinician might change his or her threshold to be more sensitive and less specific in order to capture more disease. This has not been addressed in these studies however it may change the sensitivity and specificity of RCM in actual clinical practice. Another area where it may be helpful is if the clinician is suspicious of a lesion, especially featureless pink lesions, and are considering monitoring it the clinician may well use RCM and find that it is positive and proceed with excision. The risk here is if it is RCM negative and the clinician does not follow it up with some monitoring procedure they may miss a melanoma.

To gauge the trade off between the reduction in unnecessary biopsies and the missed melanoma diagnoses the sensitivity and specificity can be applied to an estimated prevalence of melanoma in the spectrum of patients that would be selected for RCM examination.

The average prevalence of melanoma in the studies included in the review was 36%. In a 2002 systematic review of dermoscopy the mean frequency of melanoma was 28%. Previous research has suggested a malignant to benign ration of 1:4 with the expert use of dermoscopy [32]. This translates to a frequency melanoma in dermoscopy positive lesions of 20%. If we assume that in real clinical practice the clearly positive melanomas would not be examined with RCM, we can gauge an estimated frequency of disease in dermoscopy positive lesions would be slightly lower.

Figure 4 demonstrates a flow diagram of the impact of RCM in its proposed role as an add on test to dermoscopy using a sensitivity of 93% and a specificity of 75% as calculated in the meta-analysis and a melanoma frequency of 20%. For 1000 dermoscopy positive lesions, there would be 200 melanomas. RCM would correctly identify 186 of these and miss [14]. There would be 192 benign lesions excised and 608 benign lesions not excised.

The only benefit of RCM in this pathway is to increase the specificity of diagnosis and reduce the number of benign lesions excised. The value of reflectance confocal microscopy as an add-on test in the diagnosis of melanoma depends on the trade off between the harms associated with excising benign lesions and the harms associated with misdiagnosing a melanoma as benign. If RCM is to be used in clinical practice a decision has to be made weighing up the consequences.
A weakness of this review is that the current studies may not have focused on the pertinent patient populations to test the ability of RCM as an add-on test to dermoscopy. It is noted that in three of the five studies included in this review the main operators using RCM were Giovanni Pellacani and...
References


