Systematic Review and Meta-Analysis of the Accuracy of Confocal Microscopy in the Diagnosis of Skin Cancer

Rahman A1,*, Miller S2, Whittington B2, Tee C2, Mclellan G2, Cole D2, Spence R2, Rajpara S1,2

1Aberdeen Royal Infirmary, Foresterhill Aberdeen, Scotland, UK
2University of Aberdeen, Aberdeen, UK

*Corresponding author: Dr. Atiya Rahman, 311, Street No 39, G 9/1 Islamabad, Pakistan ; Tel: 0923345252383; Email: atiya_rahman7@yahoo.com

Introduction

The incidence of non melanoma skin cancer, BCC and SCC has increased by 10% per annum; with 2-3 million new cases diagnosed worldwide each year [1]. Worldwide the incidence of melanoma is increasing faster than any other cancer, with rates doubling every 10-20 years in the Caucasian population. In the UK the incidence of melanoma has quadrupled over the last 40 years [2].

Early diagnosis is of utmost importance in the management of skin cancers to prevent or minimize the morbidity and mortality associated with them. In specialized centres the accuracy of clinical diagnosis, on the basis of an unaided eye, of melanoma is only around 60% [3]. Histology remains the gold standard for the diagnosis of skin cancers. Various non invasive methods like dermoscopy [4,5], Raman spectroscopy [6], CLSM [7] and others like positron emission tomography, ultrasonography, Doppler, computed tomography, magnetic resonance imaging, optical coherence tomography and terahertz imaging have been used to improve the diagnostic performance. Out of these, dermoscopy has been in routine use for the diagnosis of the skin cancer. It provides a non invasive, rapid, in vivo examination of the superficial layers of the skin. It has a sensitivity and specificity of 86% and 89% respectively for cutaneous melanoma [5]. The requirement to develop a better imaging tool remains. CLSM, a non-invasive diagnostic tool, was first reported in 1995 to be used in vivo on human tissue [8]. Since then, there has been many studies evaluating its role in viewing the microscopic features of normal skin [9] and various lesions such as cutaneous neoplasms [10,11], pigmented lesions [12,13,14], actinic keratosis [15], sebaceous gland hyperplasia [16], psoriasis [17], irritant and allergic contact dermatitis [18,19]. CLSM can distinguish melanocytes from other pigmented lesions like pigmented...
keratinocytes and melanophages [20]. Typical CLSM has a lateral resolution of 0.5-1µm and axial resolution of 3-5 µm, providing horizontal sections similar to thickness of histology sections. Maximum penetration of the skin is 350µm [21,22]. Different skin cancers have different morphological features which can be distinguished by using CLSM in vivo and ex vivo, enabling it to be potentially used to give a real time diagnosis to the patients negating the need to wait for histopathology results [22,23]. There are many articles but only a limited number of studies on the use of CLSM in the diagnosis of skin cancer. A systematic review was done to do an objective assessment of the available evidence.

Aim and Objective
The scope of this systematic review and meta analysis is to examine the diagnostic accuracy of CLSM in skin cancers in the selected studies which have met the inclusion criteria.

Materials and Methods

Search strategy
A comprehensive literature search of Medline (from 1950 to July 2013), Embase (from 1980 to August 2013), and all EBM reviews (ACP Journal Club, Cochrane Controlled Trials Register, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects) and Journals@Ovid was carried out using the following key words or corresponding MeSH terms: confocal microscopy, confocal laser scanning microscopy, reflectance confocal microscopy, confocal microscopy and skin cancer, confocal microscopy and histology, confocal microscopy and skin cancer and histology, to identify original studies, abstracts and systematic reviews. The resultant hits performed on 17th August 2013 are shown in Table 1. Studies fulfilling the following inclusion criteria were included without any language or time restrictions.

Type of studies
Original prospective and retrospective studies were included. These were the principal published reports of original data from case control or cohort studies and were independent from other studies to avoid including the results from the same study twice.

Target conditions
Melanoma and nonmelanoma skin cancers i.e. BCC and SCC.

Study population
The selected studies involved humans of any age, sex and ethnic group, and performed at primary, secondary or tertiary levels of healthcare involving equal to or more than 10 subjects.

Intervention
Confocal laser scanning microscopy (CLSM).

Reference standard
Histopathology of the excised skin lesions.

Study results
The studies reporting results in the form of true positives (TP), true negatives (TN), false negatives (FN) and false positives (FP) or results from which these values could be computed was included. Where a study showed different sensitivity and specificity values by taking into account different criteria, the results with the highest values were included for review.

Methodological quality
Assessment of the quality of articles was undertaken by AR, SM, CT, BW, GM, DC, RS and SR. 120 citations were identified as relevant to the topic. The titles and abstracts of all 120 retrieved articles were divided amongst the first 7 authors and were reviewed independently by them and then in the group with SR. Where the abstract was not available, full-text articles were retrieved for further evaluation. 56 citations were excluded after duplication or the title and/or abstract were not relevant for the end point of the study. In this way, 64 full text articles were assessed for eligibility. The methodological quality of the studies was assessed by the group independently and then together with SR with Standards for Reporting of Diagnosis Accuracy (STARD) initiative [24]. This checklist and National Institute for Health and Clinical Excellence (NICE) guidelines on methodology, checklist for diagnostic studies [25] formed the basis for this review and meta-analysis. Reviews, editorials, case reports, small case series and studies that were derivation studies, i.e. not done on new set of patients were excluded. Finally eight studies [21,26-32] met the criteria to be included for meta analysis. The evidence level of the studies was determined as per the NICE guidelines on development methods [33].

Data collection
Data from the eligible studies were extracted in a table adapted from NICE guidelines [34] (Table 1 and 2).

Data analysis
Meta-analysis was carried out using Meta-DiSc (available at: http://www.hrc.es/investigacion/metadisc_en.htm) version 1.4. For sensitivity and specificity of the included studies, heterogeneity was tested by using likelihood ratio tests, shown as a x² value on the forest plots. For the positive and negative likelihood ratios and DORs homogeneity was tested using Cochran's Q. The possibility of heterogeneity being due to use of different thresholds in different studies were investigated by plotting sensitivities and specificities from the studies as points on an SROC curves.

<table>
<thead>
<tr>
<th>Key words</th>
<th>Ovid embase</th>
<th>Ovid medline</th>
<th>Pubmed</th>
<th>Cochrane</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confocal microscopy</td>
<td>11769</td>
<td>9221</td>
<td>57839</td>
<td>0</td>
</tr>
<tr>
<td>Confocal laser scanning microscopy</td>
<td>2010</td>
<td>1436</td>
<td>57839</td>
<td>0</td>
</tr>
<tr>
<td>Reflectance confocal microscopy</td>
<td>102</td>
<td>146</td>
<td>607</td>
<td>0</td>
</tr>
<tr>
<td>Confocal microscopy and skin cancer</td>
<td>34</td>
<td>50</td>
<td>461</td>
<td>0</td>
</tr>
<tr>
<td>Confocal microscopy and histology</td>
<td>500</td>
<td>399</td>
<td>32157</td>
<td>0</td>
</tr>
<tr>
<td>Confocal microscopy and skin cancer and histology</td>
<td>3</td>
<td>4</td>
<td>358</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: The number of hits obtained from each database when searching for articles using key words
Table 2: Tabulated form of variables collected from the included studies

<table>
<thead>
<tr>
<th>Ser. No</th>
<th>Study Author, Year</th>
<th>Study Type</th>
<th>Evidence Level</th>
<th>Type of Technique Used</th>
<th>Number of Skin Cancer Lesions</th>
<th>Total Number Of Lesions*</th>
<th>Reference Standard</th>
<th>Blinding</th>
<th>Number Of Patients</th>
<th>Patient Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nori S et al., 2004 [21]</td>
<td>Retrospective</td>
<td>Ib</td>
<td>In vivo CM</td>
<td>83</td>
<td>137</td>
<td>Histology</td>
<td>yes</td>
<td>145</td>
<td>Lesions 98 male specimens, 54 female specimen</td>
</tr>
<tr>
<td>2</td>
<td>Horn M et al., 2006 [26]</td>
<td>Retrospective</td>
<td>II</td>
<td>Ex vivo CM</td>
<td>20</td>
<td>40</td>
<td>Histology</td>
<td>yes</td>
<td>20</td>
<td>Not described</td>
</tr>
<tr>
<td>3</td>
<td>Gerger A et al., 2005 [27]</td>
<td>Prospective</td>
<td>II</td>
<td>In vivo CM</td>
<td>27</td>
<td>57</td>
<td>Histology</td>
<td>yes</td>
<td>57</td>
<td>Gender and age not mentioned. Location of lesions given as 15 on back, 6 on legs, 4 on head and 1 on arm skin</td>
</tr>
<tr>
<td>4</td>
<td>Langley R et al., 2007 [28]</td>
<td>Prospective</td>
<td>II</td>
<td>In vivo CM</td>
<td>37</td>
<td>125</td>
<td>Histology</td>
<td>no</td>
<td>125</td>
<td>Age range 16-84 yrs, mean age 44.2, no gender statistics</td>
</tr>
<tr>
<td>5</td>
<td>Guitera P et al., 2010 [29]</td>
<td>Retrospective</td>
<td>II</td>
<td>In vivo CM</td>
<td>29</td>
<td>73</td>
<td>Histology</td>
<td>yes</td>
<td>73</td>
<td>Not described for test set</td>
</tr>
<tr>
<td>6</td>
<td>Curchin C et al., 2011 [30]</td>
<td>Retrospective</td>
<td>Ib</td>
<td>In vivo CM</td>
<td>13 MM 9 BCC</td>
<td>50</td>
<td>Histology</td>
<td>yes</td>
<td>42</td>
<td>Not described</td>
</tr>
<tr>
<td>7</td>
<td>Longo C et al., 2013 [31]</td>
<td>Retrospective</td>
<td>II</td>
<td>In vivo CM</td>
<td>32 MM 28 BCC 6 SCC</td>
<td>140</td>
<td>Histology</td>
<td>yes</td>
<td>140</td>
<td>64 male, 76 female</td>
</tr>
<tr>
<td>8</td>
<td>Guitera P et al., 2012 [31]</td>
<td>Retrospective</td>
<td>II</td>
<td>In vivo CM</td>
<td>105 MM 52 BCC</td>
<td>356</td>
<td>Histology</td>
<td>yes</td>
<td>Not Available</td>
<td>Total 663 patients, 309 females and 354 males. Median age 53 yrs. But patient characteristics not described separately for test set</td>
</tr>
</tbody>
</table>

*both cancerous and non cancerous

Results

Methodological quality

Eight studies fulfilling majority, which is 13 or more, of the criteria in STARD checklist [24] were included in the systematic review. In a study by Nori S et al. sensitivity and specificity were calculated by using 1-5 criteria’s of CLSM [21]. We used sensitivity and specificity values for 4 or more criteria’s from this study as they give optimum sensitivity and specificity.

Included studies

All studies were published as original articles in the English language and were conducted at tertiary care settings in the Europe, North America and Australia. Demography and study characteristics are shown in Table 1. Average sensitivity, specificity, positive and negative predictive values from these studies is shown in Table 2. All selected studies had CLSM done on either a single or more than one type of skin cancer and normal peri-lesional skin [26] and/or benign lesions [21,27-32]. The skin cancers included melanoma, BCC and SCC and the benign lesions included seborrheic keratosis [21,30], benign melanocytic lesions[21,27-32], actinic keratosis [21,29,31,32], sebaceous hyperplasia[21], acne[21], rosacea[21], wart[21], molluscum contagiosum[21], cutaneous T cell lymphoma[21], eczema[21], dermatofibroma[21,30,31], tinea[21], vascular lesions[30] and Bowen's disease[31] etc. Five studies [21,26-29] were done on one type of skin cancer. In 3 studies more than one skin cancer was included in the study population [30-32]. The study by Longo C et al.[30] included all 3 types of skin cancer i.e. malignant melanoma, BCC and SCC. Guitera P et al. [31] devised MM and BCC algorithms. Studies by Curchin C et al. [32] and Guitera P et al. [31] grouped lesions of SCC together with pre malignant condition, actinic keratosis, so the results were not used in meta-analysis.

All the included studies compared CLSM results with histopathology, which is considered the "gold standard". In the studies [21,27] where histopathology was not done on all the lesions only the histopathologically confirmed lesions were included in the meta-analysis. In order to carry out meta-analysis the studies were split into three groups depending on the type of cancer that CLSM was used to detect. There were only two studies for the diagnosis of SCC [26,30], four and six studies for BCC [21,30-32] and melanoma [27-32] respectively. The results of the studies for BCC and melanoma were pooled separately, as they were of clinically similar patients and methodologies. The two included studies on SCC only had a total of 26 lesions; as the data is scanty meta-analysis was not done. However, SCC has been included in the Tables 2 and 3. Results of meta-analysis are shown in Table 4 and Figures 1-6.

Sensitivity

Sensitivity is the proportion of positives among people with disease. The pooled sensitivity for melanoma diagnosis was 91.4% [95% Confidence Interval (CI), 87.1%-94.6%] and for BCC diagnosis was 90.1% [95% CI, 84.6%-94.1%] (figure 1 a&b).

Specificity

Specificity can be defined as proportion of negatives among people without disease. Pooled specificities were 79.9% [95% CI, 76.4%-83.2%] and 92.6% [95% CI, 89.9% - 94.7%] for malignant melanoma and BCC respectively (figure 2 a&b).

Likelihood ratios

The positive likelihood ratio (LR) indicates how many times more likely a particular test result is, in subjects with the disease compared to those without the disease. The higher the
Figure 1: Sensitivity of CLSM in the diagnosis of (a) Malignant Melanoma and (b) Basal Cell Carcinoma.

Figure 2: Specificity of CLSM in the diagnosis of (a) Malignant Melanoma and (b) Basal Cell Carcinoma.

Figure 3: Positive likelihood ratio of CLSM in the diagnosis of (a) Malignant Melanoma and (b) Basal Cell Carcinoma.

Figure 4: Negative likelihood ratio of CLSM in the diagnosis of (a) Malignant Melanoma and (b) Basal Cell Carcinoma.

Figure 5: Diagnostic odds ratio of CLSM in the diagnosis of (a) Malignant Melanoma and (b) Basal Cell Carcinoma.

Figure 6: SROC curve of CLSM for diagnosing (a) Malignant Melanoma and (b) Basal Cell Carcinoma.
positive LR the test the more indicative of disease it is. It is considered a good diagnostic test if the value is 10 or above [36]. Pooled positive LR for BCC was 25.18 and for melanoma it was 5.94 (Figure 3a & Figure 3b). The negative LR is a good indicator for a tests ability to rule out a diagnosis. A good diagnostic test has a value of <0.1 and as shown in the figures 4 a&b, CLSM has a negative LR of 0.11 and 0.09 for melanoma and BCC respectively.

**Diagnostic Odds Ratio**

DOR expresses how much greater the odds of having the disease are for the people with a positive test result than for the people with a negative test result. DOR is a convenient measure when combining studies in a systematic review. The pooled DOR can range from <1 to infinity with the higher results showing a higher discriminatory test [37]. DOR for melanoma and BCC were 80.08 and 358.05 respectively (Figure 5a & Figure 5b).

**Summary Receiver Operating Characteristic (SROC) Curve**

The shape of a ROC curve and the area under the curve (AUC) helps us estimate how high the discriminative power of a test is. The closer the curve is located to upper-left hand corner and the larger the area under the curve, the better the test is at discriminating between diseased and non-diseased. The area under the curve can have any value between 0 and 1 and it is a good indicator of the goodness of the test. A perfect diagnostic test has an AUC 1.0. Separate SROC curves were plotted for the studies on malignant melanoma and BCC (Figure 6a & Figure 6b). Each point in the figures represents a different study and the size of the circle reflects the number of lesions included in the study. CLSM has an AUC of 0.97 for melanoma and 0.99 for BCC which indicates an excellent diagnostic accuracy of the instrument.

**Discussion**

Skin cancer is the most prevalent cancer in the light-skinned population. Histopathology remains the gold standard for the diagnosis till date. Over a period of time there have been various diagnostic tests and technologies developed and tried [38]. Reflectance CM or CLSM has the highest resolution of all optical techniques used in skin cancer diagnosis research [39].

In the studies included in meta-analysis the authors have used different pre defined scoring system or system of diagnosis. Gerger et al. [27] validated CLSM in diagnosing melanoma using the diagnostic morphologic features previously described by Busam et al. [20,40,41] and Langley et al. [42]. Langley et al. [28] used the criteria to establish a CLSM diagnosis of melanoma and naevi they had described previously in their initial series [42]. Curchin C et al. [32] followed the algorithms developed by Pellacani et al. [43] for melanoma and by Guitera P et al. [29] for lentigo maligna. Guitera P et al. [29] worked on lentigo maligna (LM) and identified which CM features mentioned by Pellacani et al. [44] can distinguish LM from benign macules. Their study showed six features were independently correlated with malignancy by means of discriminant analysis and they termed them as LM Score. Longo C et al. [30] used 36 RCM criteria according to distinct diagnosis including melanoma, BCC, SCC, seborrhoeic keratosis, dermatofibroma, vascular lesions and xanthogranuloma. Significant RCM features for nodular melanoma and metastatic nodules were pagetoid distribution, cytological atypia and dermal nesting. Guitera P et al. [31] identified 7 independently significant features for the diagnosis of MM. They include cerebriform nests, atypical cobblestone with small nucleated cells, marked cytologic atypia, pagode cells, epidermal disarray, large interpapillary space and dense nest. Interestingly, pagetoid cells should be viewed with caution as work by Hashemi P [45] emphasizes that intraepidermal Langerhans cells are difficult to differenti-
ate from melanocytes under Reflectance Confocal Microscopy (RCM) and their presence may simulate pagetoid spread of melanocytes on RCM images.

A model for the diagnosis of Basal Cell Carcinoma (BCC) with RCM described by Nori et al. [21] is based on the RCM features of an early version of the microscope, which is not in current widespread use. Segura S et al. [46,47] defined a two-step method for diagnosis, whereby at first, melanocytic lesions were distinguished from non-melanocytic lesions, and second, criteria were then applied to differentiate melanomas (MMs) from naevi. Although the authors described RCM features of BCCs, based on the analysis of 26 cases, the typical clinical cases of BCCs were not excised, and hence, the diagnosis of each case was not confirmed pathologically and could not be included in the review. Guitera P et al. [31] defined a model for accurately diagnosing BCCs studying 119 lesions. Their work represents the largest such series reported to date. The significant features distinguishing BCC from other lesions include polarized honeycomb appearance, linear telangiectasia-like horizontal vessels, basaloid cord or nodule, epidermal shadow, convoluted glomerular-like vessels, non-visible papillae, crenbriform nests, disarray of the epidermal layer. Longo C et al. [30] identified cauliflower architecture, BCC islands (dark silhouettes / tightly packed basaloid islands), bright filaments within tumour islands and collagen as significant features.

Because of the paucity of cases of SCC the results should be viewed with caution. However, work by Horn M et al. [26] and Longo C et al. [30] have provided the foundation for the future researchers. Morphological CLSM features of SCC were assessed according to standard criteria used in conventional histopathology by Horn M et al. [26]. The most frequent finding in SCC found by Longo C et al. [30] was the presence of a disarrayed epidermis, present in five out of six cases (83.3%).

Amongst non-invasive diagnostic methods for skin cancers only dermoscopy is in regular clinical use and has been shown to have pooled sensitivity of 0.88 and specificity of 0.86 and pooled diagnostic odds ratio of 51.5 for the diagnosis of melanoma [5]. Although used for clinical diagnosis of BCC, dermoscopy has not been evaluated systematically for this purpose. As compared to naked eye examination, dermoscopy has ability to see through the stratum corneum and outer layers of the epidermis. The advantage of dermoscopy is that it is a convenient hand held instrument. It is a simple and quick method which takes up to a minute to be performed, while CLSM takes considerably more time about 5-15 minutes for a single lesion [9].

In this review the pooled sensitivities, specificities and DORs of melanoma and BCC show the superiority of CLSM over dermoscopy on accurate diagnosis of these two types of skin cancers. Results for BCC are more impressive than for melanoma. This relies on the researches done by Longo C et al. [30] and Guitera P et al. [31], published in 2013 and 2012 respectively both of whom have shown 100% sensitivity of CLSM for BCC diagnosis. Curchin et al. [32] found CLSM to be 100% specific for BCC. DOR of 358.05 for BCC is quite remarkable and so is DOR of 80.08 for melanoma. The results of the review, thus, show that CLSM has the potential to be an aid in the diagnosis of skin cancers. Whereas treatment has hitherto been almost synonymous with surgery, new noninvasive treatment strategies call for noninvasive diagnostics. CLSM stands out among the new armamentarium of dermatologists because of its high resolution providing a "quasi histologic" skin examination at the bedside [40,41]. CLSM is limited by the depth of imaging restricted to the papillary dermis. Processes in the reticular dermis and tumour invasion depth cannot be evaluated reliably. Pathologic changes such as stratum corneum disruption, parakeratosis, ulceration and spongiosis may all further increase spherical aberration and laser light scattering and, therefore, reduce the resolution and maximum depth of imaging [19,30].

Stevenson et al. [48] have conducted a systematic review on the diagnostic accuracy of CLSM for melanoma diagnosis recently and they found CLSM 93% (95% CI 89 – 96) sensitive and 76% (95% CI 68 - 83) specific for melanoma diagnosis indicating the potential of the instrument to be used as an add-on test for melanoma diagnosis. Our review shows slightly higher specificity as it includes the 2013 study by Longo C et al. [30] which found the instrument 100% specific. CLSM has impressive diagnostic performance, however, for the instrument to be routinely used CLSM needs to be smaller in size, less costly and be an operator friendly instrument.

Conclusion
CLSM enables in vivo and ex vivo imaging of human skin at a quasi-histologic resolution. Our review has shown that CLSM has very good diagnostic accuracy for BCC and melanoma based on limited number of studies. CLSM has the potential to be a useful diagnostic aid for skin cancers in the clinical settings.

Larger better designed studies should evaluate this potential further and also compare that with dermoscopy.

Limitations of this review
1. Although there is a good deal of literature on confocal microscopy use for the skin cancer diagnosis, only a small number of good quality studies have been done.
2. All studies were done at tertiary referral centers and not in the community.

References


25) NICE Diagnostics Assessment programme manual.


33) Methods - Chapter 7: Reviewing and grading the evidence (revised).

34) J2: Example of an evidence table for studies of diagnostic test accuracy.


