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

[Link to publication on Research at Birmingham portal](#)

Systematic review of diagnostic accuracy of reflectance confocal microscopy for melanoma diagnosis in patients with clinically equivocal skin lesions

Alexander D. Stevenson[1], Sharon Mickan[2], Susan Mallett[3], Mekhala Ayya[1]

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, **Oxfo**rd, England, United Kingdom

ABSTRACT

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 microscope*, 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:

Type of study

Cohort studies of diagnostic test accuracy with a predefined threshold that was established on separate data are eligible for inclusion.

Target condition

Melanoma of the skin.

Study population

Patients presenting with lesions suspicious for melanoma that were equivocal to clinical and dermoscopic diagnosis. No restriction was placed upon participant characteristics such as age, sex, ethnicity etc.

Index test

Reflectance confocal microscopy. There was no restriction on the type of algorithm or diagnostic process.

Reference standard

Histopathology of the excised skin lesion or long-term clinical follow-up.

Data extraction and management

Per lesion data was extracted onto a study specific data extraction sheet by two authors independently. The following data was collected: the details of the study population, details of the reference standard and index test, blinding of the reference standard and the index test. Prevalence of melanoma, information to complete the 2 x 2 table.

Statistical analysis and data synthesis

Data were extracted by two reviewers independently. Hierarchical bivariate random-effects meta-analysis [19] was used to perform the statistical meta-analysis as this has been demonstrated to be the most robust method [19].

If there appeared to be no or minimal threshold differences between the studies clinically or on the receiver operator characteristic (ROC) plot then a summary statistic in the form of sensitivity and specificity was planned [20]. If there were, clinically and visually, the appearance of a threshold effect then the summary ROC curve was planned as the most appropriate summary measure [20].

If a study presented several sensitivity and specificity estimates on a receiver operator characteristic curve (ROC) 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.

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 textural 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].

This resulted in a low risk of bias regarding the applicability of the included patients and the appropriateness of the index test. In this study, the reporting of patient selection was generally poor however all domains were graded as low risk of bias. The methodological quality assessment is shown graphically in Table 2.

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-
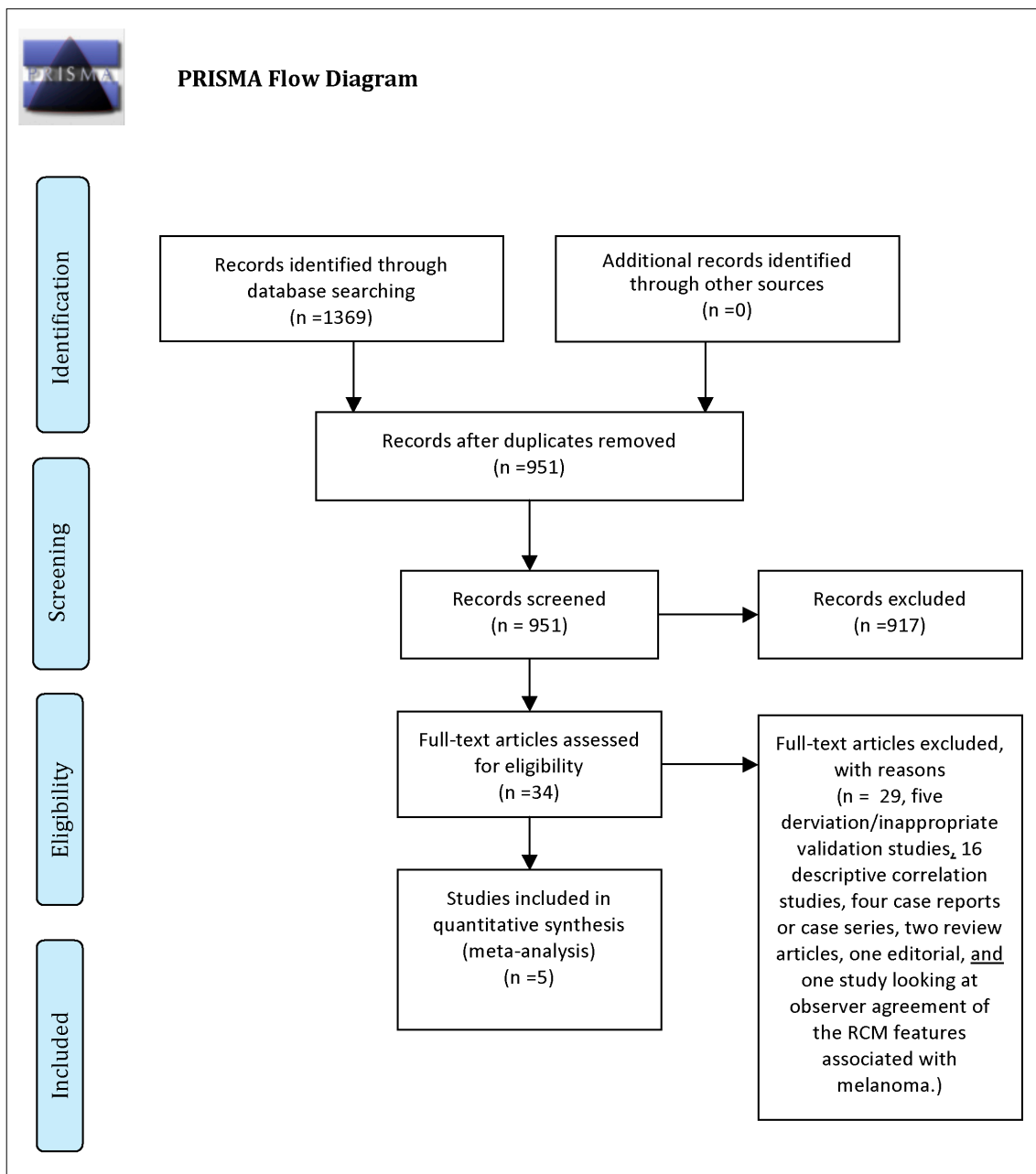
PRISMA Flow Diagram

Identification

Records identified through database searching
(n =1369)

Additional records identified through other sources
(n =0)

Records after duplicates removed
(n =951)

Screening

Records screened
(n = 951)

Records excluded
(n =917)

Eligibility

Full-text articles assessed for eligibility
(n =34)

Full-text articles excluded, with reasons
(n = 29, five derviation/inappropriate validation studies, 16 descriptive correlation studies, four case reports or case series, two review articles, one editorial, and one study looking at observer agreement of the RCM features associated with melanoma.)

Included

Studies included in quantitative synthesis (meta-analysis)
(n =5)

Figure 1. PRISMA flow diagram. [Copyright: ©2013 Stevenson et al.]

ing system developed by Pellacani 2005[26] (Curchin 2011, Guitera 2009, Pellacani 2007), two used a scoring system for lentigo maligna developed by Guitera 2010[27] (Guitera 2010, Curchin 2011) and one did not use a specific diagnostic algorithm but made RCM diagnoses based upon pre-specified melanoma associated features (Langley 2007). The operators were self identified as experienced in four out of the five studies and inexperienced in one (Curchin 2011). There were no explicit differences in the spectrum of disease of patients being examined with RCM. All studies examined equivocal skin lesions. One study was exclusively limited to equivocal skin lesions on the face (Guitera 2010) and two studies excluded lesions on the face (Pellacani 2007 Guitera 2009).

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%].

TABLE 1. Characteristics of included studies

First author and year	Country of study	Algorithm	Patient Characteristics	Number of lesions	Reference standard	Number of patients	Melanoma frequency
Curchin 2011[17]	Australia	Pellacani 2005, Guitera 2010	Data not available	35	Histology	Not available	37%
Guitera 2009[15]	Italy and Australia	Pellacani 2005	Median age and IQR 47(36-60). F:M 149:177	326	Histology	326	37%
Guitera 2010[27]	Australia Italy and USA	Guitera 2010	Data not available	73	Histology/ long term follow up	Not available	37%
Pellacani 2007[14]	Italy and Australia	Pellacani 2005	Median age and IQR 47.7(35.9-60.4). F:M 158:174	351	Histology	332	37%
Langley 2007[33]	Canada	No named diagnostic algorithm	Average age: 44 range: 16-84	125	Histology	125	29%

[Copyright: ©2013 Stevenson et al.]

TABLE 2. QUADAS-2 Risk of bias assessment.

Study	Risk Of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow And Timing	Patient Selection	Index Test	Reference Standard
Curchin 2011	☺	☺	☺	☺	☺	☺	☺
Guitera 2009	☺	☺	☺	☺	☺	☺	☺
Guitera 2010	?	☺	☺	☺	☺	☺	☺
Pellacani 2007	?	☺	☺	☺	☺	☺	☺
Langley 2007	☺	☺	☺	☺	☺	☺	☺

☺ Low Risk   ☻ High Risk   ? Unclear Risk

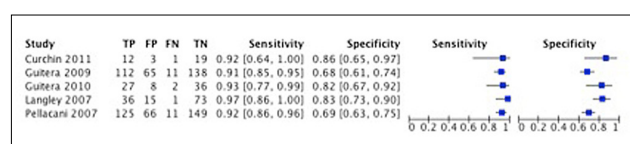[Copyright: ©2013 Stevenson et al.]

Figure 2. Forest plot of the studies. [Copyright: ©2013 Stevenson et al.]

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
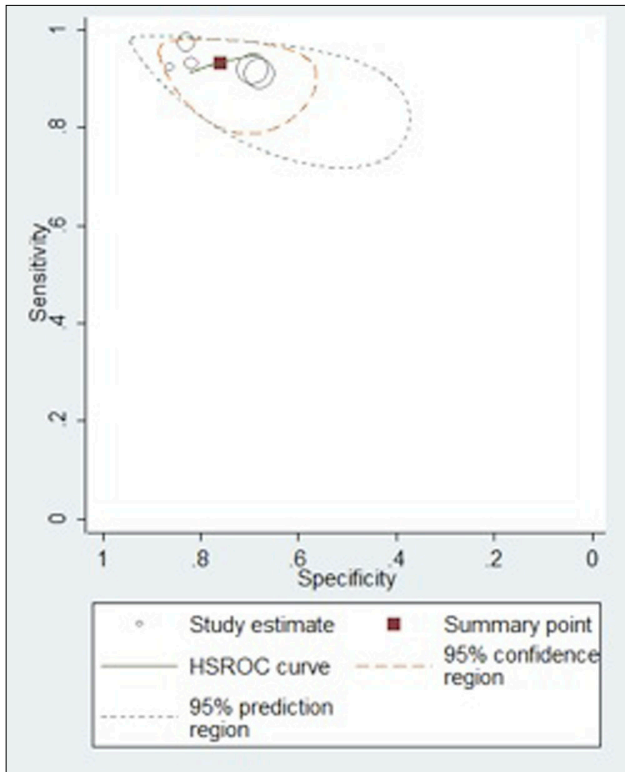
Figure 3. Hierarchical summary receiver operator characteristic curve. [Copyright: ©2013 Stevenson et al.]

with dermoscopy and RCM then this is no doubt the case. This, however, is not helpful for clinical practice. It takes seconds to examine a lesion with dermoscopy and minutes to examine a lesion with RCM. RCM is not going to take on the role of dermoscopy. Therefore it is not useful to compare RCM to the sensitivity and specificity of dermoscopy. Instead, it should be considered as an add-on test to the best current clinical diagnostic tool, which in this case is dermoscopy.

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
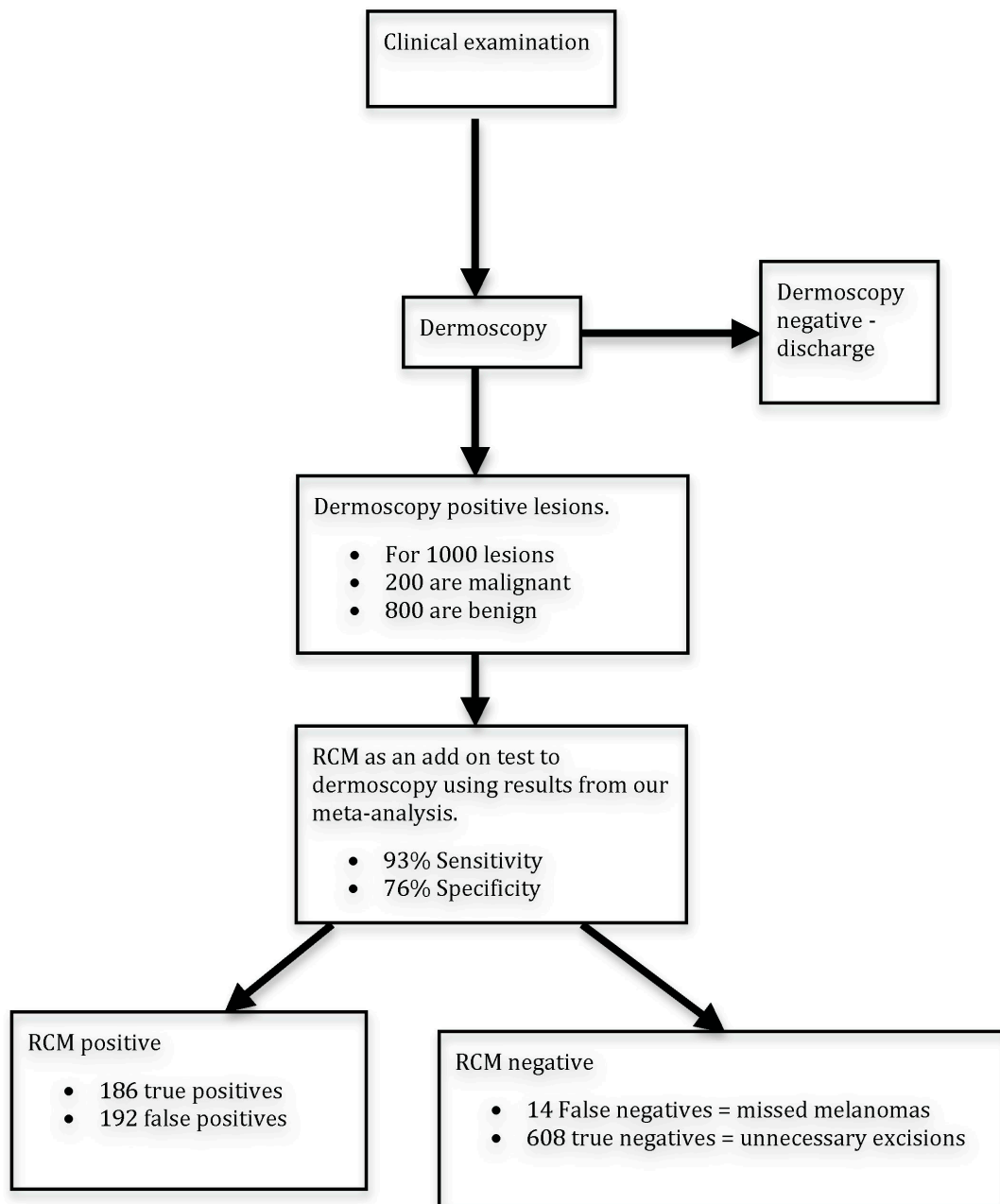
Figure 4. Proposed role of RCM in diagnostic pathway: Hypothetical example based on 1000 lesions positive with dermoscopy. [Copyright: ©2013 Stevenson et al.]

of missing a melanoma and the harms averted by avoiding performing un-necessary excisions.

Excision of skin lesions on most areas of the body is often a quick and easy process that does not carry a great risk of morbidity. The situations where this is not the case are when lesions are on cosmetically sensitive areas of the body such as the face, head and neck or where skin surgery becomes complex, involving the use of skin grafts or flaps. It is these lesions where the reduction in benign lesion excision would have the most impact.

The algorithms that have been developed for use in melanoma diagnosis are based upon several features observed in the images obtained. A study looking at the agreement between observers in identifying these features found high overall levels of reproducibility [16].

Weaknesses

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

Pascale Guitera. In addition the small number of studies and poor reporting in the primary studies limited the scope of the review.

Summary

Reflectance confocal microscopy may contribute to the diagnosis of melanoma as an add-on test in the diagnostic pathway to reduce over-diagnosis following dermoscopy. Reduction in the excision rate of benign lesions that look suspicious on clinical examination may be important particularly where treatment by removal is potentially difficult or harmful. As no diagnostic test is 100% accurate, each clinician and patient will have to decide if the trade off between missing a small number of melanomas is worth the reduction in excision of benign lesions.

References

1. de Vries E, Coebergh JW. Cutaneous malignant melanoma in Europe. Eur J Cancer. 2004;40(16):2355-66.
2. UK CR. Melanoma statistics and outlook. [29/09/2012]; Available from: http://cancerhelp.cancerresearchuk.org/type/melanoma/treatment/melanoma-statistics-and-outlook.
3. Lucas R, McMichael T, Smith W, Armstrong BK. Solar ultraviolet radiation: global burden of disease from solar ultraviolet radiation: World Health Organization; 2006.
4. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol. 2009;27(36):6199-206.
5. Sladden MJ, Balch C, Barzilai DA, et al. Surgical excision margins for primary cutaneous melanoma. Cochrane Database Syst Rev. 2009(4):CD004835.
6. Wheatley K, Ives N, Hancock B, et al. Does adjuvant interferon-alpha for high-risk melanoma provide a worthwhile benefit? A meta-analysis of the randomised trials. Cancer Treat Rev. 2003;29(4):241-52.
7. Marsden JR, Newton-Bishop JA, Burrows L, et al. Revised U.K. guidelines for the management of cutaneous melanoma 2010. Br J Dermatol. 2010;163(2):238-56.
8. Ives NJ, Stowe RL, Lorigan P, Wheatley K. Chemotherapy compared with biochemotherapy for the treatment of metastatic melanoma: a meta-analysis of 18 trials involving 2,621 patients. J Clin Oncol. 2007;25(34):5426-34.
9. Menzies SW, Zalaudek I. Why perform dermoscopy? The evidence for its role in the routine management of pigmented skin lesions. Arch Dermatol. 2006;142(9):1211-2.
10. Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/artificial intelligence for the diagnosis of melanoma. Br J Dermatol. 2009;161(3):591-604.
11. de Giorgi V, Trez E, Salvini C, et al. Dermoscopy in black people. Br J Dermatol. 2006;155(4):695-9.
12. Rajadhyaksha M, Grossman M, Esterowitz D, Webb RH, Anderson RR. In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. J Invest Dermatol. 1995;104(6):946-52.
13. Gerger A, Hofmann-Wellenhof R, Samonigg H, Smolle J. In vivo confocal laser scanning microscopy in the diagnosis of melanocytic skin tumours. Br J Dermatol. 2009;160(3):475-81.
14. Pellacani G, Guitera P, Longo C, et al. The impact of in vivo reflectance confocal microscopy for the diagnostic accuracy of melanoma and equivocal melanocytic lesions. J Invest Dermatol. 2007;127(12):2759-65.
15. Guitera P, Pellacani G, Longo C, et al. In vivo reflectance confocal microscopy enhances secondary evaluation of melanocytic lesions. J Invest Dermatol. 2009;129(1):131-8.
16. Pellacani G, Vinceti M, Bassoli S, et al. Reflectance confocal microscopy and features of melanocytic lesions: An internet-based study of the reproducibility of terminology. Arch Dermatol. 2009;145(10):1137-43.
17. Curchin CES, Wurm EMT, Lambie DLJ, et al. First experiences using reflectance confocal microscopy on equivocal skin lesions in Queensland. Australas J Dermatol. 2011;52(2):89-97.
18. Egger M, Smith GD, O'Rourke K. Rationale, potentials, and promise of systematic reviews. In: Egger M, Smith GD, Altman DG (eds.). *Systematic Reviews In Health Care: Meta-analysis In Context*. 2nd ed. London: BMJ Books, 2001.
19. Harbord RM, Whiting P, Sterne JA, et al. An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. J Clin Epidemiol. 2008;61(11):1095-103.
20. Macaskill P Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ Bossuyt PM, Gatsonis C (eds.). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. The Cochrane Collaboration, 2010.
21. Copenhagen: The Nordic Cochrane Centre TCC. Review Manager (RevMan). 5.1 ed, 2011.
22. Harbord R. METANDI: Stata module to perform meta-analysis of diagnostic accuracy. Statistical Software Components S456932, Boston College Department of Economics, 2008.
23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Int Med. 2011;155(8):529-36.
24. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282(11):1061-6.
25. Buntinx F, Aertgeerts B, Aerts M, et al. Chapter 8. Multivariable analysis in diagnostic accuracy studies: what are the possibilities? In: Knottnerus JA, Buntinx F (eds.). *The Evidence Base of Clinical Diagnosis: Theory and Methods of Diagnostic Research*. 2nd ed. Oxford: Wiley-Blackwell, 2009.
26. Pellacani G, Cesinaro AM, Seidenari S. Reflectance-mode confocal microscopy of pigmented skin lesions—improvement in melanoma diagnostic specificity. J Am Acad Dermatol. 2005;53(6):979-85.
27. Guitera P, Pellacani G, Crotty KA, et al. The impact of in vivo reflectance confocal microscopy on the diagnostic accuracy of lentigo maligna and equivocal pigmented and nonpigmented macules of the face. J Invest Dermatol. 2010;130(8):2080-91.
28. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? Ann Int Med. 2006;144(11):850-5.
29. Duff CG, Melsom D, Rigby HS, Kenealy JM, Townsend PL. A 6 year prospective analysis of the diagnosis of malignant melanoma in a pigmented-lesion clinic: even the experts miss malignant melanomas, but not often. Br J Plas Surg. 2001;54(4):317-21.

30. Rampen FH, Casparie-van Velsen JI, van Huystee BE, Kiemeney LA, Schouten LJ. False-negative findings in skin cancer and melanoma screening. J Am Acad Dermatol. 1995;33(1):59-63.

31. Gur D, Bandos AI, Cohen CS, et al. The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. Radiology. 2008;249(1):47-53.

32. Carli P, De Giorgi V, Crocetti E, et al. Improvement of malignant/benign ratio in excised melanocytic lesions in the 'dermoscopy era': a retrospective study 1997-2001. Br J Dermatol. 2004;150(4):687-92.

33. Langley RGB, Walsh N, Sutherland AE, et al. The diagnostic accuracy of in vivo confocal scanning laser microscopy compared to dermoscopy of benign and malignant melanocytic lesions: A prospective study. Dermatology. 2007;215(4):365-72.