Reflectance confocal microscopy for the diagnosis of keratinocyte skin cancers in adults

Review information

Review type: Diagnostic test accuracy

Review number: #165b

Authors

Jacqueline Dinnes[1], Jonathan J Deeks[1], Naomi Chuchu[1], Daniel Saleh[2,3], Susan E Bayliss[1], Yemisi Takwoingi[1], Clare Davenport[1], Lopa Patel[4], Rubeta N Matin[5], Colette O'Sullivan[6], Rakesh Patalay[7], Hywel C Williams[8], Cochrane Skin Cancer Diagnostic Test Accuracy Group[1]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[2]Newcastle Hospitals, Newcastle Hospitals NHS Trust, Royal Victoria Infirmary, Newcastle, UK
[3]School of Clinical Medicine, The University of Queensland, PA-Southside Clinical Unit, Brisbane, Australia
[4]Plastic Surgery, Royal Stoke Hospital, Staffordshire, UK
[5]Department of Dermatology, Churchill Hospital, Oxford, UK
[6]c/o Cochrane Skin Group, The University of Nottingham, Nottingham, UK
[7]Department of Dermatology, Guy's and St Thomas' NHS Foundation Trust, London, UK
[8]Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK

Citation example: Dinnes J, Deeks JJ, Chuchu N, Saleh D, Bayliss SE, Takwoingi Y, Davenport C, Patel L, Matin RN, O'Sullivan C, Patalay R, Williams HC, Cochrane Skin Cancer Diagnostic Test Accuracy Group. Reflectance confocal microscopy for the diagnosis of keratinocyte skin cancers in adults. Cochrane Database of Systematic Reviews , Issue . Art. No.: . DOI: .

Contact person

Jacqueline Dinnes

Institute of Applied Health Research
University of Birmingham
Birmingham
B15 2TT
UK

E-mail: j.dinnes@bham.ac.uk

Dates

Assessed as Up-to-date:	29 August 2016
Date of Search:	29 August 2016
Next Stage Expected:	Not provided
Protocol First Published:	Not specified
Review First Published:	Not specified
Last Citation Issue:	Not specified

What's new

Date	Event	Description

History

Date	Event	Description

Abstract

Background

Early accurate detection of all skin cancer types is important to guide appropriate management and to improve morbidity and survival. Basal cell carcinoma (BCC) is usually a localised skin cancer but with potential to infiltrate and damage surrounding tissue, whereas squamous cell carcinoma (cSCC) and melanoma are higher risk skin cancers with the potential to metastasise and ultimately lead to death. When used in conjunction with clinical or dermoscopic suspicion of malignancy, or both, reflectance confocal microscopy (RCM) may help to identify those eligible for non-surgical treatment without the need for a diagnostic biopsy, particularly in people with suspected BCC. Any potential benefit must be balanced against the risk of any misdiagnoses.

Objectives

1) To determine the diagnostic accuracy of RCM for the detection of BCC, cSCC, or any skin cancer in adults with a) any

suspicious lesion and b) lesions that are difficult to diagnose (equivocal); and 2) to compare its accuracy with that of usual practice (visual inspection or dermoscopy, or both).

Search methods

We undertook a comprehensive search of the following databases from inception up to August 2016: Cochrane Central Register of Controlled Trials; MEDLINE; EMBASE; CINAHL; CPCI; Zetoc; Science Citation Index; US National Institutes of Health Ongoing Trials Register; NIHR Clinical Research Network Portfolio Database; and the World Health Organization International Clinical Trials Registry Platform. We studied reference lists and published systematic review articles.

Selection criteria

Studies of any design that evaluated the accuracy of RCM alone, or RCM in comparison to visual inspection or dermoscopy, or both, in adults with lesions suspicious for skin cancer compared with a reference standard of either histological confirmation or clinical follow-up, or both.

Data collection and analysis

Two review authors independently extracted all data using a standardised data extraction and quality assessment form (based on QUADAS-2). We contacted authors of included studies where information related to the target condition or diagnostic threshold were missing. We estimated summary sensitivities and specificities using the bivariate hierarchical model. For computation of likely numbers of true positive, false positive, false negative, and true negative findings in the 'Summary of findings' tables, summary sensitivity and specificity estimates were applied to lower quartile, median and upper quartiles of the prevalence observed in the study groups. We also investigated the impact of observer experience.

Main results

Ten studies reporting on a total of 11 study cohorts were included. All 11 cohorts reported data for the detection of BCC, including 2037 lesions (464 with BCC); and four cohorts reported data for the detection of cSCC, including 834 lesions (71 with cSCC). Only one study also reported data for the detection of BCC or cSCC using dermoscopy, limiting comparisons between RCM and dermoscopy. Studies were at high or unclear risk of bias across almost all methodological quality domains, and were of high or unclear concern regarding applicability of the evidence. Selective participant recruitment, unclear blinding of the reference test, and exclusions due to image quality or technical difficulties were observed. It is unclear whether studies are representative of populations eligible for testing with RCM, and test interpretation was often undertaken using images, remotely from the patient and the interpreter blinded to clinical information that would normally be available in practice.

Meta-analysis found RCM to be more sensitive but less specific for the detection of BCC in studies of participants with equivocal lesions (sensitivity 94%, 95% CI 79% to 98%; specificity 85%, 95% CI 72% to 92%; n = 3 studies) compared to studies that included any suspicious lesion (sensitivity 76%, 95% CI 45% to 92%; specificity 95%, 95% CI 66% to 99%; n = 4 studies), although confidence intervals were wide. At the median prevalence of disease of 12.5% observed in studies including any suspicious lesion, applying these results to a hypothetical population of 1000 lesions results in 30 BCCs missed with 44 false positive results (lesions misdiagnosed as BCCs). At the median prevalence of disease of 15% observed in studies of equivocal lesions, 9 BCCs would be missed with 128 false positive results in a population of 1000 lesions. Across both sets of studies, up to 15% of these false positive lesions were observed to be melanomas mistaken for BCCs. There was some suggestion of higher sensitivities in studies with more experienced observers. Summary sensitivity and specificity could not be estimated for the detection of cSCC due to paucity of data.

Authors' conclusions

There is insufficient evidence for the use of RCM for the diagnosis of BCC or cSCC in either population group. A possible role for RCM in clinical practice is as a tool to avoid diagnostic biopsies in lesions with a relatively high clinical suspicion of BCC. The potential for, and consequences of, misclassification of other skin cancers such as melanoma as BCCs requires further research. Importantly, data are lacking that compare RCM to standard clinical practice (with or without dermoscopy).

Plain language summary

What is the diagnostic accuracy of reflectance confocal microscopy (RCM) for the detection of basal or squamous cell carcinoma of the skin in adults?

What is the aim of the review?

The aim of this Cochrane Review was to find out how accurate reflectance confocal microscopy (RCM) is on its own or compared to inspection of a skin lesion with the naked eye alone or using a hand-held microscope called dermoscopy for diagnosing two common forms of keratinocyte skin cancer: basal cell carcinoma (BCC) or cutaneous squamous cell carcinoma (cSCC) in adults. Researchers in Cochrane included 10 studies to answer this question.

Why is improving the diagnosis of BCC or cSCC important?

There are a number of different types of skin cancer. BCC and cSCC are usually localised skin cancers. Making the correct diagnosis is important because mistaking one skin cancer for another can lead to the wrong treatment being used or lead to a delay in effective treatment. A missed diagnosis of BCC (known as a false negative result) can result in the missed BCC growing and causing disfigurement. A missed diagnosis of cSCC is more serious as it could spread to other parts of the body. Diagnosing a skin cancer when it is not actually present (a false positive result) may result in unnecessary biopsy or treatment and can cause discomfort and worry to patients.

What was studied in the review?

Microscopic techniques are used by skin cancer specialists to provide a more detailed, magnified examination of suspicious skin lesions than can be achieved using the naked eye alone. Currently, dermoscopy is used by doctors as part of the examination of suspicious skin lesions. RCM is a new microscopic technique to increase the magnification. It is a handheld device or static unit using infrared light that can visualise deeper layers of the skin when compared with dermoscopy. Both techniques are painless procedures, but RCM is more expensive, time consuming, and requires additional specialised training. Dermoscopy can be used by general practitioners whereas RCM is likely to only be used by hospital specialists for people who have been referred with a skin lesion that is suspected to be a skin cancer. We wanted to see if RCM should be used instead of, or as well as, inspection of a skin lesion with the naked eye alone or using dermoscopy in order to diagnose BCC or cSCC. The accuracy of the test was looked at when used on people with any suspicious skin lesion and also in those with skin lesions that were tricky to diagnose.

What are the main results of the review?

The review included 10 studies that included information on 11 groups of people with lesions suspicious for skin cancer. The main results are based on 7 of the 11 sets of data: four in any lesion suspicious for skin cancer and three in particularly difficult to diagnose skin lesions.

For the comparison of RCM versus dermoscopy, four sets of data that included 912 suspicious skin lesions were found. The results suggest that in a group of 1000 people with any suspicious lesion, of whom 125 (12.5%) really do have BCC:

- An estimated 139 will have an RCM result indicating BCC is present.

- Of these, 44 (32%) will not have BCC (false positive results) including one person with a melanoma mistaken for a BCC.

- Of the 861 people with an RCM result indicating that BCC is not present, 30 (3%) will actually have BCC.

The review also included 3 sets of data on people that had 668 particularly difficult to diagnose skin lesions, one comparing RCM to dermoscopy. The results suggest that if RCM was to be used by skin specialists in a group of 1000 people, of whom 150 (15%) really do have BCC:

- An estimated 269 will have an RCM result indicating BCC is present.

- Of these, 128 (48%) will not have a BCC (known as a false positive result), including as many as 19 people with melanomas mistaken for BCCs.

- Of the 732 people with an RCM result indicating that BCC is not present, 9 (1%) will actually have BCC.

There was not enough evidence to determine the accuracy of RCM for the detection of cSCC in either population group.

How reliable are the results of this review?

There was lots of variation in the results of the studies in this review. Poor reporting of study conduct made assessment of the reliability of studies difficult. It is unclear whether studies are representative of populations eligible for testing with RCM, and test interpretation was often undertaken using images, remotely from the patient and the interpreter blinded to clinical information that would normally be available in practice. Only one study compared the accuracy of dermoscopy and RCM. Most studies were conducted by specialist research teams with high levels of training and experience with RCM, meaning that RCM may appear better than it would be when used in everyday practice. Most studies reported diagnosis based on observers' subjective views, which might not be the same for people using the technique in everyday practice. In nine studies, the diagnosis of skin cancer was made by a skin biopsy or by following up those people over time to make sure they remained negative for skin cancer* .This is likely to have been a reliable method for deciding whether patients really had skin cancer. In one study, the absence of skin cancer was made by experts looking at the skin, a method that may be less reliable for deciding whether patients really had skin cancer.

Who do the results of this review apply to?

Five studies were carried out in Europe (61%), and the rest in Asia, Oceania, North America or more than one continent. The average ages of people who took part ranged from 41 to 65 years. The percentage of people with BCC in these studies ranged from 6% to 83% (a middle value of 12% for any suspicious lesion and 15% for difficult to diagnose skin lesions). For studies of RCM used for cSCC, the percentage of people with cSCC ranged between 4% and 13%. In many studies it was not clear what tests people taking part had received before RCM.

What are the implications of this review?

There is not enough good evidence to support the use of RCM for the diagnosis of BCC or cSCC outside of research studies. There is a lot of variation and uncertainty in results and in the ways studies were carried out, reducing the reliability of findings. Using RCM might avoid the need for a diagnostic biopsy in people who see a doctor with a high suspicion of a BCC lesion, but more research is needed to confirm this. Such research should compare RCM to dermoscopy in well-described groups of people with suspicious skin lesions and they must say whether other skin cancers end up being missed or being wrongly classified as BCC.

How up-to-date is this review?

The review authors searched for and used studies published up to August 2016.

*In these studies biopsy or clinical follow up were the reference standards.

Background

Target condition being diagnosed

The commonest skin cancers in Caucasian populations are those arising from keratinocyte cells: basal cell carcinoma (BCC) and cutaneous squamous cell carcinoma (cSCC) (Gordon 2013; Madan 2010). Basal cell carcinoma is by far the most frequent of the two keratinocyte carcinomas, and around one third of people with a BCC will develop at least a second BCC over time (Flohil 2013). In 2003, the World Health Organization estimated that between two and three million 'non-melanoma' skin cancers (of which BCC and cSCC are estimated to account for around 80% and 16% of cases respectively) and 132,000 melanoma skin cancers occur globally each year (WHO 2003).

In this diagnostic test accuracy review we collectively refer to BCC and cSCC using the new preferred and more accurate term of 'keratinocyte carcinoma' (Karimkhani 2015). We define (a) basal cell carcinoma and (b) squamous cell carcinoma as the primary target conditions for this review. We will also examine accuracy for the target condition of (c) any skin cancer, including keratinocyte skin cancer, melanoma or intraepidermal melanocytic variants and any other skin cancer. We have examined the accuracy of reflectance confocal microscopy for the diagnosis of melanoma in another review (Dinnes 2018a) which is one of a series of systematic reviews of diagnostic tests for the diagnosis of keratinocyte skin cancers (Dinnes 2015). A table of acronyms used is provided in Appendix 1.

Basal cell carcinoma

BCC can arise from multiple stem cell populations, including from the bulge and interfollicular epidermis (Grachtchouk 2011). BCC growth is usually localised, but it can infiltrate and damage surrounding tissue, sometimes causing considerable destruction and disfigurement, particularly when located on the face (Figure 1). The four main types of BCC are superficial, nodular, morphoeic, and pigmented. BCCs typically present as slow-growing asymptomatic papules, plaques, or nodules which may bleed or form ulcers that do not heal (Firnhaber 2012). It is often diagnosed incidentally rather than by people presenting with symptoms (Gordon 2013). A systematic review of the worldwide incidence of keratinocyte skin cancers found estimates for BCC generally under 100 per 100,000 across Europe, ranging from around 900 per 100,000 up to 1800 per 100,000 for some parts of Australia (Lomas 2012).

BCC most commonly occurs on sun-exposed areas of the head and neck (McCormack 1997) and are more common in men and in people over the age of 40. A rising incidence of BCC in younger people has been attributed to increased recreational sun exposure (Bath-Hextall 2007a; Gordon 2013; Musah 2013). Other risk factors include Fitzpatrick skin types I and II (Fitzpatrick 1975; Lear 1997; Maia 1995); previous skin cancer history; immunosuppression; arsenic exposure; and genetic predisposition, such as in basal cell naevus (Gorlin) syndrome (Gorlin 2004; Zak-Prelich 2004). Annual incidence is increasing worldwide; Europe has experienced an average increase of 5.5% per year over the last four decades, the USA 2% per year, while estimates for the UK show incidence appears to be increasing more steeply at a rate of an additional 6 / 100,000 persons per year (Lomas 2012). The rising incidence has been explained by an ageing population, changes in the distribution of known risk factors, particularly ultraviolet radiation, and improved detection due to the increased awareness amongst both practitioners and the general population (Verkouteren 2017).

According to National Institute for Health and Care Excellence (NICE) guidance (NICE 2010), low risk BCCs that may be considered for excision are nodular lesions occurring in patients older than 24 years old who are not immunosuppressed and without Gorlin's syndrome. Furthermore, the lesions should be located below the clavicle; should be small (< 1 cm) with well-defined margins; not recurrent following incomplete excision; and not in awkward or highly visible locations (NICE 2010). Superficial BCCs are also typically low risk and may be amenable to medical treatments such as photodynamic therapy or topical chemotherapy (Kelleners-Smeets 2017). Assigning BCCs as low or high risk influences the management options (Batra 2002; Randle 1996).

Advanced locally destructive BCC can arise from long-standing untreated lesions or from a recurrence of aggressive basal cell carcinoma after primary treatment (Lear 2012). Very rarely, BCC metastasises to regional and distant sites resulting in death, especially cases of large neglected lesions in those who are immunosuppressed or those with Gorlin syndrome (McCusker 2014). Rates of metastasis are reported at 0.0028% to 0.55% (Lo 1991), with very poor survival rates.

Squamous cell carcinoma of the skin

Primary cSCC arises from the keratinising cells of the epidermis or its appendages. People with cSCC often present with an ulcer or firm (indurated) papule, plaque, or nodule (Griffin 2016) often with an adherent crust and poorly defined margins (Madan 2010). cSCC can arise in the absence of a precursor lesion or it can develop from pre-existing lesions, such as actinic keratosis or Bowen's disease (considered by some to be cSCC *in situ*) with an estimated annual risk of progression of <1% to 20% (Alam 2001) and 5% respectively (Kao 1986). It remains locally invasive for a variable length of time, but has the potential to spread to the regional lymph nodes or via the bloodstream to distant sites, especially in immunosuppressed individuals (Lansbury 2010). High risk lesions are those arising on the lip or ear, recurrent cSCC, lesions arising on non-exposed sites, scars or chronic ulcers, tumours more than 20mm in diameter and depth of invasion more than 4mm and poor differentiation on pathological examination (Motley 2009). A systematic review of incidence studies found that the highest reported incidence of cSCC in Europe was in Switzerland, at 28.9/100,000 person-years (1997 data), with rates generally lower in Northern European countries (Lomas 2012). Incidence is higher in the USA and Australia, with rates in men of 60/100,000 person-years reported in Alberta, 290/100,000 in Arizona, and 387/100,000 person-years in Australia (Lomas 2012). Based on data from 2000 to 2006,

the annual incidence rates of cSCC in England, Scotland, and Northern Ireland were 22.7 per 100,000, 27.0 per 100,000, and 30.6 per 100,000 person-years, respectively (Lomas 2012).

Chronic ultraviolet light exposure through recreation or occupation is strongly linked to cSCC occurrence (Alam 2001). It is particularly common in people with fair skin and in less common genetic disorders of pigmentation, such as albinism, xeroderma pigmentosum, and recessive dystrophic epidermolysis bullosa (RDEB) (Alam 2001). Other recognised risk factors include immunosuppression; chronic wounds; arsenic or radiation exposure; certain drug treatments, such as voriconazole and BRAF mutation inhibitors; and previous skin cancer history (Baldursson 1993; Chowdri 1996; Dabski 1986; Fasching 1989; Lister 1997; Maloney 1996; O'Gorman 2014). In solid organ transplant recipients, cSCC is the most common form of skin cancer; the risk of developing cSCC has been estimated at 65 to 253 times that of the general population (Hartevelt 1990; Jensen 1999; Lansbury 2010). Overall, local and metastatic recurrence of cSCC at five years is estimated at 8% and 5% respectively. The five-year survival rate of metastatic cSCC of the head and neck is around 60% (Moeckelmann 2018).

Treatment

Treatment options for BCC and cSCC include surgery, other destructive techniques and topical chemotherapy. A Cochrane Review of 27 randomised controlled trials (RCTs) of interventions for BCC found very little good quality evidence for any of the interventions used (Bath-Hextall 2007b). Complete surgical excision of primary BCC has a reported five-year recurrence rate of <2% (Griffiths 2005; Walker 2006), leading to significantly fewer recurrences than treatment with radiotherapy (Bath-Hextall 2007b). Mohs micrographic surgery, whereby horizontal sections of the tumour are microscopically examined and re-excision is undertaken until the margins are tumour-free, can be considered for high risk lesions on the face where standard wider excision margins might lead to considerable functional or cosmetic impairment (Bath-Hextall 2007b; Motley 2009; Lansbury 2010; Stratigos 2015). Bath-Hextall and colleagues (Bath-Hextall 2007b) found a single trial comparing Mohs micrographic surgery with a 3mm surgical margin excision in BCC (Smeets 2004); the update of this study showed non-significantly lower recurrence at 10 years with Mohs micrographic surgery (4.4% compared to 12.2% after surgical excision, P = 0.10) (van Loo 2014).

The main treatments for high risk BCC are excision or Mohs micrographic surgery and radiotherapy. For low risk or superficial subtypes of BCC, or for those with small or multiple BCCs, or both, at low risk sites (Marsden 2010), destructive techniques other than excisional surgery may be used (e.g. electrodessication and curettage or cryotherapy (Alam 2001; Bath-Hextall 2007b)). Alternatively non-surgical (or non-destructive) treatments may be considered, including topical chemotherapy such as imiquimod and 5-fluorouracil and photodynamic therapy (Bath-Hextall 2007b; Kim 2014; Roozeboom 2016; Williams 2017; Drew 2017). Non-surgical treatments are most frequently used for superficial forms of BCC, with one head to head trial suggesting topical imiquimod is superior to PDT and 5-FU (Jansen 2017). Although non-surgical techniques are increasingly used, they do not allow histological confirmation of tumour clearance, and their use is dependent on accurate characterisation of the histological subtype and depth of tumour. The 2007 systematic review of BCC interventions found limited evidence from very small RCTs for these approaches (Bath-Hextall 2007b), which have only partially been filled by subsequent studies (Bath-Hextall 2014; Kim 2014; Roozeboom 2012). Most BCC trials have compared interventions within the same treatment class, and few have compared medical versus surgical treatments (Kim 2014).

Vismodegib, a first-in-class Hedgehog signalling pathway inhibitor is now available for the treatment of metastatic or locally advanced BCC based on the pivotal study ERIVANCE BCC (Sekulic 2012). It is licensed for use in these patients where surgery or radiotherapy is inappropriate, e.g. for treating locally advanced periocular and orbital BCCs with orbital salvage of patients who otherwise would have required exenteration (Wong 2017). However, NICE has recently recommended against the use of vismodegib based on cost-effectiveness and uncertainty of evidence (NICE 2017).

A systematic review of interventions for primary cSCC found only one RCT eligible for inclusion (Lansbury 2010). Current practice therefore relies on evidence from observational studies, as reviewed in Lansbury 2013, for example. Surgical excision with predetermined margins is usually the first-line treatment (Motley 2009; Stratigos 2015). Estimates of recurrence after Mohs micrographic surgery, surgical excision, or radiotherapy, which are likely to have been evaluated in higher risk populations, have shown pooled recurrence rates of 3%, 5.4% and 6.4%, respectively with overlapping confidence intervals; the review authors advise caution when comparing results across treatments (Lansbury 2013).

Index test(s)

Reflectance confocal microscopy (RCM), also known as confocal laser scanning microscopy or confocal microscopy, was first developed for skin imaging in the early 1990s (Rajadhyaksha 1995) and is emerging as a potential alternative or adjunct to dermoscopy for the diagnosis of skin cancer. It is a non-invasive technology, which can be used to visualise horizontally sectioned images of the skin at a cellular lateral resolution of ~1micron, in vivo to the depth of the upper dermis. The contrast for the monochrome images produced is achieved by the variation of the optical properties within the skin when illuminated by a near-infrared light (830nm) (see Figure 2). The greatest contrast is achieved from melanin, so that RCM is advocated as being particularly useful for assessing pigmented lesions.

The Caliber ID VivaScope® imaging systems are the only commercially available RCM devices (distributed by MAVIG in Europe; www.vivascope.de/en/home.html). The Vivascope 1500 (and the previously available 1000 version) is a console based unit with a dermoscopic attachment, whereas the Vivascope 3000 is a handheld device designed for superior ergonomics, allowing imaging of lesions inaccessible for the 1500 version (Figure 3). Imaging can be undertaken by clinicians or technicians following appropriate training (Edwards 2016). The length of time required for

diagnosis has been estimated at 15 minutes for Vivascope 1500 (10 minutes of a technician's time for imaging and 5 minutes of a dermatologists for image interpretation) and 10 minutes for Vivascope 3000 (Edwards 2016). The company has estimated the average cost per use of the 1500 system, including dermoscopy, as £120 based on 2014 NHS reference costs and an indicative price for Vivascope 1500 of £95,224 (Edwards 2016).

Various algorithms have been proposed for the interpretation of RCM images, particularly for the diagnosis of melanoma (Dinnes 2018a); however, evaluation of lesion characteristics associated with other types of skin cancer, especially BCC, is ongoing (Gonzalez 2002; Guitera 2012). The lesion characteristics most recently proposed to be associated with BCC include the presence of 'dark silhouettes' or 'bright tumour islands' plus at least one of: 'streaming' polarization of nuclei in neoplastic aggregates along the same axis of orientation; 'peripheral palisading' of nuclei at the tumour islands' periphery; dark 'peritumoral clefts' around the tumour islands; fibrotic stroma with 'thickened collagen bundles'; dilated and tortuous 'linear blood vessels' and 'coiled blood vessels'; 'bright dendritic structures' within tumour islands; and 'bright round cells' in the stroma. Nevertheless, BCC and cSCC specific criteria have yet to be fully established, with some suggestion that the keratotic surface of SCC may prohibit the use of imaging techniques (Edwards 2016).

Clinical Pathway

The diagnosis of skin lesions occurs in primary, secondary, and tertiary care settings by both generalist and specialist healthcare providers. In the UK, people with concerns about a new or changing lesion will present to their general practitioner rather than directly to a specialist in secondary care. If the general practitioner has concerns then he/she usually refers the patient to a specialist in secondary care – usually a dermatologist but sometimes to a surgical specialist such as a plastic surgeon or an ophthalmic surgeon. Suspicious skin lesions may also be identified in a referral setting, for example by a general surgeon, and referred for a consultation with a skin cancer specialist (Figure 4). Skin cancers identified by other specialist surgeons (such as an ear, nose, and throat (ENT) specialist or maxillofacial surgeon) will usually be diagnosed and treated without further referral.

Current UK guidelines recommend that all suspicious pigmented lesions presenting in primary care should be assessed by taking a clinical history and visual inspection using the seven-point checklist (MacKie 1990); lesions suspected to be melanoma or cSCC (London Cancer Alliance 2013) should be referred for appropriate specialist assessment within two weeks (Chao 2013; Marsden 2010; NICE 2015). Evidence is emerging, however, to suggest that excision of melanoma by general practitioners (GPs) is not associated with increased risk compared with outcomes in secondary care (Murchie 2017). In the UK, low risk BCC are usually recommended for routine referral, with urgent referral for those in whom a delay could have a significant impact on outcomes, for example due to large lesion size or critical site (NICE 2015). Appropriately qualified generalist care providers increasingly undertake management of low-risk BCC in the UK (NICE 2010) and Australia (CCAAC Network 2008).

For referred lesions, the specialist clinician will use history-taking, visual inspection of the lesion (in conjunction with other skin lesions), and palpation of the lesion and associated lymph nodes in conjunction with dermoscopic examination to inform a clinical decision. If melanoma is suspected, then urgent 2mm excision biopsy is recommended (Lederman 1985; Lees 1991); for cSCC, predetermined surgical margin excision or a diagnostic biopsy may be considered. BCC and pre-malignant lesions potentially eligible for nonsurgical treatment may undergo a diagnostic biopsy before initiation of therapy if there is diagnostic uncertainty. Equivocal (or more difficult to diagnose) melanocytic lesions for which a definitive clinical diagnosis cannot be reached may undergo surveillance to identify any lesion changes that would indicate excision biopsy or reassurance and discharge for those that remain stable over a period of time.

Theoretically, teledermatology consultations may aid appropriate triage of lesions into urgent referral; non-urgent secondary care referral (e.g. for suspected basal cell carcinoma); or where available, referral to an intermediate care setting, e.g. clinics run by GPs with a special interest in dermatology. The distinction between setting and examiner qualifications and experience is important as specialist clinicians might work in primary care settings (for example, in the UK, GPs with a special interest in dermatology and skin surgery who have undergone appropriate training), and generalists might practice in secondary care settings (for example, plastic surgeons who do not specialise in skin cancer). The level of skill and experience in skin cancer diagnosis will vary for both generalist and specialist care providers and will also impact on test accuracy.

Prior test(s)

The diagnosis of skin cancer is based on history-taking and clinical examination. In the UK, this is typically undertaken at two decision points – first in the GP surgery where a decision is made to refer or not to refer, and then a second time by a dermatologist or other secondary care clinician where a decision is made to biopsy or excise or not. Visual inspection of the skin is undertaken iteratively, using both implicit pattern recognition (non-analytical reasoning) and more explicit 'rules' based on conscious analytical reasoning (Norman 2009), the balance of which will vary according to experience and familiarity with the diagnostic question. Various attempts have been made to formalise the "mental rules" involved in analytical pattern recognition for melanoma (Friedman 1985; Grob 1998; MacKie 1985; MacKie 1990; Sober 1979; Thomas 1998); however, visual inspection for keratinocyte skin cancers relies primarily on pattern recognition. Accuracy has been shown to vary according to the expertise of the clinician. Primary care physicians have been found to miss over half of BCC (Offidani 2002) and to inappropriately diagnose one third of BCC (Gerbert 2000). In contrast, an Australian study found that trained dermatologists were able to detect 98% of BCC, but with a specificity of only 45% (Green 1988).

A range of technologies have emerged to aid diagnosis to reduce the number of diagnostic biopsies or

inappropriate surgical procedures. Dermoscopy using a hand-held microscope has become the most widely used tool for clinicians to improve diagnostic accuracy of pigmented lesions, in particular for melanoma (Argenziano 1998; Argenziano 2012; Haenssle 2010; Kittler 2002), although is less well established for the diagnosis of BCC or cSCC. The diagnostic accuracy, and comparative accuracy, of visual inspection and dermoscopy for keratinocyte skin cancer has been evaluated in a further review in this series (Dinnes 2018b).

Role of index test(s)

RCM is most likely to have a role as an additional test to better identify lesions that can be monitored or reassured as being benign, instead of being sent for urgent excision (Edwards 2016), or for low risk BCC to identify those eligible for non-surgical treatment without the need for a diagnostic biopsy. RCM could also be considered as a primary diagnostic test, i.e. as a potential replacement for dermoscopy.

Delay in diagnosis of a BCC as a result of a false-negative test is not as serious as for melanoma because BCCs are usually slow-growing and very unlikely to metastasise (Betti 2017). However, delayed diagnosis can result in a larger and more complex excision with consequent greater morbidity. Very sensitive diagnostic tests for BCC however may compromise on lower specificity leading to a higher false-positive rate, and an enormous burden of skin surgery, such that a balance between sensitivity and specificity is needed. The situation for cSCC is more similar to melanoma in that the consequences of falsely reassuring a person that they do not have skin cancer can be serious and potentially fatal. Thus, a good diagnostic test for cSCC should demonstrate high sensitivity and a corresponding high negative predictive value. A test that can reduce false positive clinical diagnoses without missing true cases of disease has patient and resource benefits. False-positive clinical diagnoses not only cause unnecessary morbidity from the biopsy, but could lead to initiation of inappropriate therapies and also increase patient anxiety.

A further postulated advantage of RCM is its ability to non-invasively differentiate seborrhoeic keratoses or non-melanocytic lesions from a population of pigmented lesions (de Carvalho 2015; Nascimento 2014; Menge 2016). RCM could also develop a role in guiding definitive therapeutic margins (Edwards 2016), both pre- and intra-operatively and to estimate response to topical chemotherapy for lentigo maligna and potentially BCCs; however, these uses are not under consideration in this review.

Alternative test(s)

A number of other tests have been reviewed as part of our series of Cochrane DTA reviews on the diagnosis of keratinocyte skin cancers, including visual inspection and dermoscopy (Dinnes 2018b), teledermatology (Chuchu 2018a), mobile phone applications (Chuchu 2018b), computer-assisted diagnosis (CAD) techniques (Ferrante di Ruffano 2018a), optical coherence tomography (OCT) (Ferrante di Ruffano 2018b), exfoliative cytology (Ferrante di Ruffano 2018c) and high frequency ultrasound (Dinnes 2018c).

OCT is an emerging optical imaging technology based on interferometry using a near infra-red light source. It exploits differences in the refractive index in the skin to create vertically sectioned images in vivo, in real time and has a relatively high depth of penetration, allowing dermal lesions to be delineated (Olsen 2015). OCT is considered to be particularly useful for the differentiation of non pigmented lesions. Pigmented lesions produce regular scattering patterns which inhibit the differentiation of malignant from benign lesions (Olsen 2015; Gambichler 2015). The use of high frequency ultrasound has been advocated in diagnosing a range of skin conditions, including skin cancer, infection, and inflammatory conditions (Kleinerman 2012), with malignant lesions reportedly appearing as hypoechogenic areas surrounded by a hyperechogenic dermis. Melanomas in particular also reportedly appear homogenous and with well-defined margins (e.g. Harland 2000). CAD or artificial intelligence-based techniques process and manipulate lesion data using predefined algorithms to identify the features that discriminate malignant from benign lesions (Rajpara 2009; Esteva 2017). These techniques have been incorporated into commercially available handheld devices for ease of use in a clinic setting, including SIAscopy™ (Moncrieff 2002; Walter 2012), MelaFind® (Monheit 2011; Wells 2012; Hauschild 2014), and the Nevisense™ Electrical Impedance Spectroscopy system (Malvehy 2014). CAD has however most commonly been applied to digital dermoscopy images (Rajpara 2009; Esteva 2017).

Evidence permitting, the accuracy of available tests will be compared in an overview review, exploiting within-study comparisons of tests and allowing the analysis and comparison of commonly used diagnostic strategies where tests may be used singly or in combination.

We are not considering the accuracy of histopathological confirmation following lesion excision or biopsy as an index test for these reviews; it is the established reference standard for skin cancer diagnosis and will be one of the standards against which we will evaluate the index tests in these reviews.

Rationale

Our series of reviews of diagnostic tests used to assist the clinical diagnosis of the keratinocyte skin cancers BCC and cSCC, aims to identify the most accurate approaches to diagnosis and provide clinical and policy decision-makers with the highest possible standard of evidence on which to base decisions. With the increasing availability of a wider range of tests, there is a need to differentiate and appropriately triage keratinocyte skin cancers to avoid sending too many people with benign or low risk lesions for a specialist opinion and possible excision or biopsy, whilst not missing those people who have lesions that require treatment.

Although a set of billing codes for the USA have been agreed since January 2016 (Rajadhyaksha 2017), RCM is not recommended for routine use in the UK (Edwards 2016), Australia (Guitera 2017), or New Zealand (Sobarun 2015). To date, the use of RCM has been limited by expense (in terms of both equipment and staff time) and the need for

specialised training. Recent studies have demonstrated high sensitivity and specificity amongst experienced RCM users, however, in at least one study, the accuracy of the group on average was higher than that of any one individual observer (Farnetani 2015). Our own systematic review of 18 studies of RCM for the diagnosis of melanoma suggested that although RCM may augment diagnostic sensitivity when used in conjunction with clinical inspection and dermoscopy, its main contribution is an increase in specificity, reducing the number of individuals receiving unnecessary surgery by up to three quarters compared to dermoscopy (Dinnes 2018a).

Available systematic reviews of RCM for keratinocyte skin cancers are limited by out of date searches and methods. Xiong 2016 failed to consider differences in study populations and varying definitions of the target condition, and used an out of date meta-analytic approach. Mogensen 2007 did not report the use of systematic methods for study inclusion or extraction and did report undertaking quality assessment, while Edwards 2016 focused on selected studies considered to be more applicable to a UK setting. In this rapidly advancing field, there is a need for an up-to-date analysis of the accuracy of RCM for the diagnosis of keratinocyte skin cancer.

This is one of a series of Cochrane Diagnostic Test Accuracy (DTA) reviews on the diagnosis and staging of melanoma and keratinocyte skin cancers as part of the National Institute for Health Research (NIHR) Cochrane Systematic Reviews Programme. Appendix 2 shows the content and structure of the programme. As several reviews for each topic area followed the same methodology, generic protocols were prepared in order to avoid duplication of effort, one for diagnosis of keratinocyte skin cancers (Dinnes 2015) and one for diagnosis of melanoma (Dinnes 2015a). The Background and Methods sections of this review therefore use some text that was originally published in the protocol concerning the evaluation of tests for the diagnosis of keratinocyte skin cancer (Dinnes 2015) and text that overlaps some of our other reviews (Dinnes 2018a).

Objectives

To determine the diagnostic accuracy of reflectance confocal microscopy (RCM) for the detection of BCC in adults, and to compare its accuracy with visual inspection or dermoscopy or both.

To determine the diagnostic accuracy of RCM for the detection of cSCC in adults, and to compare its accuracy with visual inspection or dermoscopy, or both.

Accuracy was estimated separately according to the point in the clinical pathway at which RCM was evaluated:

1. in participants with any suspicious lesion, where RCM might be used as an alternative to dermoscopy or to supplement visual inspection alone
2. in participants with equivocal (or more difficult to diagnose) lesions in whom a clear management decision could not be made following visual inspection and dermoscopy, where RCM might be used as an addition to visual inspection or dermoscopy, or both.

Studies that did not clearly fit into either of these two groups were considered as 'other lesion' studies. The terms equivocal and 'difficult to diagnose' have been used, and should be interpreted, interchangeably throughout this review.

Secondary objectives

To determine the diagnostic accuracy of RCM for the detection of any skin cancer in adults, where keratinocyte skin cancers make up at least 50% of included skin cancers, and to compare its accuracy with visual inspection or dermoscopy, or both.

Accuracy was estimated separately according to the point in the clinical pathway at which RCM is evaluated:

1. where it might be used as an alternative to dermoscopy in participants with any lesion suspicious for melanoma
2. where it might be used as an addition to visual inspection or dermoscopy, or both, in participants with equivocal lesions in whom a clear management decision could not be made following visual inspection and dermoscopy alone

For the detection of BCC or cSCC (the primary target conditions):

i. To compare the accuracy of RCM to dermoscopy where both tests have been evaluated in the same studies (direct test comparisons)
ii. To determine the diagnostic accuracy of individual algorithms for RCM
iii. To determine the effect of observer experience.

Investigation of sources of heterogeneity

We set out to address a range of potential sources of heterogeneity for investigation across our series of reviews, as outlined in our generic protocol (Dinnes 2015) and described in Appendix 3; however, our ability to investigate these was necessarily limited by the available data on each individual test reviewed.

Methods

Criteria for considering studies for this review

Types of studies

We included test accuracy studies that allow comparison of the result of the index test with that of a reference standard, including the following:

- studies where all participants receive a single index test and a reference standard;
- studies where all participants receive more than one index test(s) and reference standard;

- studies where participants are allocated (by any method) to receive different index tests or combinations of index tests and all receive a reference standard (between-person comparative studies (BPC));
- studies that recruit series' of participants unselected by true disease status (referred to as case series for the purposes of this review);
- diagnostic case-control studies that separately recruit diseased and non-diseased groups (see Rutjes 2005).
- both prospective and retrospective studies; and
- studies where previously acquired clinical or dermoscopic images were retrieved and prospectively interpreted for study purposes.

We excluded studies from which we could not extract 2x2 contingency data, or if they included less than five cases of BCC or cSCC or less than five benign lesions.

Participants

We included studies in adults with lesions suspicious for skin cancer.

We excluded studies that recruited only participants with malignant diagnoses and studies that compared test results in participants with malignancy compared with test results based on 'normal' skin as controls, due to the bias inherent in such comparisons (Rutjes 2006).

We excluded studies conducted in children or which clearly reported more than 50% of participants aged 16 and under.

Index tests

Studies evaluating RCM alone, or RCM in comparison to usual practice (visual inspection or dermoscopy, or both) were included.

All established algorithms or checklists to assist diagnosis by RCM were included. Studies developing new algorithms or methods of diagnosis (i.e. derivation studies) were included if they:

- used a separate independent 'test set' of participants or images to evaluate the new approach, or
- investigated lesion characteristics that had previously been suggested as associated with BCC or cSCC and the study reported accuracy based on the presence or absence of particular combinations of characteristics.

Studies were excluded if they:

- used a statistical model to produce a data driven equation, or algorithm based on multiple diagnostic features, with no separate test set.
- used cross-validation approaches such as 'leave-one-out' cross-validation (Efron 1983)
- evaluated the accuracy of the presence or absence of individual lesion characteristics or morphological features, with no overall diagnosis of malignancy.

No exclusions were made according to test observer.

Target conditions

Two primary target conditions were defined as the detection of:

- basal cell carcinoma (BCC), including all subtypes;
- invasive cutaneous squamous cell carcinoma (cSCC) (we did not consider cutaneous SCC in situ or Bowen's disease as disease positive)

An additional definition of the target condition was considered in secondary analysis, the detection of:

- any skin cancer, including BCC, cSCC, melanoma or any rare skin cancer (e.g. Merkel cell cancer), as long as skin cancers other than melanoma made up more than 50% of the disease positive group. Data from studies in which melanoma accounted for more than 50% of skin cancers were included in the review of RCM for the diagnosis of melanoma (Dinnes 2018a).

Reference standards

The ideal reference standard was histopathological diagnosis in all eligible lesions. A qualified pathologist or dermatopathologist should perform histopathology. Ideally, reporting should be standardised detailing a minimum dataset to include the type of skin cancer (BCC, cSCC) and subtype of BCC and may also refer to the TNM (**t**umour, **n**ode, and **m**etastasis) classification of staging for cSCC (Royal College of Pathologists 2014). We did not apply the reporting standard as a necessary inclusion criterion, but extracted any pertinent information.

Partial verification (applying the reference test only to a subset of those undergoing the index test) was of concern given that lesion excision or biopsy are unlikely to be carried out for all benign-appearing lesions within a representative population sample. Therefore, we accepted clinical follow-up of benign-appearing lesions as an eligible reference standard, whilst recognising the risk of differential verification bias (as misclassification rates of histopathology and follow-up will differ).

Additional eligible reference standards included cancer registry follow-up and 'expert opinion' with no histology or clinical follow-up. Cancer registry follow-up is considered less desirable than active clinical follow-up, as it is not carried out within the control of the study investigators. Furthermore, if participant-based analyses as opposed to lesion-based analyses are presented (as for cancer registry follow up), it may be difficult to determine whether the detection of a malignant lesion during follow-up is the same lesion that originally tested negative on the index test.

All of the above were considered eligible reference standards with the following caveats:

- all study participants with a final diagnosis of the target disorder must have a histological diagnosis, either subsequent to the application of the index test or after a period of clinical follow-up, and
- at least 50% of all participants with benign lesions must have either a histological diagnosis or clinical follow-up to confirm benignity.

Search methods for identification of studies

Electronic searches

The Information Specialist (SB) carried out a comprehensive search for published and unpublished studies. A single large literature search was conducted to cover all topics in the programme grant (see Appendix 2 for a summary of reviews included in the programme grant). This allowed for the screening of search results for potentially relevant papers for all reviews at the same time. A search combining disease-related terms with terms related to the test names, using both text words and subject headings was formulated. The search strategy was designed to capture studies evaluating tests for the diagnosis or staging of skin cancer. As the majority of records were related to the searches for tests for staging of disease, a filter using terms related to cancer staging and to accuracy indices was applied to the staging test search, to try to eliminate irrelevant studies, for example, those using imaging tests to assess treatment effectiveness. A sample of 300 records that would be missed by applying this filter was screened and the filter adjusted to include potentially relevant studies. When piloted on MEDLINE, inclusion of the filter for the staging tests reduced the overall numbers by around 6000. The final search strategy, incorporating the filter (Appendix 4), was subsequently applied to all bibliographic databases as listed below. The final search result was cross-checked against the list of studies included in five systematic reviews; our search identified all but one of the studies, and this study is not indexed on MEDLINE. The Information Specialist devised the search strategy, with input from the Information Specialist from Cochrane Skin. No additional limits were used.

We searched the following bibliographic databases to 29 August 2016 for relevant published studies:

- MEDLINE via OVID (from 1946);
- MEDLINE In-Process & Other Non-Indexed Citations via OVID; and
- EMBASE via OVID (from 1980).

We searched the following bibliographic databases to 30 August 2016 for relevant published studies:

- the Cochrane Central Register of Controlled Trials (CENTRAL) Issue 7, 2016, in the Cochrane Library;
- the Cochrane Database of Systematic Reviews (CDSR) Issue 8, 2016 in the Cochrane Library;
- Cochrane Database of Abstracts of Reviews of Effects (DARE) Issue 2, 2015;
- CRD HTA (Health Technology Assessment) database Issue 3, 2016;
- CINAHL (Cumulative Index to Nursing and Allied Health Literature via EBSCO from 1960).

We searched the following databases for relevant unpublished studies:

- CPCI (Conference Proceedings Citation Index) via Web of Science™ (from 1990);
- Zetoc (from 1993)
- SCI Science Citation Index Expanded™ via Web of Science™ (from 1900, using the "Proceedings and Meetings Abstracts" Limit function).

We searched the following trials registers:

- The US National Institutes of Health Ongoing Trials Register (www.clinicaltrials.gov);
- NIHR Clinical Research Network Portfolio Database (http://www.nihr.ac.uk/research-and-impact/nihr-clinical-research-network-portfolio/);
- The World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch/).

We aimed to identify all relevant studies regardless of language or publication status (published, unpublished, in press, or in progress). No date limits were applied.

Searching other resources

We have included information about potentially relevant ongoing studies in the 'Characteristics of ongoing studies' tables. We have screened relevant systematic reviews identified by the searches for their included primary studies, and included any missed by our searches. We have checked the reference lists of all included papers, and subject experts within the author team have reviewed the final list of included studies. No citation searching has been conducted.

Data collection and analysis

Selection of studies

Titles and abstracts were screened by at least one author (JDi or NC), with any queries discussed and resolved by consensus. A pilot screen of 539 MEDLINE references showed good agreement (89% with a kappa of 0.77) between screeners. Primary test accuracy studies and test accuracy reviews (for scanning of reference lists) of any test used to investigate suspected melanoma, BCC, or cSCC were included at initial screening. Inclusion criteria (Appendix 3) were applied independently by both a clinical reviewer (from one of a team of twelve clinician reviewers) and a methodologist reviewer (JDi or NC) to all full text articles, disagreements were resolved by consensus or by a third party (JDe, CD, HW, and RM). Authors of eligible studies were contacted when insufficient data were presented to allow for the construction of 2x2 contingency tables.

Data extraction and management

One clinical (as detailed above) and one methodologist reviewer (JDi, NC or LFR) independently extracted data concerning details of the study design, participants, index test(s) or test combinations and criteria for index test positivity, reference standards, and data required to populate a 2x2 diagnostic contingency table for each index test using a piloted data extraction form. Data were extracted at all available index test thresholds. Disagreements were resolved by consensus or by a third party (JDe, CD, HW, and RM).

Authors of included studies were contacted where information relating to the diagnostic threshold was missing. Authors of conference abstracts published from 2013 to 2015 were contacted to ask whether full data were available. If no full paper was identified, we marked conference abstracts as 'pending' and will revisit them in a future review update.

Dealing with multiple publications and companion papers

Where multiple reports of a primary study were identified, we maximised yield of information by collating all available data. Where there were inconsistencies in reporting or overlapping study populations, we contacted study authors for clarification in the first instance. If this contact with authors was unsuccessful, we used the most complete and up-to-date data source where possible.

Assessment of methodological quality

We assessed risk of bias and applicability of included studies using the QUADAS-2 checklist (Whiting 2011), tailored to the review topic (see Appendix 6). The modified QUADAS-2 tool was piloted on a small number of included full text articles. One clinical and one methodologist reviewer (JDi, NC or LFR) independently assessed quality for the remaining studies; any disagreement was resolved by consensus or by a third party where necessary (JDe, CD, HW, and RM).

Statistical analysis and data synthesis

For the detection of each definition of the target condition, we conducted separate analyses according to the point in the clinical pathway that RCM was applied. Three groups of studies were formed:

i. RCM used in participants with any lesion suspicious for skin cancer, i.e. no attempt to exclude those diagnosed as obvious BCCs or SCCs or as clearly benign on visual inspection or dermoscopy was described (denoted as studies in ‘any suspicious lesion’)

ii. RCM used as an addition to dermoscopy in participants with equivocal lesions in whom a clear management decision could not be made following visual inspection and dermoscopy (denoted as studies in ‘equivocal’ lesions)

iii. ‘Other’ studies which did not fit into either of these categories

Our unit of analysis for all analyses was the lesion rather than the patient. This is because (i) in skin cancer initial treatment is directed to the lesion rather than systemically (thus it is important to be able to correctly identify cancerous lesions for each person), and (ii) it is the most common way in which the primary studies reported data. Although there is a theoretical possibility of correlations of test errors when the same people contribute data for multiple lesions, most studies include very few people with multiple lesions and any potential impact on findings is likely to be very small, particularly in comparison with other concerns regarding risk of bias and applicability. For each analysis undertaken, only one dataset was included per study to avoid over-counting of lesions. Where multiple thresholds were assessed in an individual study, datasets for correct diagnosis of each type of malignancy were selected as opposed to data for the decision to excise lesions. If data for multiple observers was reported, data for the most experienced observer was used, and data for a single observer’s diagnosis was used in preference to a consensus or average across observers. If we were unable to choose a dataset based on the above ‘rules’, a random selection of one dataset per study was made.

For each index test, algorithm or checklist under consideration, estimates of sensitivity and specificity were plotted on coupled forest plots and in receiver operating characteristic (ROC) space. For tests where commonly used thresholds were reported we estimated summary operating points (summary sensitivities and specificities) with 95% confidence and prediction regions using the bivariate hierarchical model (Chu 2006; Reitsma 2005). Where inadequate data were available for the model to converge the model was simplified, first by assuming no correlation between estimates of sensitivity and specificity and secondly by setting estimates of near zero variance terms to zero (Takwoingi 2017). Where all studies reported 100% sensitivity (or 100% specificity) the number with disease (or no disease) was summed across studies and used to compute a binomial exact 95% confidence interval. Heterogeneity in estimates of sensitivity and specificity was assessed by inspection of the magnitude and statistical significance of the estimates of variance terms in the bivariate model.

Data on the accuracy of visual inspection or dermoscopy, to allow comparisons of tests, was included only if reported in the studies of RCM due to the known substantial unexplained heterogeneity in all studies of the accuracy of these tests (Dinnes 2018a). Comparisons were made between tests using RCM data from all RCM studies, and then only using RCM data from studies that also reported visual inspection or dermoscopy data for the same patients to enable a robust direct comparison (Takwoingi 2013).

For computation of likely numbers of true positive, false positive, false negative and true negative findings in the Summary of Findings tables, summary sensitivity and specificity estimates were applied to lower quartile, median and upper quartiles of the prevalence observed in the study groups.

Bivariate models were fitted using the meqrlogit command in STATA 13.

Investigations of heterogeneity

We initially examined heterogeneity between studies by visually inspecting the forest plots of sensitivity and specificity and summary ROC plots. Where a sufficient number of studies were identified, meta-regression was performed by adding the potential source of heterogeneity as a covariate to a hierarchical model.

Sensitivity analyses

No sensitivity analyses were performed.

Assessment of reporting bias

Because of uncertainty about the determinants of publication bias for diagnostic accuracy studies and the inadequacy of tests for detecting funnel plot asymmetry (Deeks 2005), no tests to detect publication bias were performed.

Results

Results of the search

A total of 34,347 unique references were identified and screened for inclusion. Of these, 1051 full text papers were reviewed for eligibility for any one of the suite of reviews of tests to assist in the diagnosis of melanoma or keratinocyte skin cancer. Of the 1051 full text papers assessed, 848 were excluded from all reviews in our series (see Figure 5 PRISMA flow diagram of search and eligibility results).

Of the 83 studies tagged as potentially eligible for the two RCM reviews, 22 publications were included, reporting 22 individual studies: 10 in this review and 18 in the review of RCM for the detection of melanoma (6 were included in both). Reasons for exclusion included publications not being primary test accuracy studies (n=14), lack of test accuracy data (6 studies), because they were derivation studies developing new algorithms or approaches to diagnosis without the use of separate training and test sets of data (n=5), included ineligible populations, e.g. including only malignant lesions (n=6), did not assess eligible target conditions or did not adequately define the target condition (n=20), inadequate sample size (n=15), assessed the accuracy of individual RCM characteristics (n=4) or used ineligible reference standards (i.e. less than 50% of benign group with final diagnosis established by histology or follow-up; n=4). A list of the 73 studies excluded from this review with reasons for exclusion is provided in Characteristics of excluded studies, with a list of all studies excluded from the full series of reviews available as a separate pdf.

The corresponding authors of 6 publications were contacted and asked to supply further information for the purposes of this review. To date, responses have been received from only one author to allow study inclusion (Incel 2015). In addition, Professor Pellacani provided information on lesion overlap between several identified studies that were co-authored by him.

This review reports on a total of 11 cohorts of participants with lesions suspected of skin cancer, published in 10 study publications, providing 91 datasets for RCM and 4 for usual practice (visual inspection (n=1) or dermoscopy (n=3)). A total of 2037 lesions with 464 BCCs were included in the 11 datasets reporting data for BCC, and 834 lesions with 71 cSCCs were included in the 4 datasets reporting data for cSCC. The total number of study participants cannot be estimated due to lack of reporting in study publications. The Pellacani 2014 study was split into two cohorts for the purposes of this review: one cohort of lesions equivocal on dermoscopy denoted as the RCM 'consultation' group by the study authors (Pellacani 2014a (cons)); and the other cohort of lesions recommended for excision on the basis of clear cut clinical or dermoscopic findings, denoted as the RCM 'documentation' group by the study authors (Pellacani 2014b (doc)). A description of the various algorithms and thresholds used for diagnosis across the studies is provided in Appendix 7.

Methodological quality of included studies

The overall methodological quality of all included study cohorts (n=11) is summarised in Figure 6 and Figure 7. Studies were generally at high or unclear risk of bias across all domains apart from the index test and of high or unclear concern regarding applicability of the evidence.

Almost two thirds of cohorts were at high (n=4) or unclear (n=3) risk of bias for participant selection due to exclusion of poor quality images (n=3), use of a case-control type design (n=1) or unclear participant selection (n=3). All cohorts were at high (n=8) or unclear (n=3) concern regarding applicability of included participants and setting, due to restricted study populations (with 4 studies including only participants with lesions suspected of melanoma, two including only those with high clinical suspicion of BCC, and two with more narrowly defined populations such as nodular lesions or proliferative lesions) and inclusion of multiple lesions per patient (n=5). Eight of the 11 cohorts included lesions selected for excision based on the clinical or dermoscopic diagnosis or selected retrospectively from histopathology databases; this was not considered of high concern regarding applicability for RCM studies as the primary role for RCM is to reduce unnecessary excisions.

All cohorts were at low risk of bias in the index test domain. Over half of studies were high concern for the applicability of the index test (n=7), due to remote RCM interpretation (n=5), blinding to clinical information (n=3), presentation of consensus diagnoses only (n=1), lack of detail regarding the diagnostic threshold used (n=2), or interpretation by a non-expert observer (n=2). It is of note that 8 of the 11 cohorts were produced by, or in collaboration with, the same expert research team, led by Prof Pellacani which may further reduce the generalisability of results.

One cohort was at low risk of bias for the reference standard, two at high risk due to inadequate reference standards (>20% of the disease negative group with final diagnosis by follow-up or expert opinion), and 8 at unclear risk due to unclear blinding of the reference standard to the RCM result. None of the cohorts reported blinding of histology to the referral diagnosis (based on clinical examination or dermoscopy), but this was not incorporated into the overall risk of bias for this domain. For the applicability of the reference standard, one was at high risk due to the use of expert observer diagnosis as the reference standard and 9 were unclear regarding histopathology interpretation by an experienced histopathologist or by a

dermatopathologist.

For participant flow and timing, three cohorts were at low risk of bias, 5 at high risk and three at unclear risk. Three cohorts did not use the same reference standard for all participants (differential verification), 7 were unclear on the interval between the application of the index test and excision for histology, and four did not include all participants in the analysis primarily due to technical difficulties in imaging.

Findings

1 Target condition: basal cell carcinoma

In this section we present the results for studies of RCM versus visual inspection or dermoscopy for the target condition of BCC, according to the study population: studies in participants with any lesions suspicious for melanoma versus those in participants with equivocal lesions. A number of different approaches to RCM diagnosis were used across the included studies; these are described in detail in Appendix 7. Summary characteristics of studies are provided in Appendix 8. Results for the primary analyses are presented in Table 1. Forest plots of study data for each analysis in this table are given for each analysis in Figure 8 with studies plotted in ROC space in Figure 9 and Figure 10. Table 2 and Figure 11 compare results between observers.

Any suspicious lesion

The following section documents studies where RCM appeared to have been evaluated in participants with any lesion scheduled for excision. These populations are likely to include both clinically or dermoscopically obvious BCCs, along with a proportion of more difficult to diagnose (equivocal) lesions so that, RCM was being evaluated as an addition to visual inspection alone or visual inspection with dermoscopy.

Four studies provided data for the detection of BCC with RCM (Curchin 2011; Guitera 2012; Pellacani 2014b (doc); Rao 2013). A total of 912 lesions were included with 107 cases of BCC. One study provided data for expert and non-expert observers; however, only 284 of the 334 included lesions were evaluated by both readers (Rao 2013). The total number of participants cannot be reported due to lack of reporting in two of the four studies (Rao 2013 and Guitera 2012; the latter reporting overall number of patients but not the number with lesions included in the test set of data).

All studies were case series and undertaken in secondary or specialist clinic settings. Lesions were scheduled for excision reportedly for cosmetic or medical reasons (Rao 2013), reasons not reported (Curchin 2011), to rule out an 'epithelial tumour' or melanoma (Guitera 2012), or due to clinical or dermoscopy suspicion of melanoma (Pellacani 2014b (doc)). Three studies included any type of lesion, and one restricted to pigmented lesions only (Pellacani 2014b (doc)). Sample sizes ranged from 50 to 356 lesions. The median lesion to patient ratio in three studies was 1.19 (range 1.07 to 1.20). The mean prevalence of BCC was 13% (range 8% to 18%); the mean prevalence of any malignancy (BCC, cSCC, or melanoma) was 34% (range 22% to 56%). All studies also reported data for the diagnosis of melanoma (Dinnes 2018a). Studies generally included a varied spectrum of benign lesions including benign melanocytic naevi (Guitera 2012; Pellacani 2014b (doc); Rao 2013), Spitz naevi (Guitera 2012; Pellacani 2014b (doc); Rao 2013), and seborrhoeic or actinic keratoses (Guitera 2012; Pellacani 2014b (doc); Rao 2013), or both. In all studies the reference standard diagnosis was made by histology alone (i.e. all lesions either excised or biopsied).

All four studies used the Vivascope 1500 imaging system; three reporting the use of dermoscopic images to help guide acquisition of RCM images (Curchin 2011; Guitera 2012; Rao 2013). Diagnosis was reported for a single observer rather than for a consensus of observers or average value. Observer qualifications were not reported apart from in Guitera 2012 (dermatologists). Three studies were considered to have presented data for expert observers (Guitera 2012; Pellacani 2014b (doc); Rao 2013) and one for novice observers (Curchin 2011). Diagnosis was undertaken in-person with real time interpretation of RCM images (Curchin 2011; Pellacani 2014b (doc)) or remotely based on RCM images alongside the dermoscopic image of the same lesion (Guitera 2012; Rao 2013). Rao 2013 also presented data for in-person diagnosis by a less experienced observer but this was not included in the primary analysis for detection of BCC.

One study developed a new algorithm for detection of melanoma and BCC (data for the BCC element are reported here) (Guitera 2012), and the other three reported data for the correct diagnosis of each type of malignancy (Curchin 2011; Pellacani 2014b (doc); Rao 2013). Estimates of sensitivities ranged from 52% to 100% and specificities from 45% to 100% (Figure 8). The high sensitivity of 100% and low specificity of 45% in Pellacani 2014b (doc) appear as outliers, all other studies having sensitivities at or below 67% and specificities above 95% (Figure 9). Pellacani 2014b (doc) was the only study to restrict inclusion to pigmented lesions, and all lesions had 'consistent clinical and or dermoscopic criteria for melanoma diagnosis'; it also included no cSCCs in the disease negative group.

Summary sensitivity and specificity for the detection of BCC were 76% (95% CI 45% to 92%) and 95% (95% CI 66% to 99%) (Table 1). Two studies incorrectly identified other skin cancers as BCCs (8/114 false positive diagnoses), including two melanomas and two cSCCs in Guitera 2012 and four cSCCs in Rao 2013. The other study which included cSCCs (Curchin 2011) reported correct diagnosis of all 6 cSCCs or cSCC precursors.

Equivocal lesion studies

We defined equivocal lesion studies as those in which RCM was used in participants with equivocal lesions in whom a clear management decision could not be made following visual inspection or dermoscopy, i.e. RCM was being evaluated as a potential addition to dermoscopy.

Three studies provided data for the detection of BCC with RCM (Farnetani 2015; Pellacani 2014a (cons); Witkowski 2016), one providing data for nine different observers (Farnetani 2015), and one comparing the diagnosis of the same

lesions with RCM and using dermoscopic images (Witkowski 2016). A total of 668 lesions were included with 148 cases of BCC; the total number of participants cannot be reported due to lack of reporting in two studies.

All studies were case series, two of which re-interpreted previously acquired RCM images (Farnetani 2015; Witkowski 2016), and were undertaken in secondary or specialist clinic settings. Two studies included lesions suspected of being melanoma. Farnetani 2015 included any clinically equivocal lesion excised due to clinical or dermoscopic suspicion of melanoma, and Pellacani 2014a (cons) included pigmented lesions from patients requesting a mole check or with suspicion of melanoma for which an outcome decision could not be reached based on clinical or dermoscopic criteria. One study (Witkowski 2016) included clinically equivocal 'pink' cutaneous lesions with no pigmentation or containing less than 10% pigment and the absence of pigment network. Sample sizes ranged from 100 to 308 lesions. The prevalence of BCC was 6% (Pellacani 2014a (cons)), 15% (Farnetani 2015) and 44% (Witkowski 2016), and prevalence of any malignancy (BCC, cSCC, or melanoma) was 8%, 35% and 53%, respectively (only Witkowski 2016 included any cSCC). Two studies also reported data for the diagnosis of melanoma (Farnetani 2015; Pellacani 2014a (cons); see also Dinnes 2018a). Studies included a varied spectrum of benign lesions, Farnetani 2015 and Pellacani 2014a (cons) including predominantly benign melanocytic naevi and Witkowski 2016 included a relatively larger proportion of benign keratotic lesions and dermatofibromas (Appendix 8). In two studies the reference standard diagnosis was made by histology alone (Farnetani 2015; Witkowski 2016); Pellacani 2014a (cons) reported histological diagnosis for the 81 lesions initially recommended for excision, with sequential digital follow-up in the remaining 74% (227/308) of lesions; 28 of these (all found to be benign) were later excised due to changes identified on follow-up.

All studies used the Vivascope 1500 imaging system and diagnosis was reported for single observers, with Farnetani also reporting the average across 9 observers and for the majority diagnosis (5 of 9 evaluators in agreement). Observers were dermatologists (Farnetani 2015), assumed to be dermatologists (Witkowski 2016), or RCM described as conducted in a 'confocal unit' (Pellacani 2014a (cons)). Diagnosis was undertaken in-person with real time interpretation of RCM (Pellacani 2014a (cons)) or remotely based on RCM images either alongside the dermoscopic image of the same lesion (Farnetani 2015) or blinded to all other clinical information (Witkowski 2016).

All three studies reported data for observer diagnosis of BCC. Estimates of sensitivities ranged from 85% to 100% and specificities from 76% to 94% (Figure 8). The high specificity of 94% (95% CI 89% to 97%) in Witkowski 2016 appeared as an outlier (non-overlapping confidence intervals), the other two studies having specificities of 76% (95% CI 66% to 85%) (Farnetani 2015) and 79% (95% CI 73% to 83%] (Pellacani 2014a (cons)). Of note, Witkowski 2016 had a markedly different patient population to the other two studies, including only non-pigmented lesions with a markedly different spectrum of lesion types (see above).

Summary sensitivity and specificity for the detection of BCC were 94% (95% CI 79% to 98%) and 85% (95% CI 72% to 92%) (Table 1). Two studies incorrectly identified other skin cancers as BCCs (15/91 false positive diagnoses) including 14 melanomas in Farnetani 2015 and 1 cSCC in Witkowski 2016.

Witkowski 2016 also presented data for the diagnosis of nonpigmented lesions based on the dermoscopic image alone (different observers interpreting the RCM and dermoscopic images). Sensitivity and specificity estimates were almost identical, with test sensitivity 85% (95% CI 77% to 91%) for both tests and specificity 94% (95% CI 89% to 97%) for RCM compared to 92% (95% CI 87% to 96%) for dermoscopy (Table 1 and Figure 10).

Analyses by algorithms used to assist RCM

The 11 included cohorts of lesions provided 12 datasets evaluating the accuracy of different approaches to diagnosis with RCM for the detection of BCC. A description of these approaches is provided in Appendix 7.

Only one eligible study was identified that used a formally developed algorithm for the detection of BCC in an any suspicious lesion population. Guitera 2012 randomly allocated lesions to a training set for algorithm development and a test set for validation to develop a new two-step algorithm for the detection of melanoma and BCC. Lesions were reported to be predominantly melanocytic or suspicious for BCC. Applying the features found to be independently significant for BCC, sensitivity was 65% (95% CI 51% to 78%) and specificity 95% (95% CI 92% to 97%). These results are largely similar to those of the other studies in 'any suspicious lesion' (Figure 8) all of which reported observers' correct diagnosis of BCC. All data for 'equivocal lesion' populations is also based on observers' correct diagnosis of BCC without the use of any formal algorithm.

Two studies reported accuracy for features found to be independently significant for BCC but did not use a separate training set to ascertain the relevant features (Castro 2015; Longo 2013). Two studies selected lesion characteristics thought to assist the correct diagnosis of BCC based on previously published literature (Appendix 7) (Incel 2015; Nori 2004). All four studies were classified as 'other lesion population' studies, and are covered in more detail below. All studies reported sensitivities and specificities at or above 90%, apart from specificities of 78% (95% CI 40% to 97%) reported in Castro 2015, which included only 9 'benign' lesions, and of 78% (95% CI 67% to 87%) in Nori 2004, which reported only that control group lesions had a 'range of common diagnoses' to BCC (Figure 8).

Analyses by observer experience

The 11 included studies provided 19 datasets evaluating the accuracy of observers with different levels of expertise; 9 datasets coming from the same study (Farnetani 2015).

Figure 11 provides forest plots of all studies by observer experience, separately for in-person and image-based studies. Meta-analytical estimates for each group are presented in Table 2. Data for two of the 9 observers (one for high experience and one for low experience) were randomly sampled from Farnetani 2015. One further study (Rao 2013) provides a

comparison of a less experienced (in-person diagnosis) observer compared to a more experienced (but image-based diagnosis) observer; however, the two observers did not examine the same lesions (overlap of 284/334 lesions). We did not formally make any comparisons between subgroups due to the small number of studies available.

Seven cohorts presented data for observers judged to be expert or experienced in RCM: three were based on in-person evaluations (Pellacani 2014a (cons); Pellacani 2014b (doc)), or assumed to be in-person (Castro 2015); four were from image-based evaluations, two where observers were provided with the dermoscopic image of the same lesion (Farnetani 2015; Rao 2013) and two where observers were blinded to all clinical information (Guitera 2012; Longo 2013). The pooled sensitivity for the seven datasets was 98% (95% CI 74% to 100%) and pooled specificity was 87% (95% CI 71% to 95%) (Table 2). Sensitivities were at or above 90% in all studies apart from Guitera 2012 (65%, 95% CI 51% to 78%) and Rao 2013 (52%, 95% CI 32% to 71%). Specificities were more variable (45% to 98%), likely due to variations in the spectrum of disease (Figure 11).

Four cohorts presented for observers judged to be less experienced or novice: two were based on in-person evaluations (Curchin 2011; Rao 2013) and two were image-based, one providing observers with the dermoscopic image of the same lesion (Farnetani 2015) and one blinding observers to all clinical information (Nori 2004). The pooled sensitivity for the four datasets was 85% (95% CI 69% to 93%) and specificity 91% (95% CI 81% to 96%) (Table 2).

Two studies did not report the experience of RCM observers (Incel 2015; Witkowski 2016) (Table 2).

Investigations of heterogeneity

We were unable to undertake investigations of heterogeneity for other characteristics listed in the protocol due to lack of data.

2 Target condition: cutaneous squamous cell carcinoma

Two studies reported data for RCM for the target condition of cSCC: one conducted in participants with any lesion suspicious for melanoma (Rao 2013) and one in comparison to dermoscopy participants with equivocal lesions (Witkowski 2016). Summary characteristics of studies are provided in Appendix 8. Study results are presented in Table 3 with forest plots of study data in Figure 12. Two further studies present data for cSCC in 'other' lesion populations (see 'Other lesion populations' section below).

Rao 2013 included lesions scheduled for excision for cosmetic or medical reasons and presented results for the correct diagnosis of cSCC for two observers with varying levels of experience based on in-person diagnosis and interpretation of RCM images alongside dermoscopic images. For the experienced observer assessing RCM images (42 of 323 assessed were cSCC), sensitivity for the detection of cSCC was 74% (95% CI 58% to 86%) and specificity was 92% (95% CI 88% to 95%). For the less experienced observer, in-person RCM interpretation (39/318 assessed had cSCC) had a lower sensitivity of 41% (95% CI 26% to 58%) and higher specificity of 97% (95% CI 95% to 99%).

Witkowski 2016 included 260 clinically equivocal 'pink' cutaneous lesions and presented results for the correct diagnosis of cSCC (n = 13) for one observer based on RCM image interpretation and for another observer based on the dermoscopic image alone; no other clinical information was provided. Sensitivity was the same for both tests 77% (95% CI 46% to 95%), and specificities were almost identical at 98% for RCM (95% CI 96% to 100%) and 99% for dermoscopy (95% CI 96% to 100%).

3 Target condition: any skin cancer

Four studies reported data for RCM for the target condition of any skin cancer: two were conducted in participants with any lesion suspicious for melanoma (Curchin 2011; Rao 2013) and two in participants with equivocal lesions (Farnetani 2015; Witkowski 2016). Summary characteristics of studies are provided in Appendix 8. Study results are presented in Table 4 with forest plots of study data in Figure 13.

Both studies in the any suspicious lesion group included lesions scheduled for excision, with diagnosis undertaken in-person by a novice RCM reader (Curchin 2011) or remotely by an RCM expert based on RCM images (Rao 2013). Both studies reported data for the observer's correct diagnosis of each malignancy, and Rao 2013 reported data for the correct diagnosis of each type of malignancy and for the decision to excise a lesion. Rao 2013 also reported data for in-person evaluation with RCM by a less experienced observer (data not included due to a prior stated preference for more experienced observer data). A total of 373 lesions were included with 100 skin cancers (22 melanoma, 36 BCC, and 42 cSCCs), pooled sensitivity was 85% (95% CI 0.77% to 0.91%) and specificity 86% (95% CI 82% to 98%).

One of the two studies in equivocal lesions was conducted in participants with lesions excised due to suspicion for melanoma (Farnetani 2015) and one in non-pigmented or 'pink' lesions (Witkowski 2016). A total of 360 lesions were included with 175 malignant cases (32 melanomas, 129 BCCs – 114 of which were from the Witkowski 2016 dataset, 13 cSCCs and one syringoid eccrine carcinoma). Despite difference in the spectrum of included lesions, results from the two studies were similar (Figure 13), and pooled sensitivity was 89% (95% CI 0.82% to 0.94%) and specificity 85% (95% CI 75% to 92%).

One study from each of these two groups (Rao 2013; Witkowski 2016) provided data both for correct diagnosis of each malignancy and for the decision to excise suspicious lesions. Figure 14 and Figure 15 demonstrate the trade-off between higher sensitivity and lower specificity from the lower excision threshold.

4 Other lesion populations

Four evaluations of RCM in other study populations were identified. Summary characteristics of studies are provided in

Appendix 8, and forest plots and ROC plots of study data are provided in Figure 8 and Figure 9.

Two studies included lesions with a high index of suspicion for BCC. Castro 2015 included excised lesions suspicious for BCC based on clinical and dermoscopic examination and that were amenable to RCM examination using a handheld RCM probe (Vivascope 3000) to allow comparison with the standard approach (Vivascope 1500); 83% (45/54) of included lesions were histological proven to be BCC. The presence of RCM lesion characteristics selected from previous studies was assessed; however, it was not clear whether this was an image-based or in-person evaluation. Sensitivity was 100% (95% CI 92% to 100%) using the Vivascope 1500 system compared to 93% (95% CI 82% to 99%) using Vivascope 3000, and specificity estimates were both 78% (95% CI 40% to 97%). No melanomas or cSCCs were included in this study.

Nori 2004 included 83 biopsy-confirmed BCCs and a convenience sample of non-BCC with 'range of common diagnoses'; the prevalence of BCC was 55% (83/152). Diagnosis based on images acquired using the Vivascope 1000 and based on the presence of morphologic RCM characteristics previously investigated by the same group was compared to visual inspection of clinical images (latter reported for only 105 of the 152 lesions). Sensitivity and specificity were both higher using RCM (based on the presence of 3 or more RCM criteria) compared to visual inspection: sensitivity was 94% (95% CI 86% to 98%) versus 48% (95% CI 35% to 62%) and specificity 78% (95% CI 67% to 87%) compared to 62% (95% CI 46% to 75%). Results for the 4 included cSCCs could not be disaggregated from the benign diagnoses; no melanomas were included.

Incel 2015 examined 122 nonpigmented suspected malignant lesions or proliferative skin lesions with a vascular structure on dermoscopic examination with the handheld Vivascope 3000 system, using selected characteristics considered to be indicative of BCC and characteristics considered to be indicative of cSCC. The prevalence of BCC was 46% (56/122); of cSCC was 7% (9/122); with keratoacanthoma, seborrhoeic, actinic keratosis, or Bowen's disease making up half of the benign group (29/57). Sensitivity for the detection of BCC was 91% (95% CI 80% to 97%) and specificity 100% (95% CI 95% to 100%). All 9 SCCs were considered test negative (i.e. not mistaken for BCCs). Sensitivity for the detection of cSCC was 82% (95% CI 48% to 98%) and specificity 96% (95% CI 91% to 99%). Similarly, no BCCs were mistaken for SCCs in this study.

Longo 2013 included 140 clinically nodular lesions that underwent excision including 23 nodular melanomas (16%), 28 BCCs (20%), 6 cSCC (5%), and 9 with cutaneous melanoma metastases (7.5%). An experienced dermatologist interpreted RCM images blinded to dermoscopy using RCM ‘pattern analysis. Excluding non-evaluable results (including 1 BCC and 1 SCC), sensitivities were 100% for detection of BCC and SCC, and specificities were 97% (95% CI 92% to 99%) for BCC and 100% (95% CI 97% to 100%) for SCC. For the detection of any malignant lesion (excluding melanoma metastases), sensitivity was 100% (95% CI 93% to 100%) and specificity 85% (95% CI 75% to 92%).

Discussion

Summary of main results

RCM has been evaluated in a range of study populations and using a number of different approaches to assist diagnosis. Most of the data relate to the detection of BCC, with few studies recruiting sufficient numbers of participants with cSCC (i.e. >= 5) to allow accuracy to be reliably estimated. Both sensitivity and specificity for the detection of BCC appeared to vary with the spectrum of included lesions. Sensitivity was relatively low in participants with any suspicious lesion but was higher in studies of more selected populations. Studies were generally at high or unclear risk of bias across almost all domains and of high or unclear concern regarding applicability of the evidence, limiting the strength of conclusions that can be drawn. The Summary of findings table 1 presents key results for the primary target conditions of BCC and cSCC.

For the detection of BCC in participants with any suspicious lesion, RCM summary sensitivity was 76% and specificity 95%. Applying these estimates to a hypothetical cohort of 1000 lesions at the median prevalence of BCC of 12.5%, the Summary of findings table 1 shows that RCM would miss 30 of 125 BCCs with 44 false positive diagnoses. On the evidence observed, only one of these false positive results might be a melanoma with up to two misdiagnosed cSCCs. The wide confidence intervals for both estimates mean that the number of BCCs missed could range from 10 to 69, and number of false positives from 9 to 298. A single cohort of lesions with a high clinical or dermoscopic suspicion of being melanomas was responsible for most of the variation in results (Pellacani 2014b (doc)). The other studies in this group had sensitivities between 52% and 67% and specificities above 95%, which would correspond with numbers of BCCs missed at the higher end of the 10 to 69 per 1000 range, and false positives at the lower end of the 9 to 298 per 1000 range (with a corresponding reduction in the potential for melanomas being misclassified as BCCs).

RCM sensitivity was higher (94%) for the detection of BCC in participants with equivocal lesions, but with a lower specificity of 85%. Applying these estimates to a hypothetical cohort of 1000 lesions at the median prevalence of BCC of 15%, the Summary of findings table 1 shows that RCM would miss 9 of 150 BCCs with 128 false positive diagnoses. On the evidence observed for equivocal lesions, there is a much greater potential for misdiagnosis of melanomas as BCCs, with up to 19 of these false positive results potentially being melanomas. The confidence intervals around these estimates are not as wide as for any suspicious lesion: the number of BCCs missed at this disease prevalence could range between 3 and 32, and number of false positives from 68 to 238. The lowest sensitivity (85%) and highest specificity (94%) in this group were produced from the cohort of non-pigmented lesions (Witkowski 2016) as opposed to the other two studies recruiting lesions equivocal for melanoma diagnosis.

Insufficient data were available to compare RCM with visual inspection or dermoscopy or to consider the effect of using formally developed algorithms to assist RCM diagnosis of BCC. There was however some evidence of higher sensitivity from more experienced observers.

Data for the detection of cSCC were limited but suggest sensitivity in the range of 74 to 77% with high specificity of 92 to

98%.

Strengths and weaknesses of the review

The strengths of this review include an in-depth and comprehensive electronic literature search, systematic review methods including double extraction of papers by both clinicians and methodologists, and contact with authors to allow study inclusion or clarify data. A clear analysis structure according to patient pathway was adopted to allow test accuracy in different study populations to be estimated, and a detailed and replicable analysis of methodologic quality was undertaken.

The main concerns for the review are a result of the small number of studies, variation in the spectrum of included lesions, and poor reporting of primary studies, hindering the assessment of study quality and limiting the conclusions that can be drawn from the data. Despite some evidence of high sensitivity or specificity, or both, depending on the study population, research in the field has been dominated by a single expert group and results obtained from a more typical range of specialists in different countries, healthcare systems, and settings are needed. Our analysis by observer experience across algorithms and study populations lends support to the consensus that experience and observer familiarity with the diagnostic question is a key element of any diagnostic process that requires interpretation by the human eye (Norman 2009). Only one eligible study evaluated a formally developed algorithm to assist RCM interpretation; however, we excluded four studies from our review that examined individual RCM characteristics only or did not use separate training and test sets of data (Amjadi 2011; Eichert 2010; Peppelman 2015; Rishpon 2009). Further work in this area may be warranted.

Given these limitations, our results should be considered as exploratory rather than conclusive. Our results are in contrast to those of other recently published systematic reviews (Xiong 2016; Edwards 2016), one of which was conducted as part of a technology assessment report for NICE (Edwards 2016). Our review however extends the time period searched for eligible studies from 2014 in Edwards 2016 and from 2015 in Xiong 2016, considers the impact of different study populations and target conditions, and uses currently recommended methods for diagnostic test accuracy systematic reviews (Deeks 2013). Xiong 2016 did not consider varying definitions of the target condition in their primary analysis but pooled all studies regardless of detection of melanoma, BCC, or SCC. In a secondary analysis, three studies were pooled for the detection of BCC, producing estimates of sensitivity of 91.7% (95% CI 0.87 to 0.95) and specificity of 91.3% (95% CI 0.94 to 0.96); two of the three studies with high percentages of BCC lesions were included in our 'other population' analysis (Castro 2015; Nori 2004); and one was excluded from this review due to the presentation of individual RCM features for detection of BCC rather than for an overall diagnosis (Peppelman 2013). The Edwards 2016 review did not conduct a meta-analysis, instead selecting studies considered to be more applicable to a UK setting. Using the Castro 2015 study, which was included in our review as an 'other lesion population' study, economic modelling showed RCM to be a dominant strategy when used in populations with a high clinical suspicion of being BCCs in comparison to diagnostic biopsy, whether used in lesions positive or equivocal for BCC on dermoscopy. The potential for misdiagnosis of any melanomas or cSCCs as BCCs does not appear to have been considered.

Applicability of findings to the review question

Insufficient data were available to compare the accuracy of RCM with visual inspection or dermoscopy as planned. Similarly a lack of data to assess the use different algorithms to aid diagnosis was identified. It is not clear how applicable the data included in this review are regarding the routine use of RCM in a usual clinic setting as opposed to a highly specialist centre with expert RCM observers. Data are lacking regarding specific uses of the test, for example, to confirm a clinical diagnosis of BCC before initiation of non-surgical treatment. Most of the studies used the current version of the only commercially available RCM system, the Vivascope 1500. The use of remote image-based diagnosis largely by RCM experts may restrict the transferability of results to a clinical setting.

Authors' conclusions

Implications for practice

It is unclear whether RCM has a role in clinical practice for the diagnosis of BCC, although some studies suggest it has the potential to improve diagnoses. There are as yet insufficient data to support its use as a tool for avoidance of diagnostic biopsies in lesions with high clinical suspicion of BCC. In populations with a wider spectrum of lesions, there is potential for both missed BCCs and for misclassification of benign lesions, or other malignant skin cancers such as melanoma, as BCCs. Evidence for the detection cSCC is even more scarce; however, there is a clear suggestion that cSCCs could be missed with RCM. Importantly, data are lacking that compare RCM to usual practice (whether with or without dermoscopy), such that the diagnostic impact of RCM cannot be clearly estimated.

Implications for research

Further prospective evaluation of RCM in populations with a high clinical suspicion of BCC is warranted. Research should be conducted in a standard healthcare setting with a clearly defined and representative population of participants with dermoscopically equivocal lesions. RCM results should be interpreted in a usual care setting by healthcare staff representative of those who would be likely to interpret images in practice, in order to confirm the suggested increase in accuracy over dermoscopy. A multicentre approach would allow confirmation that results can be replicated across centres and that the technology can be implemented across a health service. Prospective recruitment of consecutive series of participants, with test interpretation blinded to the reference standard diagnosis and with pre-specified and clearly defined diagnostic thresholds for determining test positivity is easily achieved. Systematic follow-up of non-excised lesions avoids over-reliance on a histological reference standard and allows results to be more generalisable to routine clinical practice. A standardised approach to diagnosis, and clear

identification of the level of training and experience required to achieve good results is also required. Any future research study needs to be clear about the diagnostic pathway followed by study participants prior to study enrolment, and should conform to the updated Standards for Reporting of Diagnostic Accuracy (STARD) guideline (Bossuyt 2015).

Acknowledgements

Members of the Cochrane Skin Cancer Diagnostic Test Accuracy Group include:

- the full project team (Susan Bayliss, Naomi Chuchu, Clare Davenport, Jonathan Deeks, Jac Dinnes, Kathie Godfrey, Rubeta Matin, Colette O'Sullivan, Yemisi Takwoingi, Hywel Williams)
- our 12 clinical reviewers (Rachel Abbott, Ben Aldridge, Oliver Bassett, Sue Anne Chan, Alana Durack, Monica Fawzy, Abha Gulati, Jacqui Moreau, Lopa Patel, Daniel Saleh, David Thompson, Kai Yuen Wong) and 2 methodologists (Lavinia Ferrante di Ruffano and Louise Johnston) who assisted with full text screening, data extraction and quality assessment across the entire suite of reviews of diagnosis and staging and skin cancer,
- our expert advisor and co-author for this review Rakesh Patalay,
- and all members of our Advisory Group (Jonathan Bowling, Colin Fleming, Matthew Gardiner, Abhilash Jain, Susan O'Connell, Pat Lawton, John Lear, Mariska Leeflang, Richard Motley, Paul Nathan, Julia Newton-Bishop, Miranda Payne, Rachael Robinson, Simon Rodwell, Julia Schofield, Neil Shroff, Hamid Tehrani, Zoe Traill, Fiona Walter).

Cochrane Skin would like to thank Robert Dellavalle, who was the Dermatology Editor for this review; and a clinical referee who wishes to remain anonymous. We also wish to thank the Cochrane DTA editorial base and colleagues.

Contributions of authors

JD was the contact person with the editorial base.
JD co-ordinated contributions from the co-authors and wrote the final draft of the review.
JD, NC, DS and LP screened papers against eligibility criteria.
JD and NC obtained data on ongoing and unpublished studies.
JD, NC, DS and LP appraised the quality of papers.
JD, NC, DS and LP extracted data for the review and sought additional information about papers.
JD entered data into RevMan.
JD and JJD analysed and interpreted data.
JD, JJD, NC, YT and CD worked on the methods sections.
JD, DS, RP, RNM and HCW drafted the clinical sections of the background and responded to the clinical comments of the referees.
JD, JJD, CD and YT responded to the methodology and statistics comments of the referees.
KG was the consumer co-author and checked the review for readability and clarity, as well as ensuring outcomes are relevant to consumers.
JD is the guarantor of the update.

Disclaimer

This project was supported by the National Institute for Health Research, via Cochrane Infrastructure funding to the Cochrane Skin Group and Cochrane Programme Grant funding. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health.

Declarations of interest

Jac Dinnes: I am employed by the University of Birmingham under a National Institute for Health Research (NIHR) Cochrane Programme Grant (13-89-15) to produce the review.
Jonathan J Deeks: Work funded by an NIHR Cochrane Programme Grant. Jon Deeks receives support from NIHR through the Birmingham Biomedical Research Centre and a Senior Investigator Award.
Naomi Chuchu: nothing to declare.
Daniel Saleh: nothing to declare.
Susan E Bayliss: nothing to declare.
Yemisi Takwoingi: nothing to declare.
Clare Davenport: nothing to declare.
Lopa Patel: nothing to declare.
Rubeta N Matin: My institution received a grant for a BARCO NV commercially sponsored study to evaluate digital dermoscopy in the skin cancer clinic and Oxfordshire Health Services Research Charitable Funds for carrying out a study of feasibility of using the SCQOLIT tool in non melanoma skin cancer; I have received payment for UK Photopheresis Society lecture on cutaneous graft versus host disease October 2017; and royalties for Oxford Handbook of Medical Dermatology (Oxford University Press). I have no conflicts of interest to declare that directly relate to the publication of this work..
Colette O'Sullivan: nothing to declare.
Rakesh Patalay: nothing to declare.
Hywel C Williams: I am director of the NIHR HTA Programme. HTA is part of the NIHR which also supports the NIHR systematic reviews programme from which this work is funded.

Differences between protocol and review

For the primary objective, study populations that could not be clearly identified as either 'any suspicious lesion' or 'equivocal lesions' were considered separately as 'other lesion' studies.

Secondary objectives have been tailored to the individual test, with three objectives added: to compare the accuracy of RCM to dermoscopy where both tests have been evaluated in the same studies; to determine the diagnostic accuracy of individual algorithms for RCM; and to determine the effect of observer experience.

The secondary objective has been changed from "for the detection of any skin cancer" to "for the detection of any skin cancer in adults, *where keratinocyte skin cancers make up at least 50% of included skin cancers*" in order to keep the focus on keratinocyte skin cancers for this review and in order not to replicate analyses conducted for the review of RCM for melanoma. These changes also affect the definition of the secondary target condition in the Methods section. Heterogeneity investigations were limited by the data available.

Population inclusion criteria amended from inclusion of adults with lesions suspicious for *keratinocyte* skin cancer to inclusion of adults with lesions suspicious for *any* skin cancer, on the basis that studies targeting those with pigmented skin lesions or with lesions suspicious for melanoma also report 2x2 contingency data for the detection of BCC or cSCC within these populations. We added a requirement for a minimum of 5 benign lesions as well as 5 malignant lesions. The size threshold of five is arbitrary. However, such small studies are unlikely to add precision to estimates of accuracy.

We proposed to supplement the database searches by searching the annual meetings of appropriate organisations (e.g. British Association of Dermatologists Annual Meeting, American Academy of Dermatology Annual Meeting, European Academy of Dermatology and Venereology Meeting, Society for Melanoma Research Congress, World Congress of Dermatology, European Association of Dermato Oncology); however, due to volume of evidence retrieved from database searches and time restrictions, we were unable to do this.

For quality assessment, the QUADAS-2 tool was further tailored according to the review topic. In terms of analysis, restriction to analysis of per patient data was not performed due to lack of data. Sensitivity analyses were not performed as planned due to lack of data.

We intended to analyse studies separately according to in-person and image-based assessments; however, we were unable to do so due to lack of data.

Studies using cross-validation, such as 'leave-one-out' cross-validation were *excluded* rather than included as these methods are not sufficiently robust and are likely to produce unrealistic estimates of test accuracy.

To improve clarity of methods and to allow inclusion of studies presenting simple algorithms consisting of lesion characteristics that had previously been suggested as associated with BCC or cSCC:, this text from the protocol, “We will include studies developing new algorithms or methods of diagnosis (i.e. derivation studies) if they use a separate independent ’test set’ of participants or images to evaluate the new approach.We will also include studies using other forms of cross validation, such as ’leave-one-out’ cross-validation (Efron 1983). We will note for future reference (but not extract) any data on the accuracy of lesion characteristics individually, e.g. the presence or absence of a pigment network or detection of asymmetry”, has been replaced with

"Studies developing new algorithms or methods of diagnosis (i.e. derivation studies) were included if they:

- used a separate independent 'test set' of participants or images to evaluate the new approach, or
- investigated lesion characteristics that had previously been suggested as associated with BCC or cSCC and the study reported accuracy based on the presence or absence of particular combinations of characteristics.

Studies were excluded if they:

- used a statistical model to produce a data driven equation, or algorithm based on multiple diagnostic features, with no separate test set.
- used cross-validation approaches such as 'leave-one-out' cross-validation (Efron 1983)
- evaluated the accuracy of the presence or absence of individual lesion characteristics or morphological features, with no overall diagnosis of malignancy."

Although we extracted any reporting of special interest or accreditation in skin cancer according to observer expertise, we were unable to analyse the effect on accuracy due to lack of data.

As per the change to secondary objectives, this text from the protocol "For our secondary objective, the target condition will include any skin lesion requiring excision. We will include studies reporting data for keratinocyte skin cancer combined, and not differentiated according to BCC or cSCC, in this analysis, along with any melanoma or rare skin cancer (e.g. Merkel or amelanotic melanoma) that may be detected. We will not consider in situ cancers or actinic keratosis as disease-positive" has been changed to "An additional definition of the target condition was considered in secondary analysis, the detection of:

- any skin cancer, including BCC, cSCC, melanoma or any rare skin cancer (e.g. Merkel cell cancer), as long as skin cancers other than melanoma made up more than 50% of the disease positive group. Data from studies in which melanoma accounted for more than 50% of skin cancers were included in the review of RCM for the diagnosis of melanoma (Dinnes 2018a)."

Added the following as possible sources of heterogeneity in Appendix 2:

- patient population: Primary /secondary / specialist unit
- lesion suspicion: general suspicion/atypical/equivocal/NR
- lesion type: any pigmented; melanocytic

- inclusion of multiple lesions per participant

Published notes

Characteristics of studies

Characteristics of included studies

Castro 2015

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Not reported **Period of data collection:** Not reported **Country:** Brazil and USA
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	High risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Patients recruited were those presenting with one or more skin lesions that were highly suspicious for BCC based on clinical and dermoscopic examination. All lesions underwent biopsy. **Setting:** Specialist unit (skin cancer/pigmented lesions clinic) **Prior testing:** Clinical or dermatoscopic suspicion, or both **Setting for prior testing:** Unspecified **Exclusion criteria:** Poor quality index test image; From discussion: "HH-RCM imaging was successfully performed in all lesions in which imaging was attempted, while TWP-RCM was technically applicable in only 59% of lesions in which imaging was attempted." **Sample size (patients):** No. eligible: 73 **Sample size (lesions):** No. eligible: 92. No. included: 54 **Participant characteristics:** Mean age 65y (30-89y). Fitzpatrick phototype: 24 patients with type II; 8 patients with type III **Lesion characteristics:** Site reported for BCCs only - Head/Neck: 9, Trunk: 26, Upper limbs/shoulder: 4, Lower limbs/hip: 6
Are the included patients and chosen study setting appropriate?	No
Did the study avoid including participants with multiple lesions?	Unclear
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** Vivascope 3000; using two different ways of assessment using hand-held (HH-RCM) and traditional wide-probe (TWP-RCM) No algorithm. Previously-published RCM criteria assessed (cites Agero 2006, Nori 2004, Guitera 2012) and selected criteria chosen **Method of diagnosis:** Not clearly reported; may be in person "All examinations, including clinical, dermoscopic and RCM imaging, were made by a dermatologist experienced with RCM examination (RPRC) with supervision by a skin cancer expert (GGR or HR)." However, also states that "All RCM images were evaluated jointly by two readers blinded to the results of the histopathological examination." Not clear whether this was undertaken at the time of RCM examination or subsequently. **Prior test data:** Clinical examination and dermoscopy "All examinations, including clinical, dermoscopic and RCM imaging, were made by a dermatologist experienced with RCM examination (RPRC) with supervision by a skin cancer expert (GGR or HR)." **Diagnostic threshold:** ≥3 RCM criteria present, **Diagnosis based on:** Consensus (2 observers); (n= 2) **Observer qualifications:** Dermatologist **Experience in practice:** High experience or 'Expert' **Experience with index test:** High experience /'Expert' users (not stated but both observers co-authored studies developing RCM) **Other detail:** attempted imaging with HH-RCM and TWP-RCM using a standardized protocol, however, TWP-RCM imaging is restricted to anatomic locations that allow contact and is not feasible in some anatomic locations such as the eyelids **Derivation aspect to study:** Images were evaluated for the presence of previously-published RCM criteria for identification of BCC (2,6,7); Approach to selection of characteristics indicative of skin cancer was not described. Characteristics selected: "at least one of the criteria had to be the presence of 'dark silhouettes' or 'bright tumor islands'; these latter criteria denote the presence of neoplastic aggregates of BCC and hence need to be observed in all cases identified as BCC by RCM." Additional criteria assessed were: • 'streaming' polarization of nuclei in neoplastic aggregates along the same axis of orientation; • 'peripheral palisading' of nuclei at the tumor islands' periphery; • dark 'peritumoral clefts' around the tumor islands; • fibrotic stroma with 'thickened collagen bundles'; • dilated and tortuous 'linear blood vessels' and 'coiled blood vessels'; • 'bright dendritic structures' within tumor islands; and • 'bright round cells' in the stroma.

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	No
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	Yes
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 45 BCCs; Disease negative: 9 **Target condition (Final diagnoses)** BCC: 45 'Benign' diagnoses: 9
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Yes
Are there concerns that the target condition as defined by the reference standard does not match the question?	Low concern

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: Imaging with both TWP-RCM and HH-RCM was attempted in92 lesions from 73 patients; however, 38 of the lesions (41%),mostly facial, were excluded as they were only accessible to HH-RCM imaging. Time interval to reference test: Not reported
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	High risk

Notes

Notes	-

Curchin 2011

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Prospective **Period of data collection:** Jan 2010 to May 2010 **Country:** Australia
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Patients from Dermatology department's minor excision booking list; not further described **Setting:** Secondary (general dermatology) **Prior testing:** Selected for excision (no further detail) **Setting for prior testing:** Unspecified **Exclusion criteria:** None reported **Sample size (patients):** No. included: 42 **Sample size (lesions):** No. included: 50 **Participant characteristics:** None reported **Lesion characteristics:** None reported
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	No
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** Vivascope 1500; RCM score (Pellacani 2007) and LM score for suspected lentigo maligna of the face (Guitera 2010) **Method of diagnosis:** In person diagnosis **Prior test data:** Dermoscopy "dermoscopic and RCM images were aligned over the top of each other so that correlation between the two could be made" **Diagnostic threshold:** For melanoma - RCM score: >=3; threshold for LM score for suspected lentigo maligna of the face was not described (Guitera 2010). Observer diagnosis for SCC/BCC; RCM features listed **Diagnosis based on:** Single observer; (n= 1?) **Observer qualifications:** Not reported **Experience in practice:** Not described **Experience with index test:** Low experience / novice users; analysis was performed by a novice to RCM analysis after completing a RCM analysis course in Modena, Italy.

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Yes
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	No
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 21; Disease negative: 29 **Target condition (Final diagnoses)** Melanoma (invasive): 12; Melanoma (in situ): 1; BCC: 9; cSCC: 6 (includes SK or AK, or both) 'Benign' diagnoses: 23
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Time interval to reference test: NR
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Low risk

Notes

Notes	-

Farnetani 2015

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series; series of cases consecutively and retrospectively selected by an expert dermoscopist for a web-based inter-observer reliability study **Data collection:** Retrospective image selection / Prospective interpretation **Period of data collection:** not reported **Country:** Italy (lesion image acquistion); Observers were located in the US (3), Europe (4), Australia (1) and Israel (1).
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Diagnostically equivocal lesions excised due to clinical or dermoscopic suspicion of melanoma, where a specific clinical and dermoscopic diagnosis could not be rendered with certainty. Lesions selected by an expert dermoscopist blinded to final diagnosis **Setting:** Secondary (general dermatology); All included RCM images were collected at the Department of Dermatology of the University of Modena and ReggioEmilia (Modena, Italy), **Prior testing:** Clinical or dermatoscopic suspicion, or both **Setting for prior testing:** Secondary (general dermatology) **Exclusion criteria:** Poor quality index test image; "No additional selection criteria were considered in case selection such as the presence or lack of pigmentation, diameter, elevation, or other clinical or dermoscopic attribute" **Sample size (patients):** No. included: NR **Sample size (lesions):** No. included: 100 **Participant characteristics:** None reported **Lesion characteristics:** None reported
Are the included patients and chosen study setting appropriate?	Yes
Did the study avoid including participants with multiple lesions?	Unclear
Are there concerns that the included patients and setting do not match the review question?	Unclear

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** Vivascope 1500; No algorithm - evaluators completed a 'pattern description' (presence/absence of a number of RCM features) **Method of diagnosis:** Confocal images (remote); 3 RCM mosaic images presented per lesion **Prior test data:** Dermoscopy "Each case for evaluation had a high-resolution dermoscopic image obtained with a dermoscopic lens that was attached to a digital camera." "No additional clinical information (eg, age and melanoma or lesion history) was provided to evaluators." **Diagnostic threshold:** Evaluators completed a 'pattern description' (presence/absence of a number of RCM features) and gave an overall diagnosis of malignant (melanoma or BCC) or benign. **Diagnosis based on:** Single observer (results presented for each of 9 observers); Consensus (≥5 of 9 evaluators); Average (across 9 observers and across 6 more experienced and 3 less experienced observers); (n= 9) **Observer qualifications:** Dermatologist **Experience in practice:** Not described **Experience with index test:** Low experience / novice users (3 with <3 years RCM experience). High experience /'Expert' users (6 with >=3 years RCM experience) **Derivation aspect to study:** In addition to pattern analysis described above, discriminant analysis was used to identified RCM features independently associated with malignancy, melanoma and BCC. Three of 6 discriminatory RCM features were more frequently observed in melanoma: the presence of pagetoid cells, the presence of atypical cells at the DEJ, and irregular epidermal architecture; 3 of 6 discriminatory RCM features were more frequently observed in BCCs: basaloid cord–like structures, presence of ulceration, and a specific DEJ pattern. Accuracy was not estimated for combinations of these particular features

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	No
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Unclear
Was the test interpretation carried out by an experienced examiner?	Yes
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 35; Disease negative: 65 **Target condition (Final diagnoses)** Melanoma (in situ and invasive, or not reported): 20; BCC: 15 Seborrheic keratosis: 7; Other: 55 melanocytic nevi, 3 actinic keratoses
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: Excised lesions only included Time interval to reference test: not reported Time interval between index test(s): N/A
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Unclear risk

Notes

Notes	-

Guitera 2012

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Not reported **Period of data collection:** NR **Country:** Australia and Italy **Test set derived:** randomly split into training and test sets
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Consecutive patients presenting or found with suspicious lesions, including all macules of the face and neck suspicious for lentigo maligna, and which would be subjected to biopsy or excision to rule out an epithelial tumor or an MM following conventional clinical and dermoscopy diagnosis and with lesion location amenable to RCM; described as predominantly melanocytic or suspicious for BCC **Setting:** Mixed, lesions recruited from Modena (general dermatology) and Sydney (skin cancer/pigmented lesions clinic) **Prior testing:** Clinical or dermatoscopic suspicion, or both **Exclusion criteria:** Location/site of lesion keratotic, sole, and palm lesions were excluded **Sample size (patients):** No. eligible: 663 **Sample size (lesions):** No. eligible: 710 / No included: 356 in test set, 253 melanocytic **Participant characteristics:** Median age (full sample): 53, IQR 39 to 66 (for full sample), Range: 6-90; Male: 354; 53.4% (of full sample) **Lesion characteristics:** Not reported
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	No
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM):** RCM score and Segura algorithm; also derived own independently significant features for MM and BCC. Vivascope 1500 **Method of diagnosis:** Confocal images (remote) **Prior test data:** Lesion site or patient age only, or both: "RCM features were described by two expert observers (GP and PG), blinded from any clinical information, dermoscopy, and clinical aspects, but not for the location and age of the patient" **Diagnostic threshold:** Two established algorithms assessed: Pellacani scoring system for melanoma (Pellacani 2007), score >3; and Segura 2-step algorithm (Segura 2009), score of zero; own new two step model identified 7 independently significant features for MM (assume presence of any one indicated T+): • cerebriform nests, • atypical cobblestone pattern with small nucleated cells in the epidermis, • marked cytological atypia, • pagetoid cells, • disarranged epidermal layer with no honey comb, • large inter-papillae spaces filled with honeycomb, • dense nest. 8 independently significant features for BCC: • polarized in the honeycomb, • linear telangiectasia-like horizontal vessels, • basaloid cord or nodule, • epidermal shadow, • convoluted glomerular-like vessels, • non-visible papillae, • cerebriform nests, • disarray of the epidermal layer. **Derivation aspect to study:** Lesion characteristics assessed a series of 48 features, corresponding to previous observations (Pellacani 2007; Guitera 2009), and new descriptors were considered at three different depth levels. Descriptions and definitions provided. Selection of characteristics indicative of skin cancer by multivariate discriminant analysis performed on the training set. **Diagnosis based on:** Single observer (n=2) **Observer qualifications:** Dermatologist **Experience in practice:** High experience or 'Expert' **Experience with index test:** High experience /'Expert' users

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Unclear
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	Unclear
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Unclear

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** not further described; full sample Disease positive: 335 / disease negative 375 **Target condition (Final diagnoses):** Test set only Melanoma (in situ and invasive, or not reported): 105; BCC: 52; cSCC: 9 Benign nevus 132; Spitz nevus 16; actinic keratosis 8; 31 benign macule of the face and 3 dermatofibroma
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	No exclusions Imaged prior to biopsy
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Low risk

Notes

Notes	-

Incel 2015

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Prospective (assumed); "Written consent was obtained from all participants before enrolment." **Period of data collection:** NR **Country:** Turkey
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Patients with nonpigmented suspected tumoral lesions or proliferative skin lesions and with a vascular structure on dermoscopic examination **Setting:** Secondary (not further described) Istanbul Training and Research Hospital **Prior testing:** Clinical or dermatoscopic suspicion, or both - all participants underwent clinical evaluations "following guidelines of the visual inspection and diagnosis of nonpigmented skin tumor"; those with a vascular structure on dermoscopic examination underwent RCM **Setting for prior testing:** Secondary (general dermatology); Specialist unit (skin cancer/pigmented lesions clinic) **Exclusion criteria:** prominent hyperkeratosis; history of significant other skin disease, kidney, liver, heart disease, surgery, or invasive procedure on the localization of tumor in the last 6 months, sunbathing or indoor tanning in the last 3 months, and subjects who are receiving therapy that has angiogenic effects such as systemic/topical steroids **Sample size (patients):** No. included: 114 **Sample size (lesions):** No. included: 122 **Participant characteristics:** Median age 61y, range 18-87y; Male: 57% **Lesion characteristics:** Site - head and neck (76.2%), extremities (10.2%), back, abdomen, and chest (13.6%).
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	No
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM):** Characteristics from previous studies selected to assist correct diagnosis of different lesion types; also assessed vascularity of lesions using RCM but this did not inform diagnosis. Vivascope 3000 **Method of diagnosis:** Unclear; Images of first 60 lesions subjected to blinded evaluation by 2 observers to identify vascular morphology; unclear whether overall diagnoses reported were based on images or in person assessments **Prior test data:** Unclear **Diagnostic threshold:** Characteristics listed for BCC included: Dark silhouettes in dermis, Bright tumour islands at DEJ and in the dermis; Cleft-like dark areas; Dendritic cells, Bright rond cells, Canalicular vessels. Characteristics listed for SCC included: Refractile squam/crust in stratum corneum and nucleated cells with dark center (parakeratotic) cells; Atypical honeycomb pattern, disarranged pattern at stratum granulosum layer; Large, round, nucleated cells at the granular layer (dyskeratotic cells); Dendritic cells at the granular layer and small edged papillae at DEJ; Dendritic cells (referenced to Malvehy 2012, Eichert 2010, Ahlgrimm-Siess 2010, Röwert-Huber 2007, Ahlgrimm-Siess 2011) **Derivation aspect to study: Study assessed** assessed vascularity of lesions with RCM but diagnoses of each lesion type reportedly based on above characteristics. **Diagnosis based on:** NR **Observer qualifications:** NR **Experience in practice:** NR **Experience with index test:** NR

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Unclear
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	Unclear
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Unclear

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** Clinically, dermoscopically, and confocally suspected malignant lesions, recurrent, and therapy resistant lesions were excised; benign appearing but suspected lesions were punch biopsied. Formalin fixed paraffin-embedded tissue sections were stained with hematoxylin–eosin. Histopathological examination was conventionally operated by light microscopy **Target condition (Final diagnoses):** BCC: 56; cSCC: 9 Keratoacanthoma 3; Sebhorrheic keratosis 11; Actinic keratosis 8; Bowen's disease 7; and 24 other nonpigmented tumors that included sebaceous hyperplasia (4), eccrine poroma (4), pyogenic granuloma (3), amelanotic melanoma (2), sebaceous adenoma (2), trichilemmoma (2), warty dyskeratoma (1), pilomatrixoma (1), kaposi sarcoma (1), fibrohistiocytic tumor (1), eccrine spiradenoma (1), and eccrine porocarcinoma (1).
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Index to reference interval: appears consecutive "Biopsy was taken for routine histology from selected patients, and was examined with RCM." No exclusions were reported
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Low risk

Notes

Notes	-

Longo 2013

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Retrospective image selection / Prospective interpretation **Period of data collection:** NR **Country:** Italy
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Clinically nodular lesions (defined as cutaneous palpable/superficial seated lesions and not subcutaneous ones) that underwent excision **Setting:** Secondary (general dermatology); Specialist unit (skin cancer/pigmented lesions clinic) **Prior testing:** Selected for excision (no further detail) **Setting for prior testing:** Secondary (general dermatology); Specialist unit (skin cancer/pigmented lesions clinic) **Exclusion criteria:** None reported (not evaluable and non specific RCM results excluded; see below) **Sample size (patients):** No. included: 140 **Sample size (lesions):** No. included: 140 **Participant characteristics:** Mean age 50 years (SD 19.7). Male: 64; 45.7% **Lesion characteristics:** All clinically nodular; Site - 'most' on the trunk; dermatofibroma mainly located on extremities. Mean thickness 2.16 mm (SD 82); 23 'pure' nodular melanomas
Are the included patients and chosen study setting appropriate?	No
Did the study avoid including participants with multiple lesions?	Yes
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** Model NR; likely Vivascope 1500. No algorithm - reports observer diagnosis and independently significant features **Method of diagnosis:** Confocal images (remote) **Prior test data:** No further information used; blinded to dermoscopic image **Diagnostic threshold:** A diagnosis was formulated based on 'RCM pattern analysis' (Longo 2011; Pellacani 2007; and several others cited) **Diagnosis based on:** Single observer (n= 1) **Observer qualifications:** Dermatologist **Experience in practice:** High experience or 'Expert'; 5 years' experience in RCM and therefore presumably in practice **Experience with index test:** High experience /'Expert' users; 5 years' experience in RCM **Derivation aspect to study:** Main study data included relate to observers interpretation of RCM characteristics and diagnostic classification. Univariate and then multivariate discriminant analysis was also performed to identify independently significant RCM criteria (total of 36 assessed) for NM+Mets vs. all other diagnoses, BCC vs. all other diagnoses, SCC vs. all other diagnoses. The data presented however relate to only 130 lesions, and the melanoma metastases cannot be separated from the nodular melanoma, therefore melanoma data not included Characteristics selected for nodular melanoma or melanoma mets were: widespread pagetoid distribution ; many atypical cells; and cerebriform nests. Characteristics selected for BCC were: tumour islands ; cauliflower architecture; bright filaments within the tumour islands; and presence of bright collagen. Not clearly reported for SCC

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	No
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Unclear
Was the test interpretation carried out by an experienced examiner?	Yes
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 66; Disease negative: 57 **Target condition (Final diagnoses)** Melanoma (invasive): 23 nodular; BCC: 28; cSCC: 6; Other malignant: 9 melanoma metastases Sebhorrheic keratosis: 14; Benign naevus: 32; (14 compound, 8 intradermal, 3 blue naevi, 7 Spitz naevi); Other: 5 vascular and 6 other benign lesions
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: 8 not evaluable and 3 'nonspecific' RCM results reported (appear to be excluded from derivation of independently significant characteristics) Not evaluable: lesions where all the three levels (epidermis, DEJ and upper dermis) were not explorable for any reason that hampered the collection of quality images or the exploration of DEJ/superficial dermis. Nonspecific: lesions where a diagnosis could not be formulated, despite the possibility of exploring all three levels, because of the impossibility of recognizing diagnostic features with enough confidence.
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Unclear risk

Notes

Notes	-

Nori 2004

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case control; appears to be case-control type design sampling BCC and non-BCC **Data collection:** Retrospective image selection / Prospective interpretation **Period of data collection:** 2 years - date range not specified **Country:** US and Spain
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	No
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	High risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Biopsy-confirmed BCC and convenience sample of non-BCC with 'range of common diagnoses' (only 7 of the 69 non-BCC had BCC on the list of possible differential diagnoses); of these 105 images were selected for clinical assessment based on superior clinical image quality **Setting:** Secondary (general dermatology) Division of Dermatology, Loma Linda Uni Med School; Dermatology Servive, Madrid, Spain; Private care (Dermatology and Dermatologic surgery private practice); Wellmann Laboratories of Photomedicine, Massachusetts General Hospital **Prior testing:** Not reported **Setting for prior testing:** Unspecified **Exclusion criteria:** None reported **Sample size (patients):** No. included: 145 **Sample size (lesions):** No. included: 152; 105 in VI analysis **Participant characteristics:** Male: 98; 64% **Lesion characteristics:** Lesion site: Face/Ears: 35%; Trunk: 13%; Limbs: Extremities 45%; Back 7%
Are the included patients and chosen study setting appropriate?	No
Did the study avoid including participants with multiple lesions?	Yes
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Visual inspection (VI)** No algorithm **Method of diagnosis:** Clinical photographs **Prior test data:** No further information used **Diagnostic threshold:** Not reported. Lesions assigned to: High probability (BCC until proven otherwise), medium probability (would biopsy to rule out BCC), and low probability (no biopsy needed). **Diagnosis based on:** Single observer (n=2) **Observer qualifications:** Dermatologist **Experience in practice:** Not described **Experience with index test:** Not described # **Reflectance confocal microscopy (RCM).** Vivascope 1000 (plus prototype device built in Wellman Laboratories (n=20)). No algorithm; selected characteristics assessed (referenced to Gonzalez 2002) **Method of diagnosis:** Confocal images (remote) **Prior test data:** No further information used; images from all 152 lesions was retrospectively analyzed in a blinded fashion **Diagnostic threshold:** >=2, 3, 4 or 5 present of 5 features selected based on prior study by same authors (?elongated monomorphic basaloid nuclei; polarization of these nuclei along the same axis of orientation; prominent inflammatory infiltrate; increased dermal vasculature; pleomorphism of the overlying epidermis indicative of actinic changes.) **Diagnosis based on:** Single observer (n= 1) **Observer qualifications:** 'Novice confocal reviewer' reviewed all images **Experience in practice:** Not described **Experience with index test:** Low experience / novice users; novice confocal reviewer **Other detail** The images produced by Vivascope 1000 and prototype device reportedly similar, with no measurable differences between them. **Derivation aspect to study:** The 5 criteria were chosen as they were "easily and unambiguously detected by non dermatopathologists and our novice reviewer so that the applicability of our results would be useful to the dermatology community as a whole."

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	No
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	No
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis plus other **Details:** 15 lesions were not biopsied (eg, lesions like seborrheic keratosis) because the clinical diagnosis was considered diagnostic Histology: Disease positive: 83; Disease negative: 54 Expert opinion; Disease positive: 0; Disease negative: 15 **Target condition (Final diagnoses)** BCC: 83; 58 in VI analysis; cSCC: 4 'Benign' diagnoses: 65
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	No
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	High

Flow and Timing

A. Risk of Bias	
Flow and timing	
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	Unclear
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	High risk

Notes

Notes	-

Pellacani 2014a (cons)

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series (consultation group documented here) **Data collection:** Prospective **Period of data collection:** January 2010 to December 2010 **Country:** Italy
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	High risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Patients requesting a mole check or with suspicion of melanoma who were referred to pigmented lesion clinic and who were then found to have atypical lesions on dermoscopy. Those in whom diagnosis could not be determined on dermoscopy were referred for an 'outcome decision' (consultation group). Patients were referred on the basis of both urgent access (melanoma suspected in a single lesion by GP or other dermatologist) and scheduled access (referred for dermoscopy and total body examination). **Setting:** Specialist unit (skin cancer/pigmented lesions clinic) **Prior testing:** Dermatoscopic suspicion in all cases. All patients underwent dermoscopy in the PLC; those with dermoscopically atypical lesions were referred for RCM, either to document a lesion already selected for excision (documentation group, reported in Pellacani 2014b (doc)) or for an 'outcome decision' (consultation group), i.e. diagnosis could not be determined on dermoscopy **Setting for prior testing:** Specialist unit (skin cancer/pigmented lesions clinic) **Exclusion criteria:** Clinically or dermatoscopically clear-cut epithelial tumours, or both, were not enrolled; Poor quality index test image - In nine cases RCM could not be performed (five RCM documentation and four RCM consultation) due to the presence of artefacts, hyperkeratosis, or ulcerations, impeding imaging. **Sample size (patients):** No. eligible: 1005 examined with dermoscopy; No. included: 252 referred for RCM consultation **Sample size (lesions):** No. eligible: NR; No. included: 308 for RCM documentation **Participant characteristics:** Median age 41.7 (IQR 31.9, 52.1); For all referred patients (n=1005): 443 male (44%); Consultation group only: History of melanoma/skin cancer 23 (7%); Family history of melanoma 30 (10%). Fitzpatrick phototype I to II: 150 (49%); Type III to IV 116 (38%) **Lesion characteristics:** Lesion site (full sample) Head/Neck: 9%; Trunk: 59%; Upper limbs/shoulder: 12%; Lower limbs/hip: 20%
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	No
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** RCM score Vivascope 1500 **Method of diagnosis:** In person **Prior test data:** Patients were "referred to confocal unit"; confocal reader was blinded to the patient pathway and aware that lesions were dermoscopically atypical for 'RCM documentation' or for 'RCM consultation'. **Diagnostic threshold:** Not reported, Pellacani 2005 cited **Diagnosis based on:** Single observer (n=1) **Observer qualifications:** Dermatologist (assumed; patients were "referred to confocal unit") **Experience in practice:** Not described **Experience with index test:** Not described but 'confocal unit' described **Other detail** Any other detail Dermatoscopy examinations were conducted using the Dermlite HR (3Gen LLC, San Juan Capistrano, CA, U.S.A.). Lesions that were scheduled for digital monitoring were also acquired by means of FotoFinder (TeachScreen GmbH, Bad Birnbach, Germany) using 20-fold magnification.

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Yes
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	Yes
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis plus FU and cancer registry follow-up Histology (not further described): 81 (consultation group); overall dataset - 292 excised (see Pellacani 2014b (doc)] Clinical FU: 227. 28 of which were subsequently excised because of observed dermatoscopic changes (all benign). Most non excised lesions (89.4% 178/199) were followed up for 1 year; the others were lost at the 1-year follow-up. Cancer registry FU: Those lost to clinical follow-up were checked on the tumour registry; no melanomas were diagnosed in patients scheduled for follow-up after baseline examinations. **Target condition (Final diagnoses)** Melanoma (invasive): 13; Melanoma (in situ): 9; BCC: 19; 1 melanoma metastasis Clark naevus 71; Spitz nevus 5; solar lentigo, seborrhoeic keratosis or lichen planus-like keratosis 0; other benign 207 (8 with histological diagnosis (25 Clark naevi, two Spitz naevi and one benign nonmelanocytic lesion) and 199 benign on FU)
Is the reference standards likely to correctly classify the target condition?	No
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: 9 excluded due to RCM failure
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Was the minimum clinical follow-up after application of index test(s) adequate?	Yes
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	High risk

Notes

Notes	-

Pellacani 2014b (doc)

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series (documentation group described here) **Data collection:** Prospective **Period of data collection:** January 2010 to December 2010 **Country:** Italy
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias?	High risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Patients requesting a mole check or with suspicion of melanoma who were referred to pigmented lesion clinic and who were then found to have atypical lesions on dermoscopy. Those in whom excision was required on dermoscopy were referred for RCM documentation (documentation group). Patients were referred on the basis of both urgent access (melanoma suspected in a single lesion by GP or other dermatologist) and scheduled access (referred for dermoscopy and total body examination). **Setting:** Specialist unit (skin cancer/pigmented lesions clinic) **Prior testing:** Dermatoscopic suspicion in all cases. All patients underwent dermoscopy in the PLC; those with dermoscopically atypical lesions were referred for RCM, either to document a lesion already selected for excision (documentation group, as reported here) or for an 'outcome decision' (consultation group, reported in Pellacani 2014a (cons)), i.e. diagnosis could not be determined on dermoscopy. **Setting for prior testing:** Specialist unit (skin cancer/pigmented lesions clinic) **Exclusion criteria:** Clinically or dermatoscopically clear-cut epithelial tumours, or both, were not enrolled; Poor quality index test image - In nine cases RCM could not be performed (five RCM documentation and four RCM consultation) due to the presence of artefacts, hyperkeratosis, or ulcerations, impeding imaging. **Sample size (patients):** No. eligible: 1005 examined with dermoscopy; No. included: 171 referred for RCM documentation **Sample size (lesions):** No. eligible: NR; No. included: 183 for RCM documentation **Participant characteristics:** Median age 41.2 (IQR 35, 63); For all referred patients (n=1005): 443 male (44%); History of melanoma/skin cancer 8 (5%); Family history of melanoma 13 (8%). Fitzpatrick phototype I to II: 99 (58%); Type III to IV 72 (42%) **Lesion characteristics:** Lesion site (full sample) Head/Neck: 9%; Trunk: 59%; Upper limbs/shoulder: 12%; Lower limbs/hip: 20%
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	No
Are there concerns that the included patients and setting do not match the review question?	High

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** RCM score Vivascope 1500 **Method of diagnosis:** In person **Prior test data:** Patients were "referred to confocal unit"; confocal reader was blinded to the patient pathway and aware that lesions were dermoscopically atypical for 'RCM documentation' or for 'RCM consultation'. **Diagnostic threshold:** Not reported Pellacani 2005 cited **Diagnosis based on:** Single observer (n=1) **Observer qualifications:** Dermatologist (assumed; patients were "referred to confocal unit") **Experience in practice:** Not described **Experience with index test:** Not described but 'confocal unit' described **Other detail** Any other detail Dermatoscopy examinations were conducted using the Dermlite HR (3Gen LLC, San Juan Capistrano, CA, U.S.A.). Lesions that were scheduled for digital monitoring were also acquired by means of FotoFinder (TeachScreen GmbH, Bad Birnbach, Germany) using 20-fold magnification.

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Yes
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	Yes
Was the test interpretation carried out by an experienced examiner?	Yes
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Low concern

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histology alone for documentation group; 227 from consultation group were referred for follow-up (see Pellacani 2014a (cons)) **Target condition (Final diagnoses)** Melanoma (invasive): 13; Melanoma (in situ): 9; BCC: 19; 1 melanoma metastasis Clark naevus 121; Spitz nevus 8; solar lentigo, seborrhoeic keratosis or lichen planus-like keratosis 7; other benign 5 (haemosiderotic dermatofibroma, xanthogranuloma, viral wart and two nonspecific inflammatory dermatoses)
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: 9 excluded due to RCM failure
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Was the minimum clinical follow-up after application of index test(s) adequate?	Yes
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	High risk

Notes

Notes	-

Rao 2013

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Not reported Appear to be prospective but not explicitly stated **Period of data collection:** Jun 2010-Sep 2011 **Country:** US
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** All lesions removed for cosmetic or medical reasons that were imaged using a confocal scanning microscope **Setting:** Secondary (general dermatology) Based on author institutions **Prior testing:** Not reported **Setting for prior testing:** Unspecified **Exclusion criteria:** None reported **Sample size (patients)**: NR **Sample size (lesions):** No. eligible: 340; No. included: 334 **Participant characteristics:** None reported **Lesion characteristics:** None reported
Are the included patients and chosen study setting appropriate?	Unclear
Did the study avoid including participants with multiple lesions?	Unclear
Are there concerns that the included patients and setting do not match the review question?	Unclear

Index Test

Index tests	**Reflectance confocal microscopy (RCM).** Vivascope 1500. No algorithm; Overall diagnosis **Method of diagnosis:** In person diagnosis US (reader 1; less experienced) Confocal images (remote) Modena, Italy; Reader 2 (more experienced) **Prior test data:** For image based "diagnosis was based on the dermoscopic image and confocal microscopy evaluation before excision." **Diagnostic threshold:** Not reported; Observers gave diagnosis and excise decision (no further details) **Diagnosis based on:** Single observer (n= 2) **Observer qualifications:** Not reported; Presume dermatologists **Experience in practice:** Not described **Experience with index test:** Low experience / novice users Reader 1 (US) had 1 year of experience at the beginning of the study; High experience /'Expert' users Reader 2 (Italy) had over 9 years of experience with RCM. **Other detail** Images were sent via Vivanet (CaliberID, Rochester, NY), a Health Insurance Portability and Accountability Act-compliant server.15

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	Unclear
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	No
Was the test interpretation carried out by an experienced examiner?	Unclear
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 78; Disease negative: 256 **Target condition (Final diagnoses)** Melanoma (invasive): 8; Melanoma (in situ); 1; BCC: 27; cSCC: 42 Benign nevi 176; seborrhoeic keratosis 22; actinic keratosis 24; other 23
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: 6 described as excluded because of insufficient information. Furthermore 318/334 reported for Reader 1 and 323/334 reported for Reader 2
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	High risk

Notes

Notes	-

Witkowski 2016

Patient Selection

A. Risk of Bias	
Patient Sampling	**Study design:** Case series **Data collection:** Retrospective image selection / Prospective interpretation **Period of data collection:** January 2009–2011 **Country:** Italy
Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Yes
Could the selection of patients have introduced bias?	Low risk

B. Concerns regarding applicability	
Patient characteristics and setting	**Inclusion criteria:** Consecutive clinically equivocal 'pink' cutaneous lesions with absent pigmentation or containing less than 10% pigment and absence of pigment network. All lesions were excised at first visit or follow-up video dermoscopy control visit and had available digital dermoscopy images and a complete standard set of RCM images, with histopathology reports **Setting:** Secondary (general dermatology) **Prior testing:** Clinical suspicion of malignancy without dermatoscopic suspicion **Setting for prior testing:** Secondary (general dermatology) **Exclusion criteria:** Unequivocal appearance/diagnosis benign diagnosis made with high confidence; lack of histological report as a result of the lesion not being excised **Sample size (patients):** NR **Sample size (lesions):** No. eligible: 3869 consecutive cases were reviewed; No. included: 260 **Participant characteristics:** None reported **Lesion characteristics:** None reported
Are the included patients and chosen study setting appropriate?	Yes
Did the study avoid including participants with multiple lesions?	Unclear
Are there concerns that the included patients and setting do not match the review question?	Unclear

Index Test

Index tests	**Dermoscopy** No algorithm **Method of diagnosis:** Dermoscopic images **Prior test data:** No further information used **Diagnostic threshold:** Not described in detail, but accuracy presented for two diagnostic decisions: correct diagnosis (of BCC, MM and SCC) and correct management decision (excise or not) **Diagnosis based on:** Single observer (n = 2; one reader evaluated only dermoscopic images while the second reader evaluated RCM images) **Observer qualifications:** Not reported; likely dermatologist **Experience in practice:** Not described **Experience with index test:** Not described **Reflectance confocal microscopy (RCM)** Vivascope 1500; No algorithm - Overall diagnosis **Method of diagnosis:** Confocal images (remote) **Prior test data:** No further information used "The first reader (JL) evaluated only dermoscopic images while the second reader (AW) evaluated RCM images." **Diagnostic threshold:** Not reported. Not described in detail, but accuracy presented for two diagnostic decisions: correct diagnosis (of BCC, MM and SCC) and correct management decision (excise or not) **Diagnosis based on:** Single observer (n= 1) **Observers** as described above

Reflectance confocal microscopy

A. Risk of Bias	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes
For studies reporting the accuracy of multiple diagnostic thresholds, was each threshold or algorithm interpreted without knowledge of the results of the others?	
Could the conduct or interpretation of the index test have introduced bias?	Low risk

B. Concerns regarding applicability	
Was the test applied and interpreted in a clinically applicable manner?	No
Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	No
Was the test interpretation carried out by an experienced examiner?	Unclear
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	High

Reference Standard

A. Risk of Bias	
Target condition and reference standard(s)	**Type of reference standard:** Histological diagnosis alone **Details:** No further details provided Disease positive: 140; Disease negative: 120 **Target condition (Final diagnoses)** Melanoma (in situ and invasive, or not reported): 12; BCC: 114; cSCC: 13; 1 syringoid eccrine carcinoma Sebhorrheic keratosis plus other: 25 (solar lentigo/seborrheic keratosis/lichen planus-like keratosis/ actinic keratosis); 47 nevi; 6 spitz nevi; 18 dermatofibromas (DF), 4 vascula rlesions, and 20 other type benign lesions (1 clear cell acanthoma, 1 discoid lupus, 10 inflammatory lesions, 1 perivascular hyperplasia, 4 granulomatous hyperacanathosis reactions, 1 papulous fibrosis, 1 eccrine poroma,and 1 eczematous lesion).
Is the reference standards likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear
Were the reference standard results interpreted without knowledge of the referral diagnosis?	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk

B. Concerns regarding applicability	
Expert opinion (with no histological confirmation) was not used as a reference standard	Yes
Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	Unclear
Are there concerns that the target condition as defined by the reference standard does not match the question?	Unclear

Flow and Timing

A. Risk of Bias	
Flow and timing	Excluded participants: Around 357 cases were excluded due to the lack of a histopathology report, as a result of the lesion not being excised, or a benign diagnosis was made with high confidence. Time interval to reference test: not reported Time interval between index test(s): not reported
Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Was the minimum clinical follow-up after application of index test(s) adequate?	
If more than one algorithm evaluated for the same test, was the interval between application of the different algorithms 1 month or less?	
Could the patient flow have introduced bias?	Unclear risk

Notes

Notes	-

Footnotes

Characteristics of excluded studies

Agero 2006

Reason for exclusion	EXCLUDE on sample size *only 5 lesions*

Ahlgrimm-Siess 2010

Reason for exclusion	EXCLUDE on study population*; BCC only* EXCLUDE on sample size; *only 2 cases*

Ahlgrimm-Siess 2011

Reason for exclusion	EXCLUDE on study population*; SCC only* EXCLUDE on sample size; *only 2 cases*

Alarcon 2014

Reason for exclusion	EXCLUDE on sample size; target condition (eligible for melanoma review only)

Amjadi 2011

Reason for exclusion	EXCLUDE on target population *Includes only BCC (82)/SCC (48) and 8 AK/SK lesions; primary aim appears to be to differentiate BCC and SCC despite describing inclusion of clinically difficult to diagnose non-pigmented lesions.* EXCLUDE on derivation study

Bassoli 2012

Reason for exclusion	EXCLUDE on target condition *The aim of this study was to identify criteria for specific diagnosis of LPLK using in vivo RCM.*

Benati 2015

Reason for exclusion	EXCLUDE if individual lesion characteristics

Braga 2009

Reason for exclusion	EXCLUDE on sample size *case reports*

Carrera 2015

Reason for exclusion	EXCLUDE not a primary study

de Carvalho 2015

Reason for exclusion	EXCLUDE if individual lesion characteristics EXCLUDE on 2x2 data

de Carvalho 2016

Reason for exclusion	EXCLUDE on target condition EXCLUDE on sample size

Edwards 2016

Reason for exclusion	EXCLUDE not a primary study systematic review

Eichert 2010

Reason for exclusion	EXCLUDE on individual lesion characteristics; l*ooks at accuracy of previously identified RCM features for melanoma, BCC and SCC in a cohort of 100 lesions but does not give accuracy for overall diagnosis of each group.*

Ferrari 2015

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Figueroa Silva 2016

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Gareau 2009

Reason for exclusion	EXCLUDE on study population *Only BCC cases*

Gerger 2005

Reason for exclusion	EXCLUDE on reference standard *only 1/3 of disease negative group had adequate ref test* EXCLUDE duplicate or related publication; d*ata reported as training set in* Koller 2011 *(#860)*

Gerger 2006

Reason for exclusion	EXCLUDE on reference standard *Only 30/120 benign were excised (30/90 benign nevi and 0/30 SK)*

Gerger 2008

Reason for exclusion	EXCLUDE on reference standard *all MMs were excised plus 14/50 benign; remainder diagnosed on clinical/dermoscopic criteria*

Giambrone 2015

Reason for exclusion	EXCLUDE on target condition EXCLUDE but contact authors *they do not give information on the target condition-only state malignant/benign cutaneous lesions??? Contacted 8-5-17*

Gill 2014

Reason for exclusion	EXCLUDE if derivation study; *looking for correlation with histological features* EXCLUDE on 2x2 data; *Looks at correlation between RCM features and histological features; not test accuracy* EXCLUDE duplicate or related publication; *Same lesions reportedly included in* Pellacani 2012

Gonzalez 2002

Reason for exclusion	EXCLUDE on population includes only BCC

Gonzalez 2013

Reason for exclusion	EXCLUDE not a primary study

Guida 2015

Reason for exclusion	EXCLUDE not a primary study systematic review

Guitera 2009

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Guitera 2010

Reason for exclusion	EXCLUDE on target condition; *only looking at LM and not LMM*

Guitera 2013

Reason for exclusion	EXCLUDE on study population; *LM and LMM only* EXCLUDE on target condition; *data only available for LM* EXCLUDE on 2x2 data

Haenssle 2006

Reason for exclusion	EXCLUDE on index test; *surveillance study estimating accuracy of different approaches to follow-up*

Hennessy 2010

Reason for exclusion	EXCLUDE on 2x2 data

Hofmann-Wellenhof 2008

Reason for exclusion	EXCLUDE on reference standard; *includes 70 melanocytic lesions - 20 MM (all histologically verified); 70 benign naevi (28% histologically verified, and the rest diagnosed with dermoscopy only).*

Hoogedoorn 2014

Reason for exclusion	EXCLUDE conference abstract

Hoogedoorn 2015

Reason for exclusion	EXCLUDE on sample size

Humphrey 2006

Reason for exclusion	EXCLUDE on study population EXCLUDE as derivation study - assesses lesion vascularity

Kadouch 2015

Reason for exclusion	systematic review

Kadouch 2015a

Reason for exclusion	EXCLUDE not a primary study *clinical trial protocol*

Koller 2011

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Kose 2014

Reason for exclusion	EXCLUDE not a test accuracy study EXCLUDE on 2x2 data

Langley 2001

Reason for exclusion	EXCLUDE on 2x2 EXCLUDE but contact authors; *contact authors for RCM 2x2 data can only get 2x2 for clinical diagnosis*

Langley 2006

Reason for exclusion	EXCLUDE on sample size

Langley 2007

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Losi 2014

Reason for exclusion	EXCLUDE if individual lesion characteristics EXCLUDE on 2x2 data

Lovatto 2015

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Maier 2013

Reason for exclusion	EXCLUDE on study population; *all study participants had final diagnosis of melanoma*

Malvehy 2012

Reason for exclusion	EXCLUDE not a primary study; *review paper*

Menge 2016

Reason for exclusion	EXCLUDE on target population; *includes participants with primary possible recurrent and or previously treated lesions and does not disaggregate results. Also includes multiple lesions per participant (63 'sites' from 17 participants; unclear how many of the 39 LM positive on histology had melanoma).*

Miller 2011

Reason for exclusion	EXCLUDE on target condition EXCLUDE on 2x2 data; *not an accuracy study*

Nobre 2011

Reason for exclusion	EXCLUDE on sample size; *case report*

Pellacani 2005

Reason for exclusion	EXCLUDE if derivation study; *uses leave one out*

Pellacani 2007

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Pellacani 2008

Reason for exclusion	EXCLUDE on 2x2 data; *no accuracy data provided in the study, looking at correlation of RCM features to dermoscopy and histology*

Pellacani 2009

Reason for exclusion	EXCLUDE on 2x2 data; *Study is testing concordance of terminology used in RCM...not accuracy.*

Pellacani 2012

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Peppelman 2013

Reason for exclusion	EXCLUDE on study population; *only present data for subtypes of BCC* EXCLUDE on 2x2 data; *does not give accuracy data*

Peppelman 2015

Reason for exclusion	EXCLUDE if derivation study EXCLUDE on 2x2 data; *no data for overall accuracy*

Peppelman 2016

Reason for exclusion	EXCLUDE not a primary study; *RCT protocol*

Puig 2012

Reason for exclusion	EXCLUDE on sample size; *case report*

Pupelli 2013

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Reggiani 2015

Reason for exclusion	EXCLUDE not a primary study; systematic review

Rishpon 2009

Reason for exclusion	EXCLUDE on sample size; *only 3 invasive SCC* EXCLUDE if derivation study *RCM characteristics for SCC*

Röwert-Huber 2007

Reason for exclusion	EXCLUDE not a primary study; *review paper*

Salerni 2011

Reason for exclusion	EXCLUDE on sample size; *<5 cases*

Scope 2009

Reason for exclusion	EXCLUDE on sample size

Scope 2014

Reason for exclusion	EXCLUDE not a primary study; e*ditorial paper*

Segura 2009

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Soyer 2013

Reason for exclusion	EXCLUDE not a primary study; *comment on a primary study (* Longo 2013 *)*

Stanganelli 2015

Reason for exclusion	EXCLUDE on target condition (eligible for melanoma review only)

Steiner 1992

Reason for exclusion	EXCLUDE on sample size *only two melanomas*

Stephens 2013

Reason for exclusion	EXCLUDE on sample size

Stevenson 2013

Reason for exclusion	EXCLUDE not a primary study *systematic review of RCM*

Tannous 2009

Reason for exclusion	EXCLUDE on sample size; *only two malignant melanomas*

Willard 2011

Reason for exclusion	EXCLUDE on sample size; *case study*

Xiong 2016

Reason for exclusion	EXCLUDE not a primary study *systematic review of RCM*

Yelamos 2016

Reason for exclusion	EXCLUDE not a primary study

Footnotes

Characteristics of studies awaiting classification

Borsari 2016

Patient Sampling	Not yet assessed
Patient characteristics and setting	Not yet assessed
Index tests	Not yet assessed
Target condition and reference standard(s)	Not yet assessed
Flow and timing	Not yet assessed
Comparative	Not yet assessed
Notes	Published October 2016; after search dates

Guitera 2016

Patient Sampling	Not yet assessed
Patient characteristics and setting	Not yet assessed
Index tests	Not yet assessed
Target condition and reference standard(s)	Not yet assessed
Flow and timing	Not yet assessed
Comparative	Not yet assessed
Notes	Published October 2016; after search dates

Jain 2017

Patient Sampling	Not yet assessed
Patient characteristics and setting	Not yet assessed
Index tests	Not yet assessed
Target condition and reference standard(s)	Not yet assessed
Flow and timing	Not yet assessed
Comparative	Not yet assessed
Notes	Published March 2017; conference abstract only

Ludzik 2016

Patient Sampling	Not yet assessed
Patient characteristics and setting	Not yet assessed
Index tests	Not yet assessed
Target condition and reference standard(s)	Not yet assessed
Flow and timing	Not yet assessed
Comparative	Not yet assessed
Notes	Published September 2016; after search dates

Footnotes

Characteristics of ongoing studies

Footnotes

Summary of results tables

1 Summary of findings table

Question:	**What is the diagnostic accuracy of reflectance confocal microscopy for the detection of keratinocyte skin cancers in adults?**
Population:	Adults with lesions suspicious for skin cancer, including: • in participants with any suspicious lesion, where RCM might be used as an alternative to dermoscopy or to supplement visual inspection alone • in participants with equivocal lesions in whom a clear management decision could not be made following visual inspection and dermoscopy, where RCM might be used as an addition to visual inspection or dermoscopy, or both
Index test:	Reflectance confocal microscopy (RCM)
Comparator test:	Visual inspection or dermoscopy, or both
Target condition:	• Basal cell carcinoma (BCC) • Cutaneous squamous cell carcinoma (cSCC)
Reference standard:	Histology with or without long-term follow-up

Question:	**What is the diagnostic accuracy of reflectance confocal microscopy for the detection of keratinocyte skin cancers in adults?**			
Action:	If accurate, negative results of RCM will stop patients having unnecessary excision of skin lesions; positive results could inform the use of nonsurgical management options			
Quantity of evidence	**Target condition**	**Number of studies**	**Total lesions**	**Total cases**
	BCC	10 (11 cohorts)	2037	464
	cSCC	4 (4 cohorts)	834	71
Limitations				
Risk of bias:	High (4/11) or unclear (3/11) risk for patient selection with exclusion on image quality and use of a case-control design. Low risk for index test (11/11). High risk from inadequate reference standard (2/11) and unclear blinding of the reference standard to the RCM result (8/11). Differential verification in 3/11, timing of tests not mentioned in 7/11 and exclusions due to technical difficulties in 4/11.			
Applicability of evidence to question:	High (8/11) or unclear (3/11) concern for participants and setting with narrowly defined populations (3/11) or multiple lesions per patient (5/11). High concern for index test (7/11) with remote RCM interpretation (5/11) blinded to clinical information (3/11), lack of detail on the diagnostic threshold (2/11) and novice RCM users (2/11). The studies are dominated by one particularly expert research group (8/11). Little information given concerning the expertise of the histopathologist.			
FINDINGS:				
A total of 10 studies providing data for 11 cohorts of lesions were eligible for inclusion, seven in our target populations of interest. All seven cohorts reported data for detection of BCC and two reported data for detection of cSCC. The findings presented are based on results from these seven cohorts. Insufficient data were available to compare RCM with visual inspection or dermoscopy or to consider the effect of using formally developed algorithms to assist RCM diagnosis. There was some evidence of melanomas or SCCs being misdiagnosed as BCCs and of higher sensitivity in cohorts using more experienced observers.				
Test: RCM for detection of BCC using any or no algorithm at any threshold in any suspicious lesion				
Datasets (n)	**Lesions (n)**	**Cases (n)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
4	912	107	76% (95% CI: 45, 92%)	95% (95% CI: 66, 99%)

Numbers in a cohort of 1000 lesions**	**True positives**	**False positives**	**False negatives**	**True negatives**	**PPV**	**NPV**
At median prevalence 12.5%	95 (56; 115)	44 (298; 9)	30 (69; 10)	831 (578; 866)	68% (16; 93)	97% (89; 99)
At lower quartile prevalence 10%	75 (45; 91)	45 (306; 9)	24 (54; 8)	856 (595; 892)	63% (13; 91)	97% (92; 99)
At upper quartile prevalence 15%	118 (70; 143)	42 (287; 8)	37 (85; 12)	803 (558; 837)	74% (20; 94)	96% (87; 99)

Test: RCM for detection of BCC using any or no algorithm at any threshold in equivocal lesions				
Datasets (n)	**Lesions (n)**	**Cases (n)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
3	668	148	94% (79, 98)	85% (72, 92)

Numbers in a cohort of 1000 lesions**	**True positives**	**False positives**	**False negatives**	**True negatives**	**PPV**	**NPV**
At median prevalence 15%	141 (119; 147)	128 (238; 68)	9 (32; 3)	723 (612; 782)	53% (33; 68)	99% (95; 100)

Question:	What is the diagnostic accuracy of reflectance confocal microscopy for the detection of keratinocyte skin cancers in adults?					
At lower quartile prevalence 11%	100 (84; 104)	134 (250; 72)	6 (22; 2)	760 (644; 822)	43% (25; 59)	99% (97; 100)
At upper quartile prevalence 29%	276 (232; 288)	106 (198; 56)	18 (62; 6)	600 (508; 650)	72% (54; 84)	97% (89; 99)

Test: RCM for detection of cSCC using any or no algorithm				
	Lesions (n)	Cases (n)	Sensitivity (95% CI)	Specificity (95% CI)
In any suspicious lesion (n = 1 study)	323	42	74% (58 to 86%)	92% (88 to 95%)
In equivocal lesions (n = 1 study)	260	13	77% (46 to 95%)	98% (96 to 100%)

Footnotes

** Number of TP, FP, FN, TN have been estimated at the median and interquartile ranges of prevalence, at average sensitivity and specificity and using the lower and upper limits of the 95% confidence intervals, denoted in brackets (lower limit; upper limit)

Additional tables

1 Comparison of RCM with dermoscopy for the detection of BCC

Test	Studies	Lesions (cases)	Sensitivity (95% CI)	Specificity (95% CI)
Any suspicious lesion studies (all studies)				
RCM	4	912	0.76 (0.45 to 0.92)	0.95 (0.66 to 0.99)
Dermoscopy	0		-	-
Any suspicious lesion studies (direct comparisons)				
RCM	0		-	-
Dermoscopy	0		-	-
Equivocal lesion studies (all studies)				
RCM	3	668	0.94 (0.79 to 0.98)	0.85 (0.72 to 0.92)
Dermoscopy	1	260	0.85 (0.77 to 0.91)	0.92 (0.87 to 0.96)
Equivocal lesion studies (direct comparisons)				
RCM	1	260	0.85 (0.77 to 0.91)	0.94 (0.89 to 0.97)
Dermoscopy	1	260	0.85 (0.77 to 0.91)	0.92 (0.87 to 0.96)

Footnotes

2 Results by observer experience

Person / image	Observer experience	Studies (n)	Lesions (cases)	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)
Detection of BCC					
In-person	High	3	545 (83)	1.00 (0.96 to 1.00)	0.67 (0.45 to 0.83)
Image	High	4	908 (119)	0.86 (0.50 to 0.97)	0.94 (0.86 to 0.98)
Both	**High**	**7**	**1453 (202)**	**0.98 (0.74 to 1.00)**	**0.87 (0.71 to 0.95)**
In-person	Low	2	368 (34)	0.75 (0.53 to 0.88)	0.97 (0.64 to 1.00)
Image	Low	2	252 (98)	0.92 (0.82 to 0.97)	0.84 (0.75 to 0.90)
Both	**Low**	**4**	**620 (132)**	**0.85 (0.69 to 0.93)**	**0.91 (0.81 to 0.96)**
In-person	NR	1	122 (56)	0.91 (0.80 to 0.97)	1.00 (0.95 to 1.00)
Image	NR	1	260 (114)	0.85 (0.77 to 0.91)	0.94 (0.78 to 0.97)
Both	**NR**	**2**	**382 (170)**	**0.88 (0.80 to 0.93)**	**0.98 (0.74 to 1.00)**
Detection of cSCC					
In-person	High	0	0	-	-
Image	High	2	452 (47)	0.95 (0.06 to 1.00)	0.99 (0.40 to 1.00)
Both	**High**	**2**	**452 (47)**	**0.95 (0.06 to 1.00)**	**0.99 (0.40 to 1.00)**
In-person	Low	1	318 (39)	0.41 (0.26 to 0.58)	0.97 (0.95 to 1.00)
Image	Low	0	0	-	-
Both	**Low**	**1**	**318 (39)**	**0.41 (0.26 to 0.58)**	**0.97 (0.95 to 1.00)**
In-person	NR	1	122 (11)	0.82 (0.48 to 0.98)	0.96 (0.91 to 0.99)
Image	NR	1	260 (13)	0.77 (0.46 to 0.95)	0.98 (0.96 to 1.00)
Both	**NR**	**2**	**382 (24)**	**0.79 (0.59 to 0.91)**	**0.98 (0.96 to 0.99)**
Detection of any skin cancer (KER)					
In-person	High	0	0	-	-
Image	High	3	552 (161)	0.94 (0.70 to 0.99)	0.86 (0.82 to 0.90)
Both	**High**	**3**	**552 (161)**	**0.94 (0.70 to 0.99)**	**0.86 (0.82 to 0.90)**
In-person	Low	2	368 (95)	0.80 (0.71 to 0.87)	0.85 (0.81 to 0.89)
Image	Low	1	90 (35)	0.83 (0.66 to 0.93)	0.85 (0.74 to 0.92)
Both	**Low**	**3**	**458 (130)**	**0.81 (0.73 to 0.87)**	**0.85 (0.81 to 0.89)**
In-person	NR	0	0	-	-
Image	NR	1	260 (140)	0.91 (0.85 to 0.95)	0.80 (0.72 to 0.87)
Both	NR	1	**260 (140)**	**0.91 (0.85 to 0.95)**	**0.80 (0.72 to 0.87)**

Footnotes

NR - Not reported

3 Comparison of RCM with dermoscopy for the detection of cSCC

Test	Studies	Lesions (cases)	Sensitivity (95% CI)	Specificity (95% CI)
All lesion studies (all studies)				
RCM	1	323 (42)	0.74 (0.58 to 0.86)	0.92 (0.88 to 0.95)
Dermoscopy	0		-	-
All lesion studies (direct comparisons)				
RCM	0		-	-
Dermoscopy	0		-	-
Equivocal lesion studies (all studies)				
RCM	1	260 (13)	0.77 (0.46 to 0.95)	0.98 (0.96 to 1.00)
Dermoscopy	1	260 (13)	0.77 (0.46 to 0.95)	0.99 (0.96 to 1.00)
Equivocal lesion studies (direct comparisons)				
RCM	1	260 (13)	0.77 (0.46 to 0.95)	0.98 (0.96 to 1.00)
Dermoscopy	1	260 (13)	0.77 (0.46 to 0.95)	0.99 (0.96 to 1.00)

Footnotes

4 Comparison of RCM with dermoscopy for the detection of any skin cancer

KER	Studies	Lesions (cases)	Sensitivity (95% CI)	Specificity (95% CI)
Any suspicious lesion studies (all studies)				
RCM	2	373 (100)	0.85 (0.77 to 0.91)	0.86 (0.82 to 0.98)
Dermoscopy	0		-	-
Any suspicious lesion studies (direct comparisons)				
RCM	0		-	-
Dermoscopy	0		-	-
Equivocal lesion studies (all studies)				
RCM	2	360 (175)	0.89 (0.82 to 0.94)	0.85 (0.75 to 0.92)
Dermoscopy	1	260 (140)	0.91 (0.86 to 0.95)	0.79 (0.71 to 0.86)
Equivocal lesion studies (direct comparisons)				
RCM	1	260 (140)	0.91 (0.85 to 0.95)	0.80 (0.72 to 0.87)
Dermoscopy	1	260 (140)	0.91 (0.86 to 0.95)	0.79 (0.71 to 0.86)

Footnotes

References to studies

Included studies

Castro 2015

* Castro RP, Stephens A, Fraga-Braghiroli NA, Oliviero MC, Rezze GG, Rabinovitz H, et al. Accuracy of in vivo confocal microscopy for diagnosis of basal cell carcinoma: a comparative study between handheld and wide-probe confocal imaging. Journal of the European Academy of Dermatology and Venereology 2015;29(6):1164-1169. [DOI: 10.1111/jdv.12780; Other: ER4:20569441; PubMed: 25338750]

Curchin 2011

* Curchin CE, Wurm EM, Lambie DLj, Longo C, Pellacani G, Soyer HP. First experiences using reflectance confocal microscopy on equivocal skin lesions in Queensland. Australasian Journal of Dermatology 2011;52(2):89-97. [Other: ER4:15465900; PubMed: 21605091]

Farnetani 2015

* Farnetani F, Scope A, Braun RP, Gonzalez S, Guitera P, Malvehy J, et al. Skin Cancer Diagnosis With Reflectance Confocal Microscopy: Reproducibility of Feature Recognition and Accuracy of Diagnosis. JAMA Dermatology 2015;

151(10):1075-80. [Other: ER4:25233569; PubMed: 25993262]

Guitera 2012

* Guitera P, Menzies SW, Longo C, Cesinaro AM, Scolyer RA, Pellacani G. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. Journal of Investigative Dermatology 2012;132(10):2386-94. [Other: ER4:15465942; PubMed: 22718115]

Incel 2015

Incel P, Gurel M S, Erdemir AV. Vascular patterns of nonpigmented tumoral skin lesions: confocal perspectives. Skin Research & Technology 2015;21(3):333-9. [PubMed: 25345376]

Longo 2013

* Longo C, Farnetani F, Ciardo S, Cesinaro AM, Moscarella E, Ponti G, et al. Is confocal microscopy a valuable tool in diagnosing nodular lesions? A study of 140 cases. British Journal of Dermatology 2013;169(1):58-67. [Other: ER4:15465992; PubMed: 23374159]

Nori 2004

* Nori S, Rius-Diaz F, Cuevas J, Goldgeier M, Jaen P, Torres A, et al. Sensitivity and specificity of reflectance-mode confocal microscopy for in vivo diagnosis of basal cell carcinoma: a multicenter study. Journal of the American Academy of Dermatology 2004;51(6):923-30. [Other: ER4:15466027; PubMed: 15583584]

Pellacani 2014a (cons)

Pellacani G, Pepe P, Casari A, Longo C. Reflectance confocal microscopy as a second-level examination in skin oncology improves diagnostic accuracy and saves unnecessary excisions: a longitudinal prospective study. British Journal of Dermatology 2014;171(5):1044-1051. [PubMed: 24891083]

Pellacani 2014b (doc)

Pellacani G, Pepe P, Casari A, Longo C. Reflectance confocal microscopy as a second-level examination in skin oncology improves diagnostic accuracy and saves unnecessary excisions: a longitudinal prospective study. British Journal of Dermatology 2014;171(5):1044-1051. [PubMed: 24891083]

Rao 2013

* Rao BK, Mateus R, Wassef C, Pellacani G. In vivo confocal microscopy in clinical practice: comparison of bedside diagnostic accuracy of a trained physician and distant diagnosis of an expert reader. Journal of the American Academy of Dermatology 2013;69(6):e295-300. [Other: ER4:15466076; PubMed: 24035553]

Witkowski 2016

* Witkowski AM, Ludzik J, DeCarvalho N, Ciardo S, Longo C, DiNardo A, et al. Non-invasive diagnosis of pink basal cell carcinoma: how much can we rely on dermoscopy and reflectance confocal microscopy? Skin Research & Technology 2016;22(2):230-7. [Other: ER4:25012281; PubMed: 26338448]

Excluded studies

Agero 2006

Agero AL, Busam KJ, Benvenuto-Andrade C, Scope A, Gill M, Marghoob AA, et al. Reflectance confocal microscopy of pigmented basal cell carcinoma. Journal of the American Academy of Dermatology 2006;54(4):638-43.

Ahlgrimm-Siess 2010

Ahlgrimm-Siess V, Cao T, Oliviero M, Hofmann-Wellenhof R, Rabinovitz HS, Scope A. The vasculature of nonmelanocytic skin tumors in reflectance confocal microscopy: vascular features of basal cell carcinoma. Archives of Dermatology 2010; 146(3):353-4.

Ahlgrimm-Siess 2011

Ahlgrimm-Siess V, Cao T, Oliviero M, Hofmann-Wellenhof R, Rabinovitz HS, Scope A. The vasculature of nonmelanocytic skin tumors on reflectance confocal microscopy: vascular features of squamous cell carcinoma in situ. Archives of Dermatology 2011;147(2):264.

Alarcon 2014

Alarcon I, Carrera C, Turegano P, Malvehy J, Puig S. Basal cell carcinoma with spontaneous regression: added value of reflectance confocal microscopy when the dermoscopic diagnosis is uncertain. Journal of the American Academy of Dermatology 2014;71(1):e7-9.

Amjadi 2011

* Amjadi M, Coventry BJ, Greenwood JE. Reflectance confocal microscopy in the diagnosis of non-melanoma skin cancer and benign lesions versus normal skin: a blinded prospective trial. Internet Journal of Plastic Surgery 2011; 7(2):1-6. [Other: ER4:21450593]

Bassoli 2012

Bassoli S, Rabinovitz HS, Pellacani G, Porges L, Oliviero MC, Braun RP, et al. Reflectance confocal microscopy criteria of lichen planus-like keratosis. Journal of the European Academy of Dermatology & Venereology 2012;26(5):578-90.

Benati 2015

Benati E, Argenziano G, Kyrgidis A, Moscarella E, Ciardo S, Bassoli S, et al. Melanoma and naevi with a globular pattern: confocal microscopy as an aid for diagnostic differentiation. British Journal of Dermatology 2015;173(5):1232-8.

Braga 2009

Braga JC, Scope A, Klaz I, Mecca P, Gonzalez S, Rabinovitz H, et al. The significance of reflectance confocal microscopy in the assessment of solitary pink skin lesions. Journal of the American Academy of Dermatology 2009;61(2):230-41.

Carrera 2015

Carrera C. High-risk melanoma patients: can unnecessary naevi biopsies be avoided? British Journal of Dermatology 2015; 172(2):313-5.

de Carvalho 2015

de Carvalho N, Farnetani F, Ciardo S, Ruini C, Witkowski AM, Longo C, et al. Reflectance confocal microscopy correlates of dermoscopic patterns of facial lesions help to discriminate lentigo maligna from pigmented nonmelanocytic macules. British Journal of Dermatology 2015;173(1):128-33.

de Carvalho 2016

deCarvalho N, Guida S, Abraham L Spagnol, Cesinaro AM, Farnetani F, Bonamonte D, et al. Pink melanocytic and non-melanocytic lesions: how reflectance confocal microscopy can help in differential diagnosis. Journal of the European Academy of Dermatology & Venereology 2016;30(6):1026-9.

Edwards 2016

Edwards SJ, Mavranezouli I, Osei-Assibey G, Marceniuk G, Wakefield V, Karner C. VivaScope 1500 and 3000 systems for detecting and monitoring skin lesions: a systematic review and economic evaluation. Health Technology Assessment (Winchester, England) 2016;20(58):1-260.

Eichert 2010

Eichert S, Mohrle M, Breuninger H, Rocken M, Garbe C, Bauer J. Diagnosis of cutaneous tumors with in vivo confocal laser scanning microscopy. Journal der Deutschen Dermatologischen Gesellschaft 2010;8(6):400-10.

Ferrari 2015

* Ferrari B, Pupelli G, Farnetani F, De Carvalho NT, Longo C, Reggiani C, et al. Dermoscopic difficult lesions: an objective evaluation of reflectance confocal microscopy impact for accurate diagnosis. Journal of the European Academy of Dermatology and Venereology : JEADV 2015;29(6):1135-1140. [DOI: 10.1111/jdv.12769; Other: ER4:20569458; PubMed: 25303304]

Figueroa Silva 2016

* Figueroa-Silva O, Cinotti E, de Almeida Silva T, Moscarella E, Lallas A, Ciardo S, et al. Diagnostic accuracy of reflectance confocal microscopy for lesions typified by dermoscopic island. Journal of the European Academy of Dermatology and Venereology : JEADV 2016;30(9):1594-8. [Other: ER4:25012335; PubMed: 27109574]

Gareau 2009

Gareau DS, Karen JK, Dusza SW, Tudisco M, Nehal KS, Rajadhyaksha M. Sensitivity and specificity for detecting basal cell carcinomas in Mohs excisions with confocal fluorescence mosaicing microscopy. Journal of Biomedical Optics 2009; 14(3):034012.

Gerger 2005

Gerger A, Koller S, Kern T, Massone C, Steiger K, Richtig E, et al. Diagnostic applicability of in vivo confocal laser scanning microscopy in melanocytic skin tumors. Journal of Investigative Dermatology 2005;124(3):493-8.

Gerger 2006

Gerger A, Kerl H, Samonigg H, Langsenlehner U, Krippl P, Smolle J. Sensitivity and specificity of confocal laser scanning microscopy for in vivo diagnosis of malignant skin tumors. Journal of Investigative Dermatology 2006;126(Supplement 3):s114.

Gerger 2008

Gerger A, Wiltgen M, Langsenlehner U, Richtig E, Horn M, Weger W, et al. Diagnostic image analysis of malignant melanoma in in vivo confocal laser-scanning microscopy: a preliminary study. Skin Research & Technology 2008; 14(3):359-63.

Giambrone 2015

Giambrone D, Alamgir M, Masud A, Bronsnick T, Rao B. The diagnostic accuracy of in vivo confocal microscopy in clinical practice. Journal of the American Academy of Dermatology 2015;73(2):317-9.

Gill 2014

Gill M, Longo C, Farnetani F, Cesinaro AM, Gonzalez S, Pellacani G. Non-invasive in vivo dermatopathology: identification of reflectance confocal microscopic correlates to specific histological features seen in melanocytic neoplasms. Journal of the European Academy of Dermatology & Venereology 2014;28(8):1069-78.

Gonzalez 2002

Gonzalez S, Tannous Z. Real-time in vivo confocal reflectance microscopy of basal cell carcinoma. Journal of the American Academy of Dermatology 2002;47:869-74.

Gonzalez 2013

Gonzalez S. Should reflectance confocal microscopy be the gold standard for basal cell carcinoma diagnosis? Imaging in Medicine 2013;5(4):299-301.

Guida 2015

Guida S, Longo C, Casari A, Ciardo S, Manfredini M, Reggiani C, et al. Update on the use of confocal microscopy in melanoma and non-melanoma skin cancer. Giornale Italiano di Dermatologia e Venereologia 2015;150(5):547-63.

Guitera 2009

* Guitera P, Pellacani G, Longo C, Seidenari S, Avramidis M, Menzies SW. In vivo reflectance confocal microscopy enhances secondary evaluation of melanocytic lesions. Journal of Investigative Dermatology 2009;129(1):131-8. [Other: ER4:15465945; PubMed: 18633444]

Guitera 2010

Guitera P, Pellacani G, Crotty KA, Scolyer RA, Li L-X, Bassoli S, 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. Journal of Investigative Dermatology 2010;130(8):2080-91.

Guitera 2013

Guitera P, Moloney FJ, Menzies SW, Stretch JR, Quinn MJ, Hong A, et al. Improving management and patient care in lentigo maligna by mapping with in vivo confocal microscopy. JAMA Dermatology 2013;149(6):692-8.

Haenssle 2006

Haenssle HA, Krueger U, Vente C, Thoms K-M, Bertsch HP, Zutt M, et al. Results from an observational trial: digital epiluminescence microscopy follow-up of atypical nevi increases the sensitivity and the chance of success of conventional dermoscopy in detecting melanoma. Journal of Investigative Dermatology 2006;126(5):980-5.

Hennessy 2010

Hennessy R, Jacques S, Pellacani G, Gareau D. Clinical feasibility of rapid confocal melanoma feature detection. Photonic Therapeutics and Diagnostics Vi 2010.

Hofmann-Wellenhof 2008

Hofmann-Wellenhof R, Langsenlehner U, Richtig E, Koller S, Weger W, Ahlgrimm-Siess V, et al. In vivo confocal laser scanning microscopy of melanocytic skin tumours: diagnostic applicability using unselected tumour images. British Journal of Dermatology 2008;158(2):329-33.

Hoogedoorn 2014

Hoogedoorn L, Peppelman M, Van Erp PE, Gerritsen MJ. The use of in vivo Reflectance Confocal Microscopy in clinical practice: Prospective differentiation of difficult to distinguish nodular basal cell carcinomas and intradermal nevi. Nederlands Tijdschrift voor Dermatologie en Venereologie 2014;24(1):48.

Hoogedoorn 2015

Hoogedoorn L, Peppelman M, Blokx WA, Erp PE, Gerritsen M-JP. Prospective differentiation of clinically difficult to distinguish nodular basal cell carcinomas and intradermal nevi by non-invasive Reflectance Confocal Microscopy: a case series study. Journal of the European Academy of Dermatology and Venereology 2015;29(2):330-6.

Humphrey 2006

Humphrey S, Walsh NM, Delaney L, Propperova I, Langley RG. Prognostic significance of vascularity in cutaneous melanoma: Pilot study using in vivo confocal scanning laser microscopy. Journal of Cutaneous Medicine and Surgery 2006; 10(3):122-7.

Kadouch 2015

Kadouch DJ, Schram ME, Leeflang MM, Limpens J, Spuls PI, de Rie MA. In vivo confocal microscopy of basal cell carcinoma: a systematic review of diagnostic accuracy. Journal of the European Academy of Dermatology & Venereology 2015;29(10):1890-7.

Kadouch 2015a

Kadouch DJ, Wolkerstorfer A, Elshot Y, Zupan-Kajcovski B, Crijns MB, Starink MV, et al. Treatment of Basal Cell Carcinoma

Using a One-Stop-Shop With Reflectance Confocal Microscopy: Study Design and Protocol of a Randomized Controlled Multicenter Trial. JMIR Research Protocols 2015;4(3):e109.

Koller 2011

* Koller S, Wiltgen M, Ahlgrimm-Siess V, Weger W, Hofmann-Wellenhof R, Richtig E, et al. In vivo reflectance confocal microscopy: automated diagnostic image analysis of melanocytic skin tumours. Journal of the European Academy of Dermatology and Venereology : JEADV 2011;25(5):554-8. [Other: ER4:15465979; PubMed: 20735518]

Kose 2014

Kose K, Cordova M, Duffy M, Flores ES, Brooks DH, Rajadhyaksha M. Video-mosaicing of reflectance confocal images for examination of extended areas of skin in vivo. British Journal of Dermatology 2014;171(5):1239-41.

Langley 2001

Langley RG, Rajadhyaksha M, Dwyer PJ, Sober AJ, Flotte TJ, Anderson RR. Confocal scanning laser microscopy of benign and malignant melanocytic skin lesions in vivo. Journal of the American Academy of Dermatology 2001;45(3):365-76.

Langley 2006

Langley RG, Burton E, Walsh N, Propperova I, Murray SJ. In vivo confocal scanning laser microscopy of benign lentigines: Comparison to conventional histology and in vivo characteristics of lentigo maligna. Journal of the American Academy of Dermatology 2006;55(1):88-97.

Langley 2007

* Langley RG, Walsh N, Sutherland AE, Propperova I, Delaney L, Morris SF, 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. [Other: ER4:15465985; PubMed: 17912001]

Losi 2014

Losi A, Longo C, Cesinaro AM, Benati E, Witkowski A, Guitera P, et al. Hyporeflective pagetoid cells: a new clue for amelanotic melanoma diagnosis by reflectance confocal microscopy. British Journal of Dermatology 2014;171(1):48-54.

Lovatto 2015

Lovatto L, Carrera C, Salerni G, Alos L, Malvehy J, Puig S. In vivo reflectance confocal microscopy of equivocal melanocytic lesions detected by digital dermoscopy follow-up. Journal of the European Academy of Dermatology and Venereology : JEADV 2015;29(10):1918-25. [Other: ER4:25012311; PubMed: 25752663]

Maier 2013

Maier T, Sattler EC, Braun-Falco M, Korting HC, Ruzicka T, Berking C. Reflectance confocal microscopy in the diagnosis of partially and completely amelanotic melanoma: report on seven cases. Journal of the European Academy of Dermatology & Venereology 2013;27(1):e42-52.

Malvehy 2012

Malvehy J, Hanke-Martinez M, Costa J, Salerni G, Carrera C, Puig S. Semiology and Pattern Analysis in Nonmelanocytic Lesions. In: Hofmann-Wellenhof Rainer, Pellacani Giovanni, Malvehy Joseph, Soyer Hans Peter, editors(s). Reflectance Confocal Microscopy for Skin Diseases. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012:239-52.

Menge 2016

Menge TD, Hibler BP, Cordova MA, Nehal KS, Rossi AM. Concordance of handheld reflectance confocal microscopy (RCM) with histopathology in the diagnosis of lentigo maligna (LM): A prospective study. Journal of the American Academy of Dermatology 2016;74(6):1114-20.

Miller 2011

Miller JH, Chrisler WB, Wang X, Sowa MB. Confocal microscopy for modeling electron microbeam irradiation of skin. Radiation & Environmental Biophysics 2011;50(3):365-9.

Nobre 2011

Nobre Moura F, Dalle S, Depaepe L, Durupt F, Balme B, Thomas L. Melanoma: early diagnosis using in vivo reflectance confocal microscopy. Clinical & Experimental Dermatology 2011;36(2):209-11.

Pellacani 2005

Pellacani G, Cesinaro AM, Seidenari S. Reflectance-mode confocal microscopy of pigmented skin lesions--improvement in melanoma diagnostic specificity. Journal of the American Academy of Dermatology 2005;53(6):979-85.

Pellacani 2007

* Pellacani G, Guitera P, Longo C, Avramidis M, Seidenari S, Menzies S. The impact of in vivo reflectance confocal microscopy for the diagnostic accuracy of melanoma and equivocal melanocytic lesions. Journal of Investigative Dermatology 2007;127(12):2759-65. [Other: ER4:15466047; PubMed: 17657243]

Pellacani 2008

Pellacani G, Longo C, Malvehy J, Puig S, Carrera C, Segura S, et al. In vivo confocal microscopic and histopathologic correlations of dermoscopic features in 202 melanocytic lesions. Archives of Dermatology 2008;144(12):1597-608.

Pellacani 2009

Pellacani G, Vinceti M, Bassoli S, Braun R, Gonzalez S, Guitera P, et al. Reflectance confocal microscopy and features of melanocytic lesions: an internet-based study of the reproducibility of terminology. Archives of Dermatology 2009; 145(10):1137-43.

Pellacani 2012

* Pellacani G, Farnetani F, Gonzalez S, Longo C, Cesinaro AM, Casari A, et al. In vivo confocal microscopy for detection and grading of dysplastic nevi: a pilot study. Journal of the American Academy of Dermatology 2012; 66(3):e109-21. [Other: ER4:15466043; PubMed: 21742408]

Peppelman 2013

Peppelman M, Wolberink EA, Blokx WA, van de Kerkhof PC, van Erp PE, Gerritsen MJ. In vivo diagnosis of basal cell carcinoma subtype by reflectance confocal microscopy. Dermatology 2013;227(3):255-62.

Peppelman 2015

Peppelman M, Nguyen KP, Hoogedoorn L, van Erp PE, Gerritsen MJ. Reflectance confocal microscopy: non-invasive distinction between actinic keratosis and squamous cell carcinoma. Journal of the European Academy of Dermatology & Venereology 2015;29(7):1302-9.

Peppelman 2016

Peppelman M, Nguyen KP, Alkemade HA, Maessen-Visch B, Hendriks JC, van Erp PE, et al. Diagnosis of basal cell carcinoma by reflectance confocal microscopy: study design and protocol of a randomized controlled multicenter trial. JMIR Research Protocols 2016;5(2):e114.

Puig 2012

Puig S, Di Giacomo TB, Serra D, Cabrini F, Alos L, Palou J, et al. Reflectance confocal microscopy of blue nevus. European Journal of Dermatology 2012;22(4):552-3.

Pupelli 2013

* Pupelli G, Longo C, Veneziano L, Cesinaro AM, Ferrara G, Piana S, et al. Small-diameter melanocytic lesions: morphological analysis by means of in vivo confocal microscopy. British Journal of Dermatology 2013;168(5):1027-33. [Other: ER4:15466070; PubMed: 23301553]

Reggiani 2015

Reggiani C, Manfredini M, Mandel VD, Farnetani F, Ciardo S, Bassoli S, et al. Update on non-invasive imaging techniques in early diagnosis of non-melanoma skin cancer. Giornale Italiano di Dermatologia e Venereologia 2015;150(4):393-405.

Rishpon 2009

Rishpon A, Kim N, Scope A, Porges L, Oliviero MC, Braun RP, et al. Reflectance confocal microscopy criteria for squamous cell carcinomas and actinic keratoses. Archives of Dermatology 2009;145(7):766-72.

Röwert-Huber 2007

Röwert-Huber J, Patel MJ, Forschner T, Ulrich C, Eberle J, Kerl H, et al. Actinic keratosis is an early in situ squamous cell carcinoma: a proposal for reclassification. British Journal of Dermatology 2007;156:8-12.

Salerni 2011

Salerni G, Lovatto L, Carrera C, Palou J, Alos L, Puig-Butille JA, et al. Correlation among dermoscopy, confocal reflectance microscopy, and histologic features of melanoma and basal cell carcinoma collision tumor. Dermatologic Surgery 2011; 37(2):275-9.

Scope 2009

Scope A, Mecca PS, Marghoob AA. skINsight lessons in reflectance confocal microscopy: rapid diagnosis of pigmented basal cell carcinoma. Archives of Dermatology 2009;145(1):106-7.

Scope 2014

Scope A, Longo C. Recognizing the benefits and pitfalls of reflectance confocal microscopy in melanoma diagnosis. Dermatology Practical & Conceptual 2014;4(3):67-71.

Segura 2009

* Segura S, Puig S, Carrera C, Palou J, Malvehy J. Development of a two-step method for the diagnosis of melanoma by reflectance confocal microscopy. Journal of the American Academy of Dermatology 2009;61(2):216-29. [Other: ER4:20569494; PubMed: 19406506]

Soyer 2013

Soyer HP, Prow TW. Reflectance confocal microscopy in the diagnosis of nodular skin lesions. British Journal of

Dermatology 2013;169(1):4.

Stanganelli 2015

* Stanganelli I, Longo C, Mazzoni L, Magi S, Medri M, Lanzanova G, et al. Integration of reflectance confocal microscopy in sequential dermoscopy follow-up improves melanoma detection accuracy. British Journal of Dermatology 2015;172(2):365-371. [Other: ER4:20569496; PubMed: 25154446]

Steiner 1992

Steiner A, Pehamberger H, Binder M, Wolff K. Pigmented Spitz nevi: improvement of the diagnostic accuracy by epiluminescence microscopy. Journal of the American Academy of Dermatology 1992;27(5 Pt 1):697-701.

Stephens 2013

Stephens A, Fraga-Braghiroli N, Oliviero M, Rabinovitz H, Scope A. Spoke wheel-like structures in superficial basal cell carcinoma: a correlation between dermoscopy, histopathology, and reflective confocal microscopy. Journal of the American Academy of Dermatology 2013;69(5):e219-21.

Stevenson 2013

Stevenson AD, Mickan S, Mallett S, Ayya M. Systematic review of diagnostic accuracy of reflectance confocal microscopy for melanoma diagnosis in patients with clinically equivocal skin lesions. Dermatology Practical & Conceptual 2013;3(4):19-27.

Tannous 2009

Tannous Z, Al-Arashi M, Shah S, Yaroslavsky AN. Delineating melanoma using multimodal polarized light imaging. Lasers in Surgery & Medicine 2009;41(1):10-6.

Willard 2011

Willard K, Warschaw KE, Swanson DL. Use of reflectance confocal microscopy to differentiate hidrocystoma from basal cell carcinoma. Dermatologic Surgery 2011;37(3):392-4.

Xiong 2016

Xiong YD, Ma S, Li X, Zhong X, Duan C, Chen Q. A meta-analysis of reflectance confocal microscopy for the diagnosis of malignant skin tumours. Journal of the European Academy of Dermatology & Venereology 2016;30(8):1295-302.

Yelamos 2016

Yelamos O, Nehal KS. Integrating clinical information, dermoscopy and reflectance confocal microscopy to improve the diagnostic accuracy and confidence of amelanotic and lightly pigmented melanomas. British Journal of Dermatology 2016; 175(6):1147-8.

Studies awaiting classification

Borsari 2016

Borsari S, Pampena R, Lallas A, Kyrgidis A, Moscarella E, Benati E, et al. Clinical indications for use of reflectance confocal microscopy for skin cancer diagnosis. JAMA Dermatology 2016;152(10):1093-98. [PubMed: 27580185]

Guitera 2016

Guitera P, Menzies SW, Argenziano G, Longo C, Losi A, Drummond M, et al. Dermoscopy and in vivo confocal microscopy are complementary techniques for diagnosis of difficult amelanotic and light-coloured skin lesions. British Journal of Dermatology 2016;175(6):1311-19. [PubMed: 27177158]

Jain 2017

Jain M, Pulijal SV, Rajadhyaksha M. The bedside diagnostic accuracy of a novice reflectance confocal microscopy reader for skin cancer detection in vivo in real-time: understanding challenges and potential pitfalls. In: Alfano RR, Demos SG, editors(s). Proceedings of SPIE. Optical Biopsy XV: Toward Real-Time Spectroscopic Imaging and Diagnosis edition. Vol. 10060. March 24, 2017. [DOI: 10.1117/12.2255685]

Ludzik 2016

Ludzik J, Witkowski AM, Roterman-Konieczna I, Bassoli S, Farnetani F, Pellacani G. Improving diagnostic accuracy of dermoscopically equivocal pink cutaneous lesions with reflectance confocal microscopy in telemedicine settings: double reader concordance evaluation of 316 cases. PloS One 2016;11(9):e0162495. [PubMed: 27606812]

Ongoing studies

Other references

Additional references

Alam 2001

Alam M, Ratner D. Cutaneous squamous-cell carcinoma. New England Journal of Medicine 2001;344(13):975-83. [PubMed: 11274625]

Argenziano 1998

Argenziano G, Fabbrocini G, Carli P, De Giorgi V, Sammarco E, Delfino M. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Archives of Dermatology 1998;134(12):1563-70. [PubMed: 9875194]

Argenziano 2012

Argenziano G, Albertini G, Castagnetti F, De Pace B, Di Lernia V, Longo C, et al. Early diagnosis of melanoma: what is the impact of dermoscopy? Dermatologic Therapy 2012;25(5):403-9. [PubMed: 23046019]

Baldursson 1993

Baldursson B, Sigurgeirsson B, Lindelof B. Leg ulcers and squamous cell carcinoma. An epidemiological study and a review of the literature. Acta Dermato-Venereologica 1993;73(3):171-4. [PubMed: 8105611]

Bath-Hextall 2007a

Bath-Hextall F, Leonardi-Bee J, Smith C, Meal A, Hubbard R. Trends in incidence of skin basal cell carcinoma. Additional evidence from a UK primary care database study. International Journal of Cancer 2007;121(9):2105-8. [PubMed: 17640064]

Bath-Hextall 2007b

Bath-Hextall Fiona J, Perkins W, Bong J, Williams HC. Interventions for basal cell carcinoma of the skin. Cochrane Database of Systematic Reviews 2007, Issue 1. Art. No.: CD003412 DOI: 10.1002/14651858.CD003412.pub2.

Bath-Hextall 2014

Bath-Hextall F, Ozolins M, Armstrong SJ, Colver GB, Perkins W, Miller PS, et al. Surgical excision versus imiquimod 5% cream for nodular and superficial basal-cell carcinoma (SINS): a multicentre, non-inferiority, randomised controlled trial. Lancet Oncology 2014;15(1):96-105. [PubMed: 24332516]

Batra 2002

Batra RS, Kelley LC. A risk scale for predicting extensive subclinical spread of nonmelanoma skin cancer. Dermatologic Surgery 2002;28(2):107-12; discussion 112. [PubMed: 11860418]

Betti 2017

Betti R, Moneghini L, Mapelli ET, Bulfamante G, Cerri A. Growth rate of different basal cell carcinoma subtypes. European Journal of Dermatology 2017;27(5):544-45. [PubMed: 29084641]

Bossuyt 2015

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527. [DOI: 10.1136/bmj.h5527]

CCAAC Network 2008

Cancer Council Australia & Australian Cancer Network. Basal Cell Carcinoma, Squamous Cell Carcinoma (and related lesions) - a guide to clinical management in Australia. www.cancer.org.au/content/pdf/HealthProfessionals/ClinicalGuidelines/Basal_cell_carcinoma_Squamous_cell_carcinoma_Guic Final_with_Corrigendums.pdf 2008 (accessed 19 May 2015).

Chao 2013

Chao D, London Cancer North and East. London Cancer, Guidelines for Cutaneous Malignant Melanoma Management August 2014. www.londoncancer.org/media/76373/london-cancer-melanoma-guidelines-2013-v0.pdf (accessed 25 February 2015).

Chowdri 1996

Chowdri NA, Darzi MA. Postburn scar carcinomas in Kashmiris. Burns 1996;22(6):477-82. [PubMed: 8884010]

Chu 2006

Chu H, Cole S. Bivariate meta-analysis for sensitivity and specificity with sparse data: a generalized linear mixed model approach (letter to the Editor). Journal of Clinical Epidemiology 2006;59(12):1331. [PubMed: 17098577]

Chuchu 2018a

Chuchu N, Dinnes J, Takwoingi Y, Matin RN, Bayliss SE, Davenport C, et al. The use of teledermatology for the diagnosis of skin cancer in adults. Cochrane Database of Systematic Reviews (in press).

Chuchu 2018b

Chuchu N, Takwoingi Y, Dinnes J, Matin RN, Bassett O, Moreau JF, et al. Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma. Cochrane Database of Systematic Reviews (in press).

Dabski 1986

Dabski K, Stoll HL Jr, Milgrom H. Squamous cell carcinoma complicating late chronic discoid lupus erythematosus. Journal of Surgical Oncology 1986;32(4):233-7. [PubMed: 3736067]

Deeks 2005

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. Journal of Clinical Epidemiology 2005;58(9):882-93. [PubMed: 16085191]

Deeks 2013

Deeks JJ, Bossuyt PM, Gatsonis C (editors). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2013. Available from srdta.cochrane.org.

Dinnes 2018a

Dinnes J, Deeks JJ, Saleh D, Chuchu N, Bayliss SE, Patel L, et al. Reflectance confocal microscopy for the diagnosis of cutaneous melanoma in adults. Cochrane Database of Systematic Reviews (in press).

Dinnes 2018b

Dinnes J, Deeks JJ, Chuchu N, Matin RN, Wong KY, Aldridge RB, et al. Visual examination and dermoscopy, alone or in combination, for the diagnosis of keratinocyte skin cancers in adults. Cochrane Database of Systematic Reviews (in press).

Dinnes 2018c

Dinnes J, Bamber J, Chuchu N, Bayliss SE, Takwoingi Y, Davenport C, et al. High frequency ultrasound for the diagnosis of skin cancer in adults. Cochrane Database of Systematic Reviews (in press).

Drew 2017

Drew BA, Karia PS, Mora AN, Liang CA, Schmults CD. Treatment patterns, outcomes, and patient satisfaction of primary epidermally limited nonmelanoma skin cancer. Dermatologic Surgery 2017;43(12):1423-30. [DOI: 10.1097/DSS.0000000000001225; PubMed: 28661992]

Efron 1983

Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. Journal of the American Statistical Association 1983;78(382):316-331. [DOI: 10.1080/01621459.1983.10477973]

Esteva 2017

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542(7639):115-118. [PubMed: 28117445]

Fasching 1989

Fasching MC, Meland NB, Woods JE, Wolff BG. Recurrent squamous-cell carcinoma arising in pilonidal sinus tract--multiple flap reconstructions. Report of a case. Diseases of the Colon & Rectum 1989;32(2):153-8. [PubMed: 2914529]

Ferrante di Ruffano 2018a

Ferrante di Ruffano L, Takwoingi Y, Dinnes J, Chuchu N, Bayliss SE, Davenport C, et al. Computer assisted diagnosis techniques (dermoscopy and spectroscopy-based) for the diagnosis of skin cancer in adults. Cochrane Database of Systematic Reviews (in press).

Ferrante di Ruffano 2018b

Ferrante di Ruffano L, Dinnes J, Deeks JJ, Chuchu N, Bayliss SE, Davenport C, et al. Optical coherence tomography for the diagnosis of skin cancer in adults. Cochrane Database of Systematic Reviews (in press).

Ferrante di Ruffano 2018c

Ferrante di Ruffano L, Dinnes J, Chuchu N, Bayliss SE, Takwoingi Y, Davenport C, et al. Exfoliative cytology for the diagnosis of basal cell carcinoma and other skin cancers in adults. Cochrane Database of Systematic Reviews (in press).

Firnhaber 2012

Firnhaber JM. Diagnosis and treatment of basal cell and squamous cell carcinoma. American Family Physician 2012; 86(2):161-168. [PubMed: 22962928]

Fitzpatrick 1975

Fitzpatrick TB. Soleil et peau. Journal de Médecine Esthétique 1975;2:33-4.

Flohil 2013

Flohil SC, van der Leest RJ, Arends LR, de Vries E, Nijsten T. Risk of subsequent cutaneous malignancy in patients with prior keratinocyte carcinoma: a systematic review and meta-analysis. European Journal of Cancer 2013;49(10):2365-75. [PubMed: 23608733]

Friedman 1985

Friedman HI, Cooper PH, Wanebo HJ. Prognostic and therapeutic use of microstaging of cutaneous squamous cell carcinoma of the trunk and extremities. Cancer 1985;56(5):1099-105. [PubMed: 4016700]

Gambichler 2015

Gambichler T, Schmid-Wendtner MH, Plura I, Kampilafkos P, Stucker M, Berking C, et al. A multicentre pilot study investigating high-definition optical coherence tomography in the differentiation of cutaneous melanoma and melanocytic naevi. Journal of the European Academy of Dermatology & Venereology 2015;29(3):537-41. [PubMed: 25073788]

Gerbert 2000

Gerbert B, Bronstone A, Maurer T, Hofmann R, Berger T. Decision support software to help primary care physicians triage skin cancer: a pilot study. Archives of Dermatology 2000;136(2):187-92. [PubMed: 10677094]

Gordon 2013

Gordon R. Skin cancer: an overview of epidemiology and risk factors. Seminars in Oncology Nursing 2013;29(3):160-9. [PubMed: 23958214]

Gorlin 2004

Gorlin RJ. Nevoid basal cell carcinoma (Gorlin) syndrome. Genetics in Medicine 2004;6(6):530-9. [PubMed: 15545751]

Grachtchouk 2011

Grachtchouk M, Pero J, Yang SH, Ermilov AN, Michael LE, Wang A, et al. Basal cell carcinomas in mice arise from hair follicle stem cells and multiple epithelial progenitor populations. Journal of Clinical Investigation 2011;121(5):1768-81. [PubMed: 21519145]

Green 1988

Green A, Leslie D, Weedon D. Diagnosis of skin cancer in the general population: clinical accuracy in the Nambour survey. Medical Journal of Australia 1988;148(9):447-50. [PubMed: 3283506]

Griffin 2016

Griffin LL, Ali FR, Lear JT. Non-melanoma skin cancer. Clinical Medicine 2016;16(1):62-5. [PubMed: 26833519]

Griffiths 2005

Griffiths RW, Suvarna SK, Stone J. Do basal cell carcinomas recur after complete conventional surgical excision? British Journal of Plastic Surgery 2005;58(6):795-805. [PubMed: 16086990]

Grob 1998

Grob JJ, Bonerandi JJ. The 'ugly duckling' sign: identification of the common characteristics of nevi in an individual as a basis for melanoma screening. Archives of Dermatology 1998;134(1):103-4. [PubMed: 9449921]

Guitera 2017

Guitera P, Menzies S, Chamberlain A, Soyer P, Cancer Council Australia Melanoma Guidelines Working Party. What is the role of confocal microscopy in melanoma diagnosis? www.wiki.cancer.org.au/australia/Guidelines:Melanoma/Confocal_microscopy (accessed prior to 28 November 2017).

Haenssle 2010

Haenssle HA, Korpas B, Hansen-Hagge C, Buhl T, Kaune KM, Rosenberger A, et al. Seven-point checklist for dermatoscopy: performance during 10 years of prospective surveillance of patients at increased melanoma risk. Journal of the American Academy of Dermatology 2010;62(5):785-93. [PubMed: 20226567]

Harland 2000

Harland CC, Kale SG, Jackson P, Mortimer PS, Bamber JC. Differentiation of common benign pigmented skin lesions from melanoma by high-resolution ultrasound. British Journal of Dermatology 2000;143(2):281-9. [PubMed: 10951134]

Hartevelt 1990

Hartevelt MM, Bavinck JN, Kootte AM, Vermeer BJ, Vandenbroucke JP. Incidence of skin cancer after renal transplantation in The Netherlands. Transplantation 1990;49(3):506-9. [PubMed: 2316011]

Hauschild 2014

Hauschild A, Chen SC, Weichenthal M, Blum A, King HC, Goldsmith J, et al. To excise or not: impact of MelaFind on German dermatologists' decisions to biopsy atypical lesions. Journal der Deutschen Dermatologischen Gesellschaft 2014; 12(7):606-14. [PubMed: 24944011]

Jansen 2017

Jansen MHE, Mosterd K, Arits AHMM, Roozeboom MH, Sommer A, Essers BAB, et al. Five-year results of a randomized controlled trial comparing effectiveness of photodynamic therapy, topical imiquimod, and topical 5-fluorouracil in patients with superficial basal cell carcinoma. Journal of Investigative Dermatology 2018;138(3):527-33. [DOI: 10.1016/j.jid.2017.09.033; PubMed: 29045820]

Jensen 1999

Jensen P, Hansen S, Moller B, Leivestad T, Pfeffer P, Geiran O, et al. Skin cancer in kidney and heart transplant recipients and different long-term immunosuppressive therapy regimens. Journal of the American Academy of Dermatology 1999;40(2

Pt 1):177-86. [PubMed: 10025742]

Kao 1986

Kao GF. Carcinoma arising in Bowen's disease. Archives of Dermatology 1986;122(10):1124-6. [PubMed: 3767398]

Karimkhani 2015

Karimkhani C, Boyers LN, Dellavalle RP, Weinstock MA. It's time for "keratinocyte carcinoma" to replace the term "nonmelanoma skin cancer". Journal of the American Academy of Dermatology 2015;72(1):186-7. [PubMed: 25497921]

Kelleners-Smeets 2017

Kelleners-Smeets NW, Mosterd K, Nelemans PJ. Treatment of low-risk basal cell carcinoma. Journal of Investigative Dermatology 2017;137(3):539-40. [PubMed: 28235442]

Kim 2014

Kim DD, Tang JY, Ioannidis JP. Network geometry shows evidence sequestration for medical vs. surgical practices: treatments for basal cell carcinoma. Journal of Clinical Epidemiology 2014;67(4):391-400. [PubMed: 24491794]

Kittler 2002

Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy [Review]. Lancet Oncology 2002; 3(3):159-65. [PubMed: 11902502]

Kleinerman 2012

Kleinerman R, Whang TB, Bard RL, Marmur ES. Ultrasound in dermatology: principles and applications. Journal of the American Academy of Dermatology 2012;67(3):478-87. [PubMed: 22285673]

Lansbury 2010

Lansbury L, Leonardi-Bee J, Perkins W, Goodacre T, Tweed JA, Bath-Hextall FJ. Interventions for non-metastatic squamous cell carcinoma of the skin. Cochrane Database of Systematic Reviews 2010, Issue 4. Art. No.: CD007869 DOI: 10.1002/14651858.CD007869.pub2.

Lansbury 2013

Lansbury L, Bath-Hextall F, Perkins W, Stanton W, Leonardi-Bee J. Interventions for non-metastatic squamous cell carcinoma of the skin: systematic review and pooled analysis of observational studies. BMJ 2013;347:f6153. [PubMed: 24191270]

Lear 1997

Lear JT, Tan BB, Smith AG, Bowers W, Jones PW, Heagerty AH, et al. Risk factors for basal cell carcinoma in the UK: case-control study in 806 patients. Journal of the Royal Society of Medicine 1997;90(7):371-4. [PubMed: 9290417]

Lear 2012

Lear JT. Oral hedgehog-pathway inhibitors for basal-cell carcinoma. New England Journal of Medicine 2012;366(23):2225-6. [PubMed: 22670909]

Lederman 1985

Lederman JS, Sober AJ. Does biopsy type influence survival in clinical stage I cutaneous melanoma? Journal of the American Academy of Dermatology 1985;13(6):983-7. [PubMed: 4078105]

Lees 1991

Lees VC, Briggs JC. Effect of initial biopsy procedure on prognosis in stage I invasive cutaneous malignant melanoma: review of 1086 patients. British Journal of Surgery 1991;78(9):1108-10. [PubMed: 1933198]

Lister 1997

Lister RK, Black MM, Calonje E, Burnand KG. Squamous cell carcinoma arising in chronic lymphoedema. British Journal of Dermatology 1997;136(3):384-7. [PubMed: 9115922]

Lo 1991

Lo JS, Snow SN, Reizner GT, Mohs FE, Larson PO, Hruza GJ. Metastatic basal cell carcinoma: report of twelve cases with a review of the literature. Journal of the American Academy of Dermatology 1991;24(5 Pt 1):715-9. [PubMed: 1869642]

Lomas 2012

Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. British Journal of Dermatology 2012;166(5):1069-80. [PubMed: 22251204]

London Cancer Alliance 2013

London Cancer Alliance. Guidelines for the treatment and referral of squamous cell carcinoma (SCC) of the skin. www.londoncancer.org/media/76391/london-cancer-scc-guidelines-2013-v1.0.pdf (accessed prior to 25 May 2018).

Longo 2011

Longo C, Rito C, Beretti F, Cesinaro AM, Piñeiro-Maceira J, Seidenari S, et al.. De novo melanoma and melanoma arising from pre-existing nevus: in vivo morphologic differences as evaluated by confocal microscopy. Journal of the American Academy of Dermatology 2011;65(3):604-14. [PubMed: 21715047]

MacKie 1985

MacKie RM, English J, Aitchison TC, Fitzsimons CP, Wilson P. The number and distribution of benign pigmented moles (melanocytic naevi) in a healthy British population. British Journal of Dermatology 1985;113(2):167-74. [PubMed: 4027184]

MacKie 1990

MacKie RM. Clinical recognition of early invasive malignant melanoma. BMJ 1990;301(6759):1005-6. [PubMed: 2249043]

Madan 2010

Madan V, Lear JT, Szeimies RM. Non-melanoma skin cancer. Lancet 2010;375(9715):673-85. [PubMed: 20171403]

Maia 1995

Maia M, Proenca NG, de Moraes JC. Risk factors for basal cell carcinoma: a case-control study. Revista de Saude Publica 1995;29(1):27-37. [PubMed: 8525311]

Maloney 1996

Maloney ME. Arsenic in Dermatology. Dermatologic Surgery 1996;22(3):301-4. [PubMed: 8599743]

Malvehy 2014

Malvehy J, Hauschild A, Curiel-Lewandrowski C, Mohr P, Hofmann-Wellenhof R, Motley R, et al. Clinical performance of the Nevisense system in cutaneous melanoma detection: an international, multicentre, prospective and blinded clinical trial on efficacy and safety. British Journal of Dermatology 2014;171(5):1099-107. [PubMed: 24841846]

Marsden 2010

Marsden JR, Newton-Bishop JA, Burrows L, Cook M, Corrie PG, Cox NH, et al. BAD Guidelines: Revised UK guidelines for the management of cutaneous melanoma 2010. British Journal of Dermatology 2010;163(2):238-56. [PubMed: 20608932]

McCormack 1997

McCormack CJ, Kelly JW, Dorevitch AP. Differences in age and body site distribution of the histological subtypes of basal cell carcinoma. A possible indicator of differing causes. Archives of Dermatology 1997;133(5):593-6. [PubMed: 9158412]

McCusker 2014

McCusker M, Basset-Seguin N, Dummer R, Lewis K, Schadendorf D, Sekulic A, et al. Metastatic basal cell carcinoma: Prognosis dependent on anatomic site and spread of disease. European Journal of Cancer 2014;50(4):774-83. [PubMed: 24412051]

Moeckelmann 2018

Moeckelmann N, Ebrahimi A, Dirven R, Liu J, Low TH, Gupta R, et al. Analysis and comparison of the 8th Edition American Joint Committee on Cancer (AJCC) nodal staging system in cutaneous and oral squamous cell cancer of the head and neck. Annals of Surgical Oncology 2018;25(6):1730-6. [DOI: 10.1245/s10434-018-6340-x; PubMed: 29352431]

Mogensen 2007

Mogensen M, Jemec GB. Diagnosis of nonmelanoma skin cancer/keratinocyte carcinoma: a review of diagnostic accuracy of nonmelanoma skin cancer diagnostic tests and technologies.. Dermatologic Surgery 2007;33(10):1158-74. [PubMed: 17903149]

Moncrieff 2002

Moncrieff M, Cotton S, Claridge E, Hall P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. British Journal of Dermatology 2002;146(3):448-57. [PubMed: 11952545]

Monheit 2011

Monheit G, Cognetta AB, Ferris L, Rabinovitz H, Gross K, Martini M, et al. The performance of MelaFind: a prospective multicenter study. Archives of Dermatology 2011;147(2):188-94. [PubMed: 20956633]

Motley 2009

Motley RJ, Preston PW, Lawrence CM. Multi-professional Guidelines for the management of the patient with primary cutaneous squamous cell carcinoma. www.bad.org.uk/healthcare-professionals/clinical-standards/clinical-guidelines/scc (accessed prior to 28 November 2017).

Murchie 2017

Murchie P, Amalraj Raja E, Brewster DH, Iversen L, Lee AJ. Is initial excision of cutaneous melanoma by General Practitioners (GPs) dangerous? Comparing patient outcomes following excision of melanoma by GPs or in hospital using national datasets and meta-analysis. European Journal of Cancer 2017;86:373-84. [PubMed: 29100192]

Musah 2013

Musah A, Gibson JE, Leonardi-Bee J, Cave MR, Ander EL, Bath-Hextall F. Regional variations of basal cell carcinoma incidence in the U.K. using The Health Improvement Network database (2004-10). British Journal of Dermatology 2013; 169(5):1093-9. [PubMed: 23701520]

Nascimento 2014

Nascimento MM, Shitara D, Enokihara MM, Yamada S, Pellacani G, Rezze GG. Inner gray halo, a novel dermoscopic feature for the diagnosis of pigmented actinic keratosis: clues for the differential diagnosis with lentigo maligna. Journal of the American Academy of Dermatology 2014;71(4):708715. [PubMed: 24947988]

NICE 2010

National Institute of Clinical Excellence. NICE Guidance on Cancer Services. Improving outcomes for people with skin tumours including melanoma (update). www.nice.org.uk/guidance/csgstim (accessed 19 May 2015).

NICE 2015

National Institute for Health and Clinical Excellence. Suspected cancer: recognition and referral. www.nice.org.uk/guidance/ng12 (accessed prior to 25 May 2018).

NICE 2017

NICE. Vismodegib for treating basal cell carcinoma. www.nice.org.uk/guidance/ta489 (accessed prior to 25 May 2018).

Norman 2009

Norman G, Barraclough K, Dolovich L, Price D. Iterative diagnosis. BMJ 2009;339:b3490. [DOI: 10.1136/bmj.b3490]

O'Gorman 2014

O'Gorman SM, Murphy GM. Photosensitizing medications and photocarcinogenesis. Photodermatology, Photoimmunology & Photomedicine 2014;30(1):8-14. [PubMed: 24393207]

Offidani 2002

Offidani A, Simonetti O, Bernardini M L, Alpagut A, Cellini A, Bossi G. General practitioners' accuracy in diagnosing skin cancers. Dermatology 2002;205(2):127-30. [PubMed: 12218226]

Olsen 2015

Olsen J, Themstrup L, Jemec GB. Optical coherence tomography in dermatology. Giornale Italiano di Dermatologia e Venereologia 2015;150(5):603-15. [PubMed: 26129683]

Pellacani 2014

Pellacani G, Pepe P, Casari A, Longo C. Reflectance confocal microscopy as a second-level examination in skin oncology improves diagnostic accuracy and saves unnecessary excisions: a longitudinal prospective study. British Journal of Dermatology 2014;171(5):1044-51. [PubMed: 24891083]

Rajadhyaksha 1995

Rajadhyaksha M, Grossman M, Esterowitz D, Webb RH, Anderson RR. In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. Journal of Investigative Dermatology 1995;104(6):946-52. [PubMed: 7769264]

Rajadhyaksha 2017

Rajadhyaksha M, Marghoob A, Rossi A, Halpern AC, Nehal KS. Reflectance confocal microscopy of skin in vivo: From bench to bedside. Lasers in Surgery and Medicine 2017;49(1):7-19. [PubMed: 27785781]

Rajpara 2009

Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. British Journal of Dermatology 2009;161(3):591-604. [PubMed: 19302072]

Randle 1996

Randle HW. Basal cell carcinoma. Identification and treatment of the high-risk patient. Dermatologic Surgery 1996; 22(3):255-61. [PubMed: 8599737]

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of Clinical Epidemiology 2005;58(10):982-90. [PubMed: 16168343]

Roozeboom 2012

Roozeboom MH, Arits AH, Nelemans PJ, Kelleners-Smeets NW. Overall treatment success after treatment of primary superficial basal cell carcinoma: a systematic review and meta-analysis of randomized and nonrandomized trials. British Journal of Dermatology 2012;167(4):733-56. [PubMed: 22612571]

Roozeboom 2016

Roozeboom MH, Arits AH, Mosterd K, Sommer A, Essers BA, de Rooij MJ, et al. Three-year follow-up results of

photodynamic therapy vs. imiquimod vs. fluorouracil for treatment of superficial basal cell carcinoma: a single-blind, noninferiority, randomized controlled trial. Journal of Investigative Dermatology 2016;136(8):1568-74. [PubMed: 27113429]

Royal College of Pathologists 2014

Royal College of Pathologists. Standards and datasets for reporting cancers. Dataset for the histological reporting of primary invasive cutaneous squamous cell carcinoma and regional lymph nodes. www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/G/G124_DatasetSquamous_May14.pdf (accessed 19 May 2015).

Rutjes 2005

Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. Clinical Chemistry 2005;51(8):1335-41. [PubMed: 15961549]

Rutjes 2006

Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ 2006;174(4):469-476. [PubMed: 16477057]

Sekulic 2012

Sekulic A, Migden MR, Oro AE, Dirix L, Lewis KD, Hainsworth JD, et al. Efficacy and safety of vismodegib in advanced basal-cell carcinoma. New England Journal of Medicine 2012;366(23):2171-9. [PubMed: 22670903]

Smeets 2004

Smeets NW, Krekels GA, Ostertag JU, Essers BA, Dirksen CD, Nieman FH, et al. Surgical excision vs Mohs' micrographic surgery for basal-cell carcinoma of the face: randomised controlled trial. Lancet 2004;364(9447):1766-72. [PubMed: 15541449]

Sobarun 2015

Sobarun P. Reflectance confocal microscopy. www.dermnetnz.org/topics/reflectance-confocal-microscopy/ (accessed 18 May 2017).

Sober 1979

Sober AJ, Fitzpatrick TB, Mihm MC, Wise TG, Pearson BJ, Clark WH, et al. Early recognition of cutaneous melanoma. JAMA 1979;242(25):2795-9. [PubMed: 501893]

Stratigos 2015

Stratigos A, Garbe C, Lebbe C, Malvehy J, del Marmol V, Pehamberger H, et al. Diagnosis and treatment of invasive squamous cell carcinoma of the skin: European consensus-based interdisciplinary guideline. European Journal of Cancer 2015;51(14):1989-2007. [PubMed: 26219687]

Takwoingi 2013

Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Annals of Internal Medicine 2013;158(7):544-54. [PubMed: 23546566]

Takwoingi 2017

Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Statistical Methods in Medical Research 2017;26(4):1896-911. [DOI: 10.1177/0962280215592269; PubMed: 26116616]

Thomas 1998

Thomas L, Tranchand P, Berard F, Secchi T, Colin C, Moulin G. Semiological value of ABCDE criteria in the diagnosis of cutaneous pigmented tumors. Dermatology 1998;197(1):11-7. [PubMed: 9693179]

van Loo 2014

van Loo E, Mosterd K, Krekels GA, Roozeboom MH, Ostertag JU, Dirksen CD, et al. Surgical excision versus Mohs' micrographic surgery for basal cell carcinoma of the face: A randomised clinical trial with 10 year follow-up. European Journal of Cancer 2014;50(17):3011-20. [PubMed: 25262378]

Verkouteren 2017

Verkouteren JAC, Ramdas KHR, Wakkee M, Nijsten T. Epidemiology of basal cell carcinoma: scholarly review. British Journal of Dermatology 2017;177(2):359-72. [DOI: 10.1111/bjd.15321.; PubMed: 28220485]

Walker 2006

Walker P, Hill D. Surgical treatment of basal cell carcinomas using standard postoperative histological assessment. Australasian Journal of Dermatology 2006;47(1):1-12. [PubMed: 16405477]

Walter 2012

Walter FM, Morris HC, Humphrys E, Hall PN, Prevost AT, Burrows N, et al. Effect of adding a diagnostic aid to best practice to manage suspicious pigmented lesions in primary care: randomised controlled trial. BMJ 2012;345:e4110. [PubMed:

22763392]

Wells 2012

Wells R, Gutkowicz-Krusin D, Veledar E, Toledano A, Chen SC. Comparison of diagnostic and management sensitivity to melanoma between dermatologists and MelaFind: a pilot study. Archives of Dermatology 2012;148(9):1083-4. [PubMed: 22986873]

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine 2011;155(8):529-36. [PubMed: 22007046]

WHO 2003

World Health Organisation. INTERSUN: The Global UV Project. A guide and compendium. www.who.int/uv/publications/en/Intersunguide.pdf (accessed 20 May 2015).

Williams 2017

Williams HC, Bath-Hextall F, Ozolins M, Armstrong SJ, Colver GB, Perkins W, et al. Surgery versus 5% imiquimod for nodular and superficial basal cell carcinoma: 5-year results of the SINS randomized controlled trial. Journal of Investigative Dermatology 2017;137(3):614-9. [DOI: 10.1016/j.jid.2016.10.019; PubMed: 27932240]

Wong 2017

Wong KY, Fife K, Lear JT, Price RD, Durrani AJ. Vismodegib for locally advanced periocular and orbital basal cell carcinoma: A review of 15 consecutive cases. Plastic and Reconstructive Surgery 2017;5(7):e1424.

Zak-Prelich 2004

Zak-Prelich M, Narbutt J, Sysa-Jedrzejowska A. Environmental risk factors predisposing to the development of basal cell carcinoma. Dermatologic Surgery 2004;30(2 Pt 2):248–252. [PubMed: 14871217]

Other published versions of this review

Dinnes 2015

Dinnes J, Wong KY, Gulati A, Chuchu N, Leonardi-Bee J, Bayliss SE, et al. Tests to assist in the diagnosis of keratinocyte skin cancers in adults: a generic protocol. Cochrane Database of Systematic Reviews 2015, Issue 10. Art. No.: CD011901 DOI: 10.1002/14651858.CD011901.

Dinnes 2015a

Dinnes J, Matin RN, Moreau JF, Patel L, Chan SA, Wong KY, et al. Tests to assist in the diagnosis of cutaneous melanoma in adults: a generic protocol. Cochrane Database of Systematic Reviews 2015, Issue 10. Art. No.: CD011902 DOI: 10.1002/14651858.CD011902.

Classification pending references

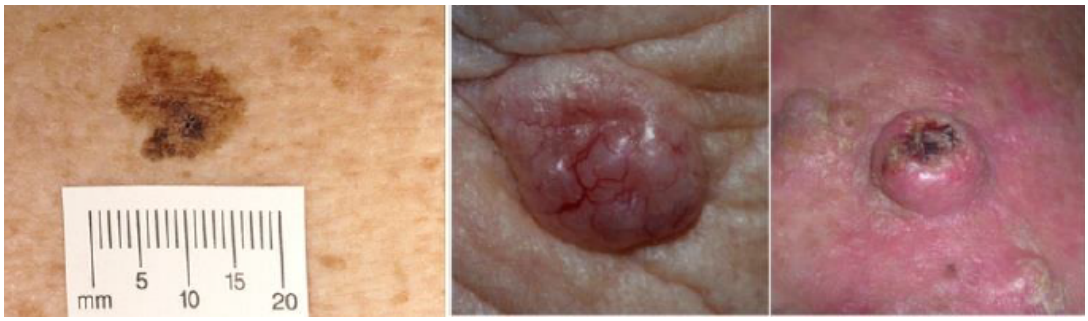
Data and analyses

Data tables by test

Test	Studies	Participants
1 BCC - any suspicious lesion	4	912
2 BCC - equivocal lesions	3	668
3 BCC - other lesion populations	4	457
4 BCC - RCM - other - Vivascope 3000	1	54
5 BCC - Dermoscopy - equivocal lesions	1	260
6 BCC - Visual inspection - other lesion populations	1	105
7 SCC - RCM - all comer	1	323
8 SCC - RCM - equivocal	1	260
9 SCC - RCM - other	2	251
10 SCC - Dermoscopy - equivocal	1	260
11 KER - RCM - all comer	2	373
12 KER - RCM - equivocal	2	360
13 KER - RCM - other	1	129
14 KER - Dermoscopy - equivocal	1	260
15 MM2 - RCM - equivocal (non-pigmented) not in melanoma review	1	260
16 BCC - RCM score at >=3 - in person	1	50
17 BCC - RCM score at NR (likely >=3) - in person	2	491
18 BCC - Guitera Two-step alg (significant chars for BCC) - image-based	1	356
19 BCC - No algorithm (significant characteristics) - in person	1	54
20 BCC - No algorithm (significant characteristics) - image-based	1	130
21 BCC - No algorithm (selected characteristics) - in person	1	122
22 BCC - No algorithm (selected characteristics) - image-based	1	152
23 BCC - No algorithm (observer diagnosis) - in person	1	318
24 BCC - No algorithm (observer diagnosis) - image-based	4	812
25 BCC - Handheld RCM - No algorithm (significant characteristics)	1	54
28 SCC - No algorithm (selected characteristics) in person	1	122
29 SCC - No algorithm (observer diagnosis) - in person	1	318
30 SCC - No algorithm (observer diagnosis) - image-based	3	712
33 KER - RCM at >=3 - in person	1	50
36 KER - No algorithm (observer diagnosis) - in person	1	318
37 KER - No algorithm (observer diagnosis) - image-based	4	812
38 KER - No algorithm (excise decision) - in person	1	318
39 KER - No algorithm (excise decision) - image-based	2	583
40 BCC - by observer - high - in person	3	545
41 BCC - by observer - high - image-based	4	908
42 BCC - by observer - low - in person	2	368
43 BCC - by observer - low - image-based	2	252
44 BCC - by observer - NR - in person	1	122
45 BCC - by observer - NR - image-based	1	260
47 SCC - by observer - low - in person	1	318
48 SCC - by observer - NR - in person	1	122
49 SCC - by observer - high - image-based	2	452
50 SCC - by observer - NR - image-based	1	260
52 KER - by observer - low - in person	2	368
53 KER - by observer - high - image-based	3	552
54 KER - by observer - low - image-based	1	100
55 KER - by observer - NR - image-based	1	260

Figures

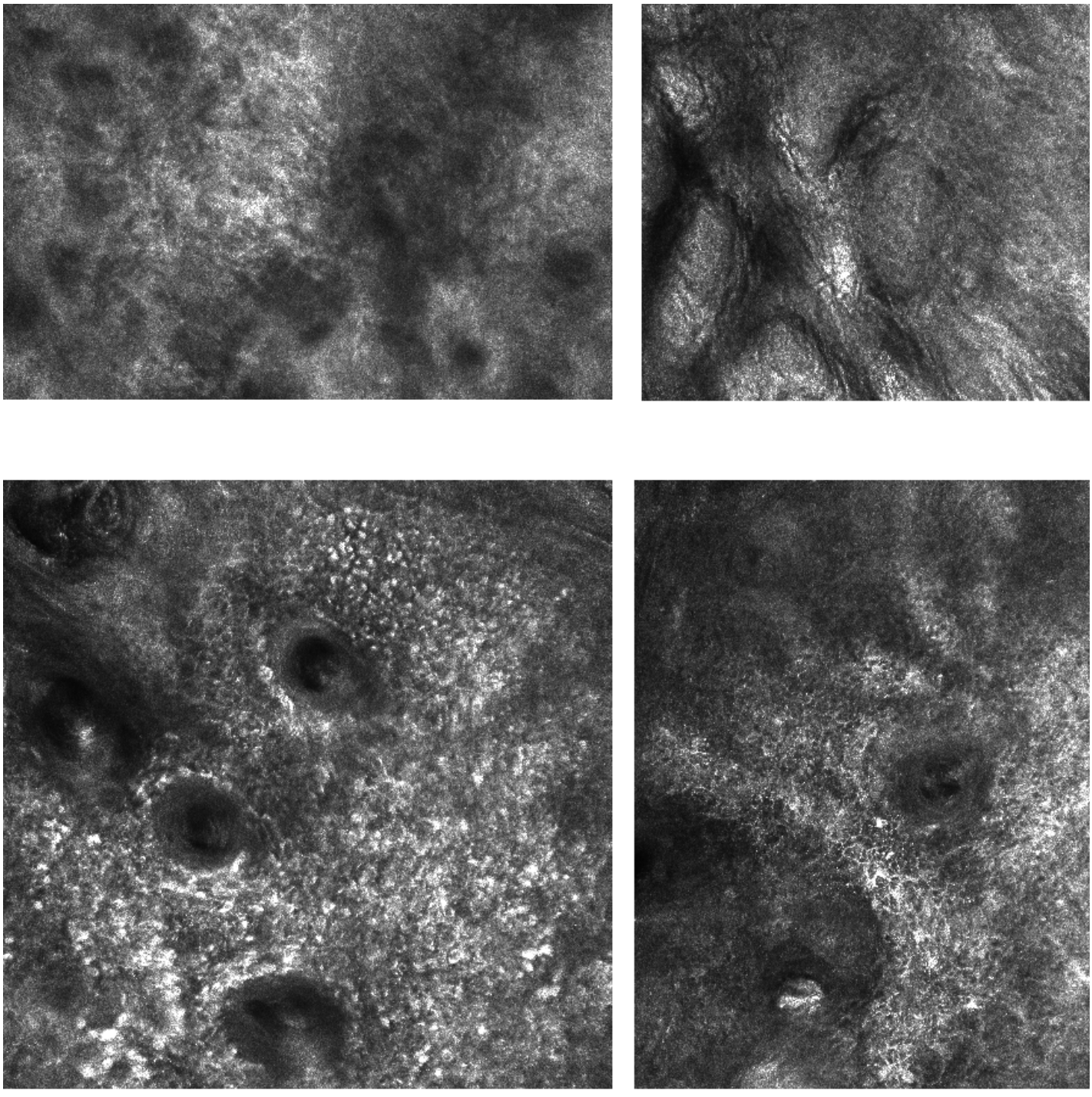
Figure 1



Caption

Sample photographs of superficial spreading melanoma (left), BCC (centre) and cSCC (right)

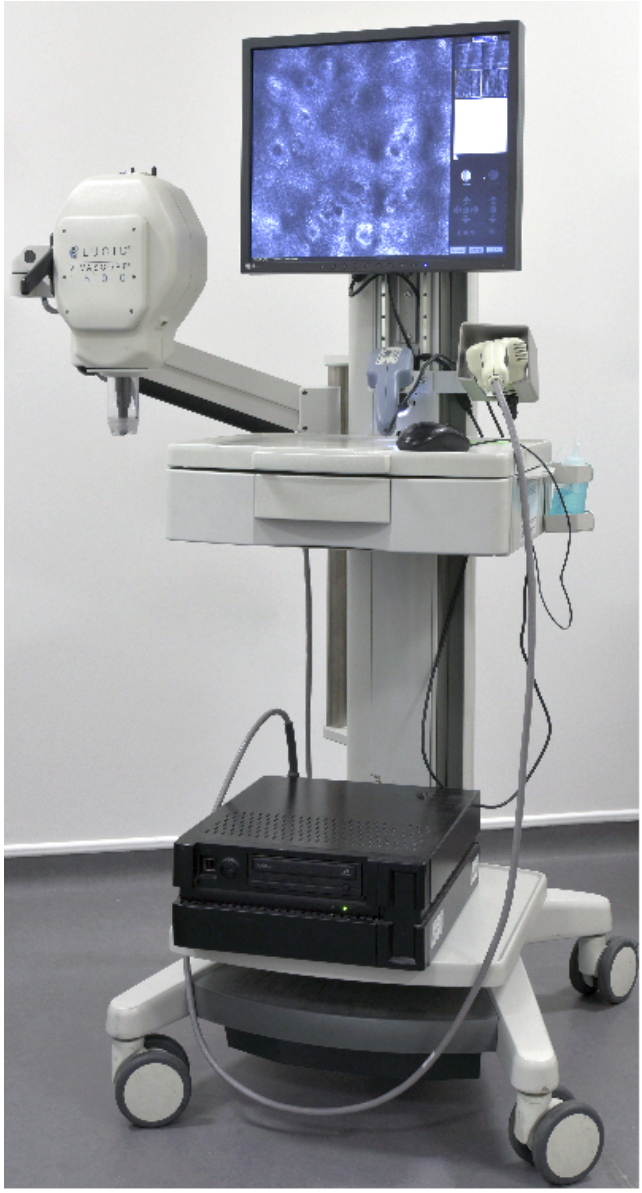
Figure 2



Caption

RCM images of normal skin (top) and of lentigo maligna (bottom)

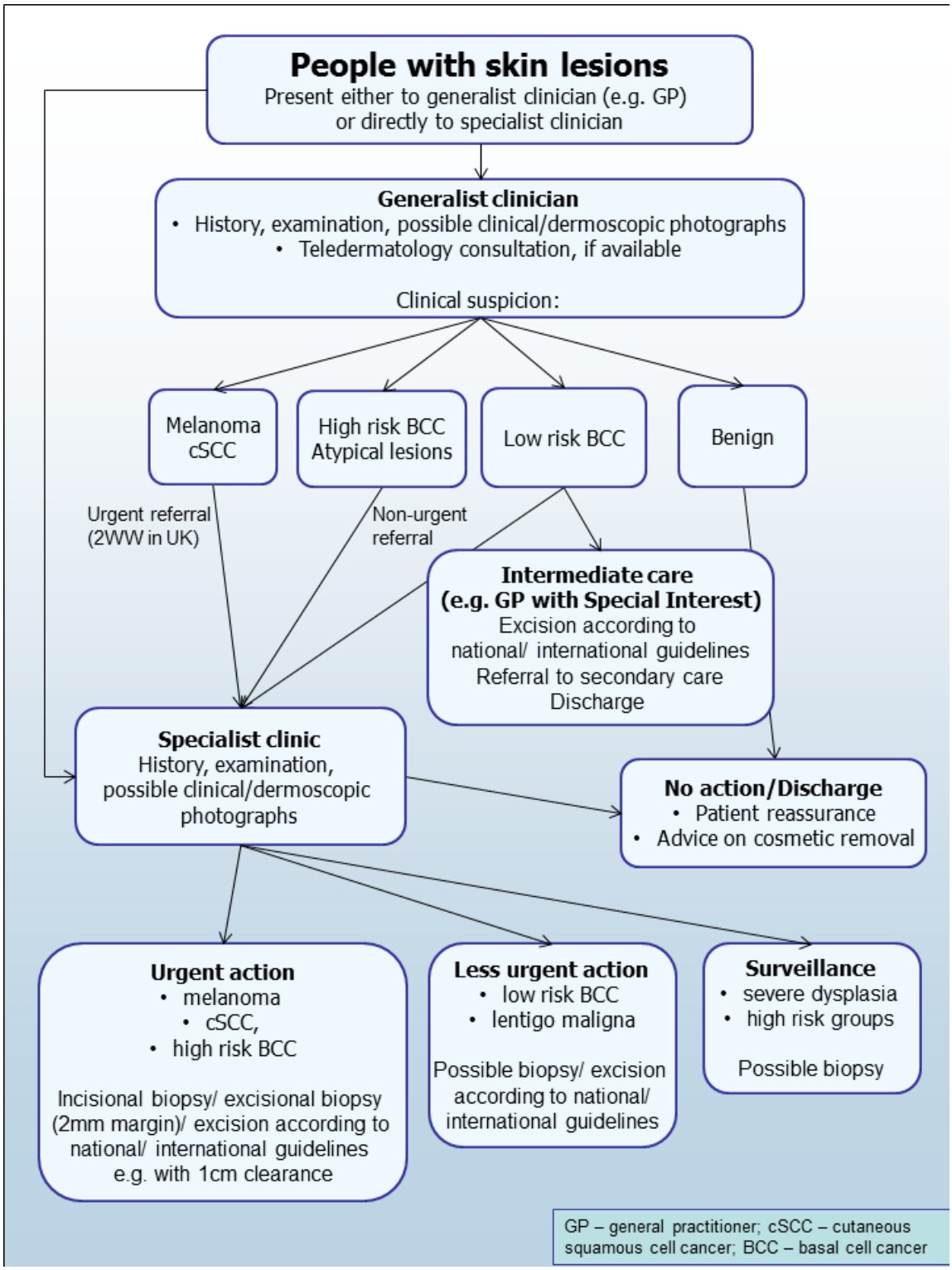
Figure 3



Caption

Caliber ID Vivascope 1500 with 3000 attachment

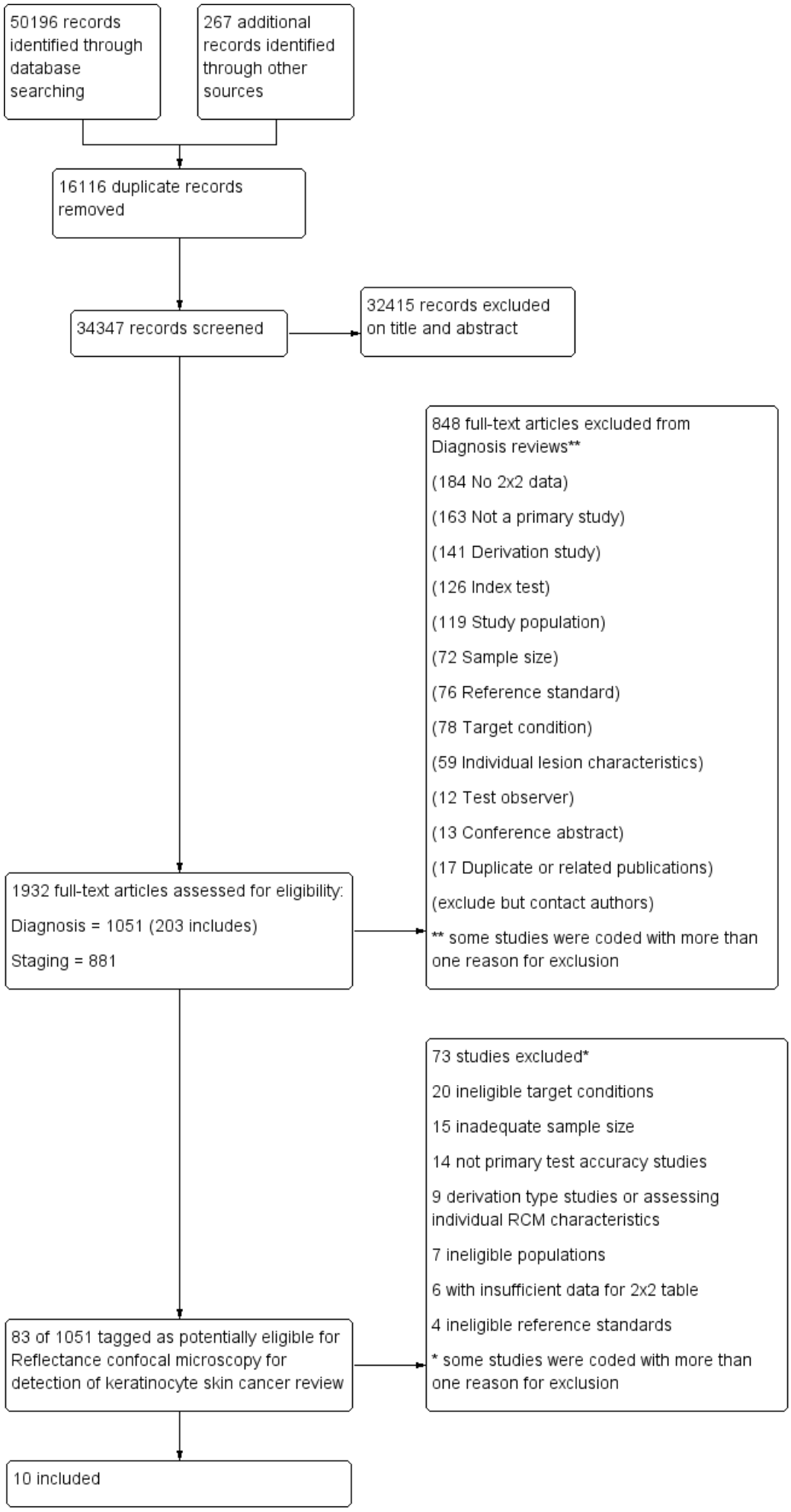
Figure 4



Caption

Current clinical pathway for people with skin lesions

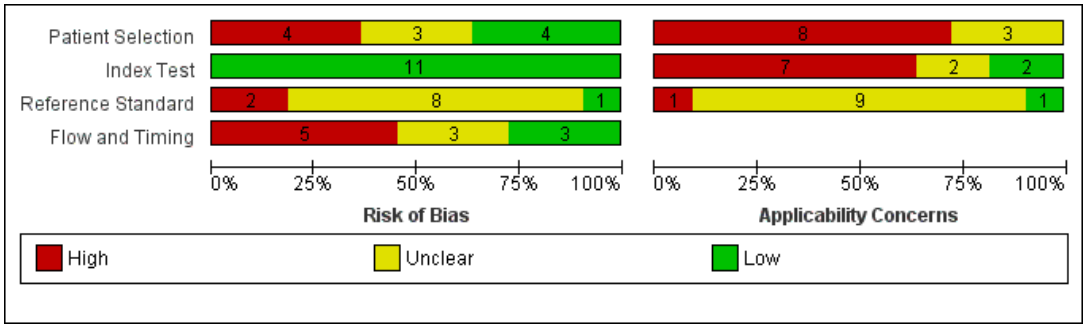
Figure 5


50196 records identified through database searching
267 additional records identified through other sources
16116 duplicate records removed
34347 records screened
32415 records excluded on title and abstract
848 full-text articles excluded from Diagnosis reviews**
(184 No 2x2 data)
(163 Not a primary study)
(141 Derivation study)
(126 Index test)
(119 Study population)
(72 Sample size)
(76 Reference standard)
(78 Target condition)
(59 Individual lesion characteristics)
(12 Test observer)
(13 Conference abstract)
(17 Duplicate or related publications)
(exclude but contact authors)
** some studies were coded with more than one reason for exclusion
1932 full-text articles assessed for eligibility:
Diagnosis = 1051 (203 includes)
Staging = 881
73 studies excluded*
20 ineligible target conditions
15 inadequate sample size
14 not primary test accuracy studies
9 derivation type studies or assessing individual RCM characteristics
7 ineligible populations
6 with insufficient data for 2x2 table
4 ineligible reference standards
* some studies were coded with more than one reason for exclusion
83 of 1051 tagged as potentially eligible for Reflectance confocal microscopy for detection of keratinocyte skin cancer review
10 included

Caption

PRISMA flow diagram.

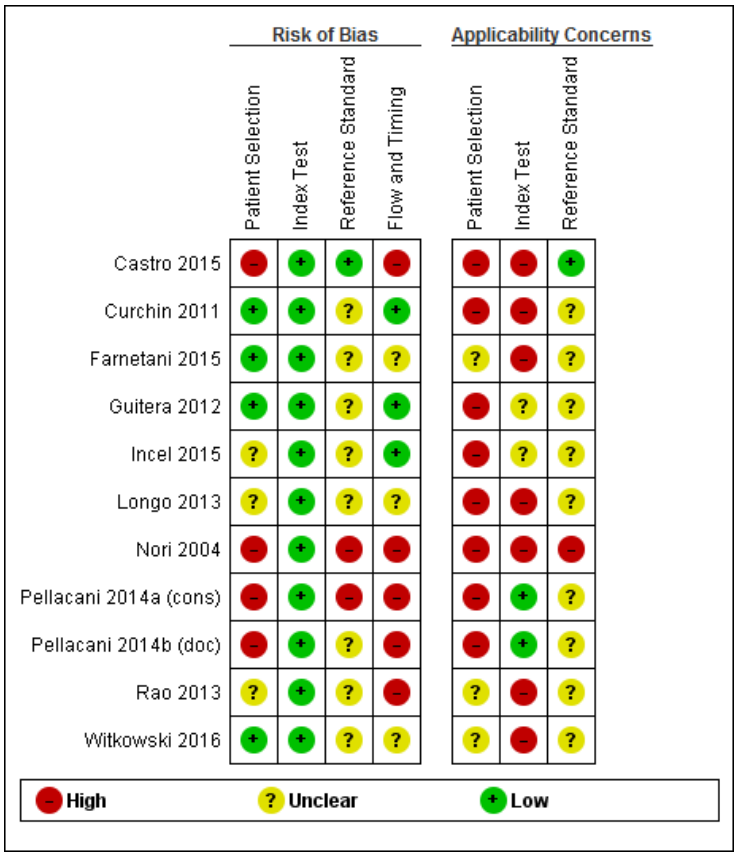
Figure 6



Caption

Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies

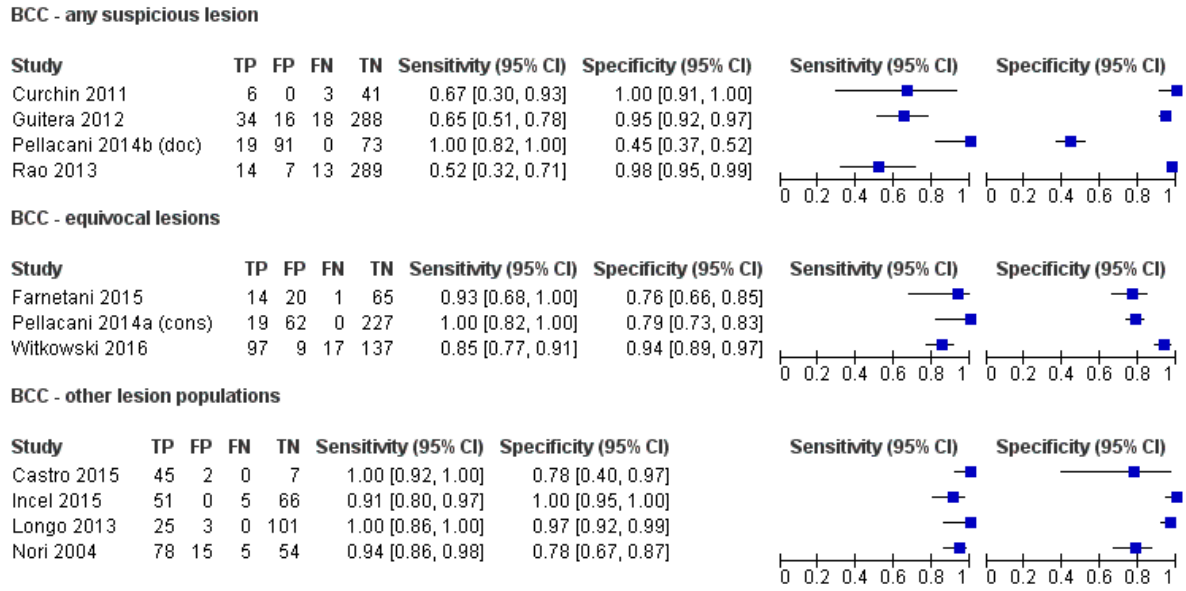
Figure 7



Caption

Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

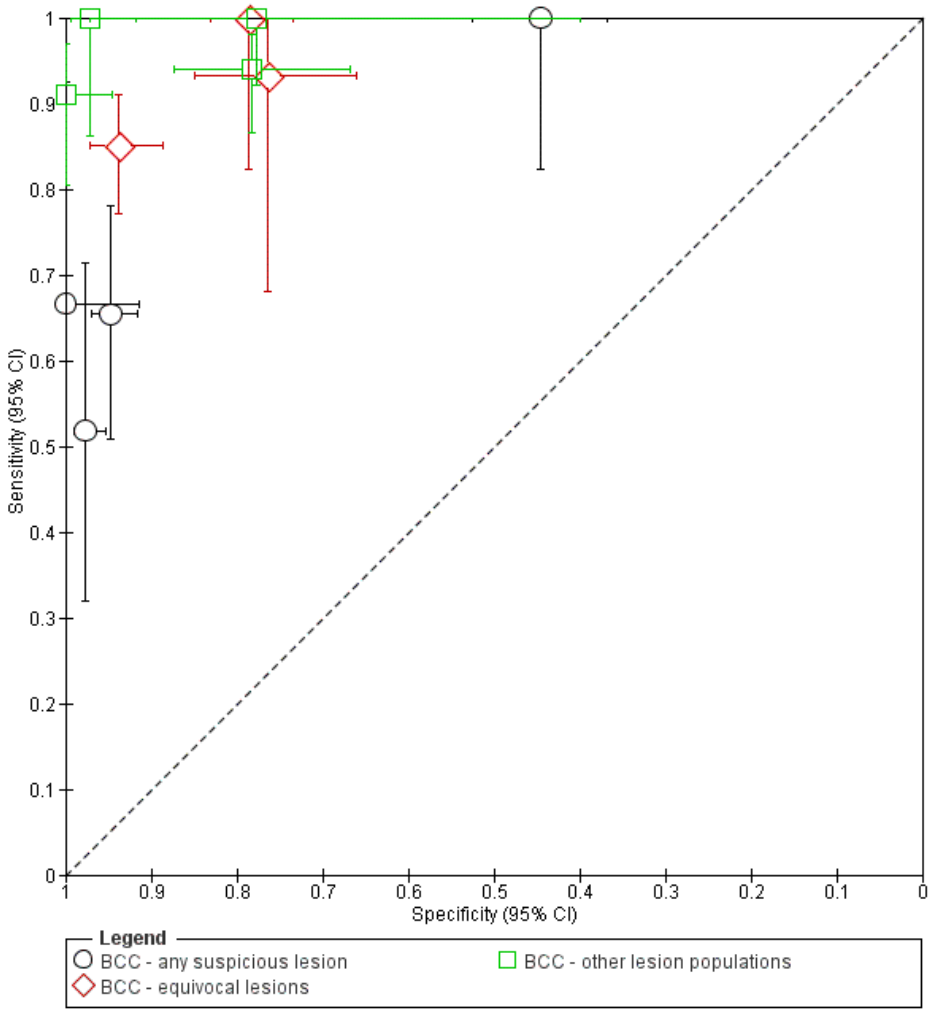
Figure 8 (Analysis 10)



Caption

Forest plot of tests: RCM for the detection of BCC in a) any suspicious lesion, b) equivocal lesions, c) other lesion populations

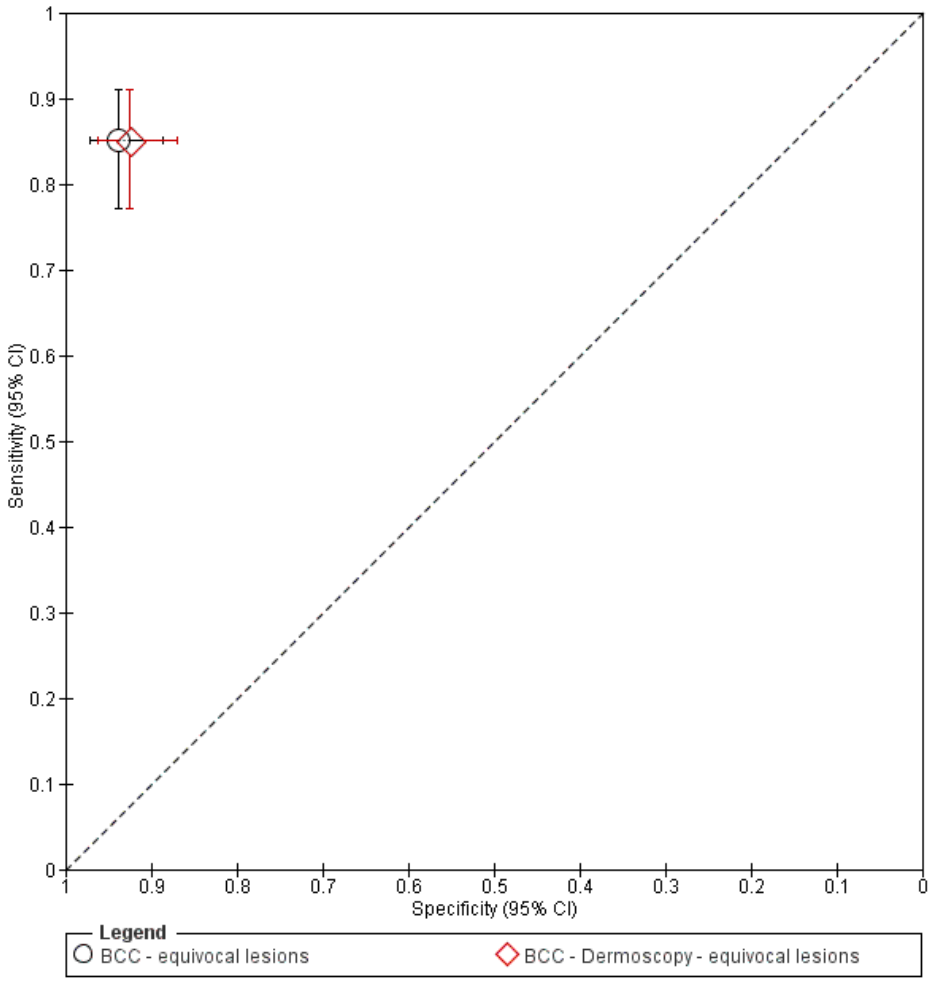
Figure 9 (Analysis 10)



Caption

ROC Plot of tests: RCM for the detection of BCC in a) any suspicious lesion, b) equivocal lesions, c) other lesion populations

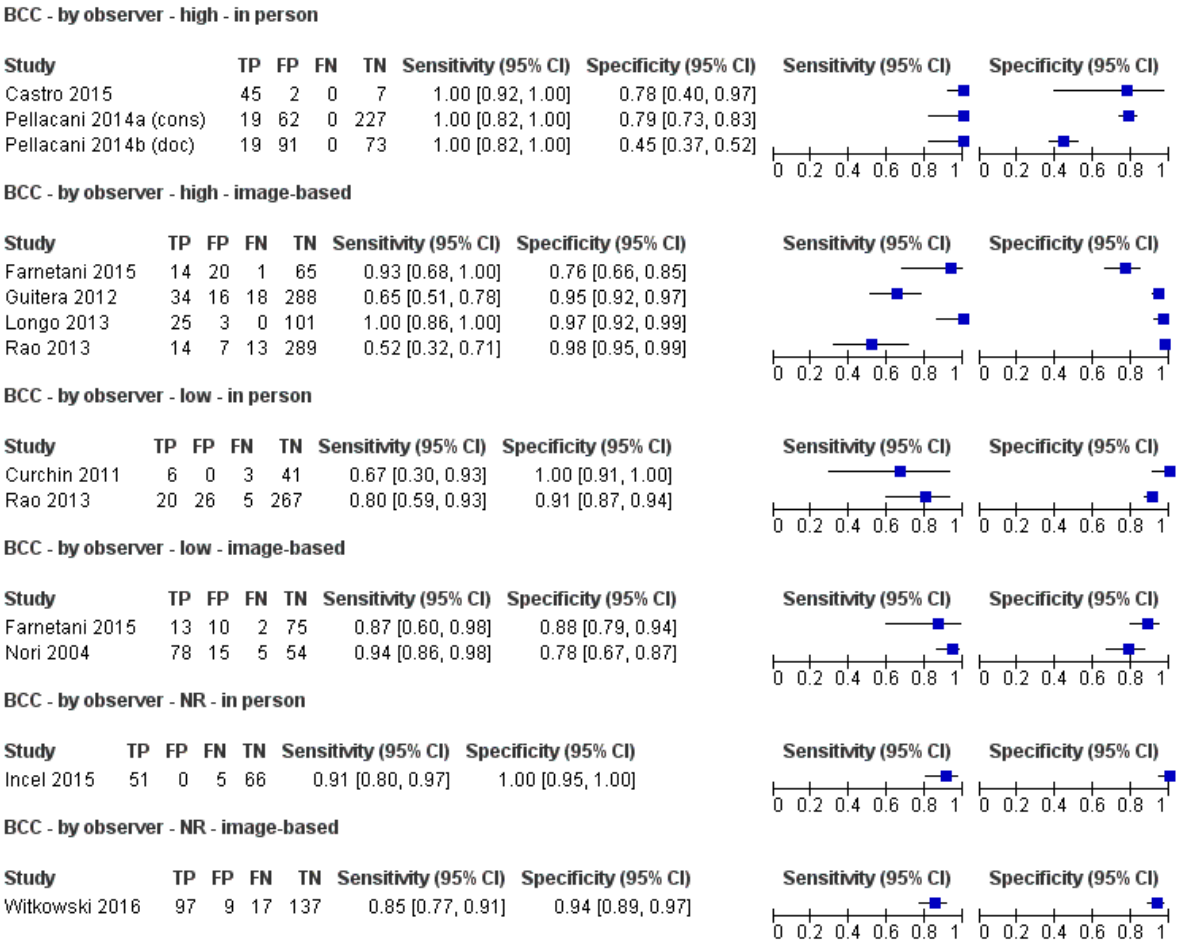
Figure 10 (Analysis 11)



Caption

ROC Plot of tests: RCM versus Dermoscopy in equivocal lesions (Witkowski 2016).

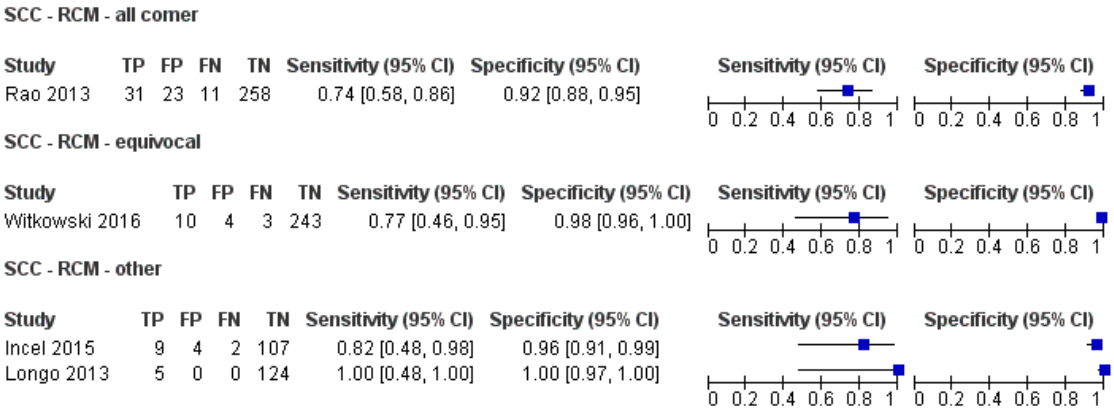
Figure 11 (Analysis 7)



Caption

Forest plot of accuracy of RCM to detect BCC by experience (separately for in person and image based studies)

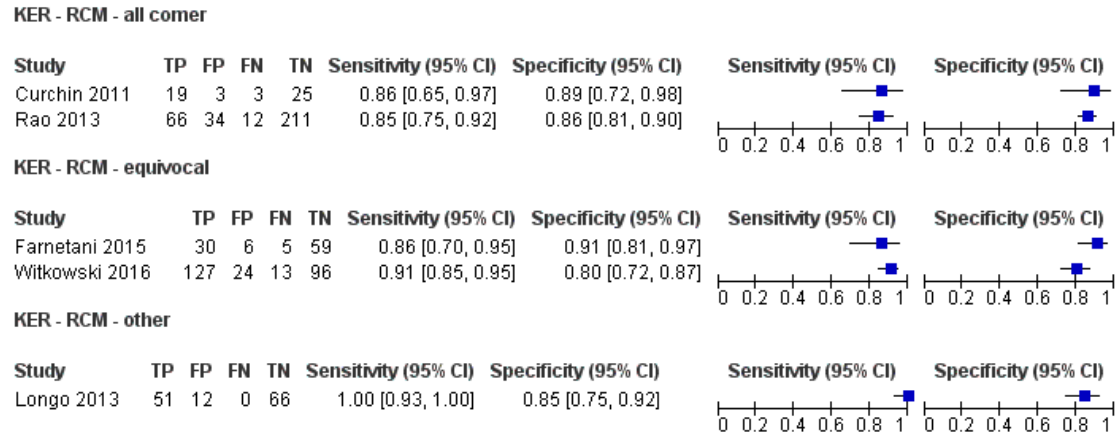
Figure 12 (Analysis 14)



Caption

Forest plot of tests: RCM for the detection of cSCC in a) any suspicious lesion, b) equivocal lesions, c) other lesion populations

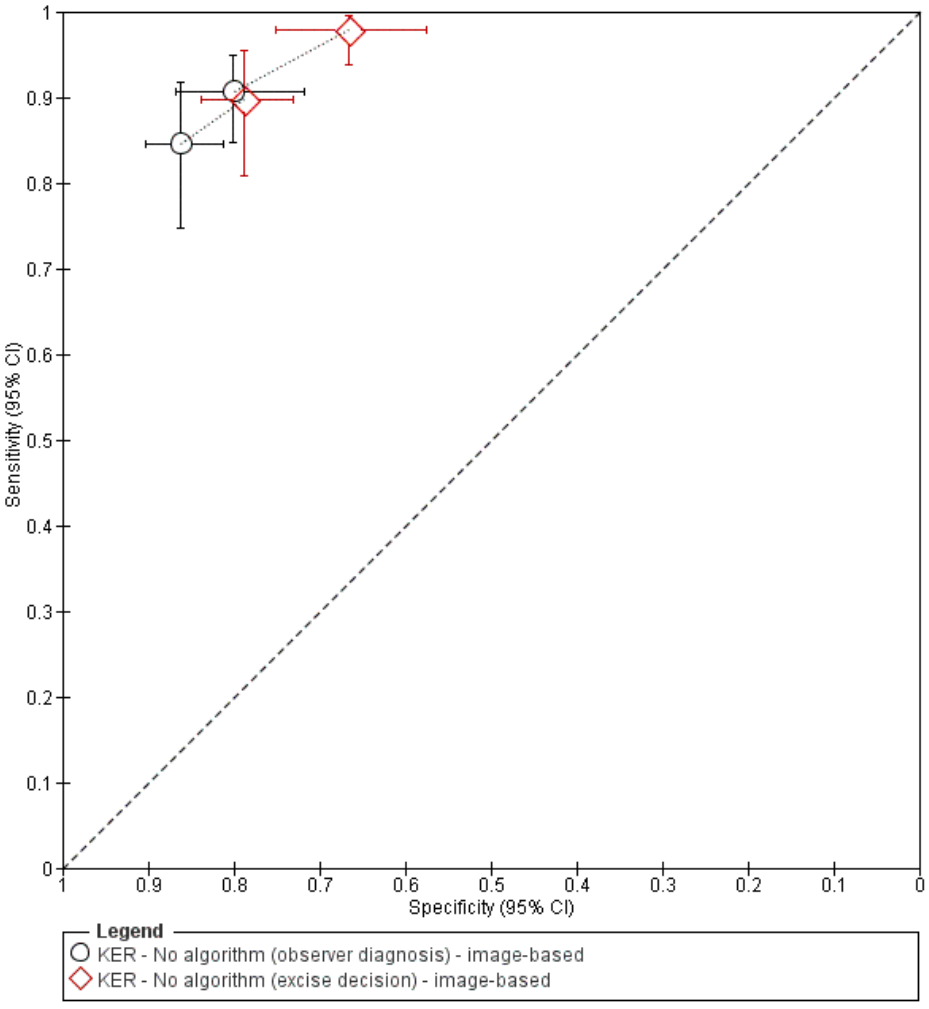
Figure 13 (Analysis 18)



Caption

Forest plot of tests: RCM for the detection of any skin cancer (KER) in a) any suspicious lesion, b) equivocal lesions, c) other lesion populations

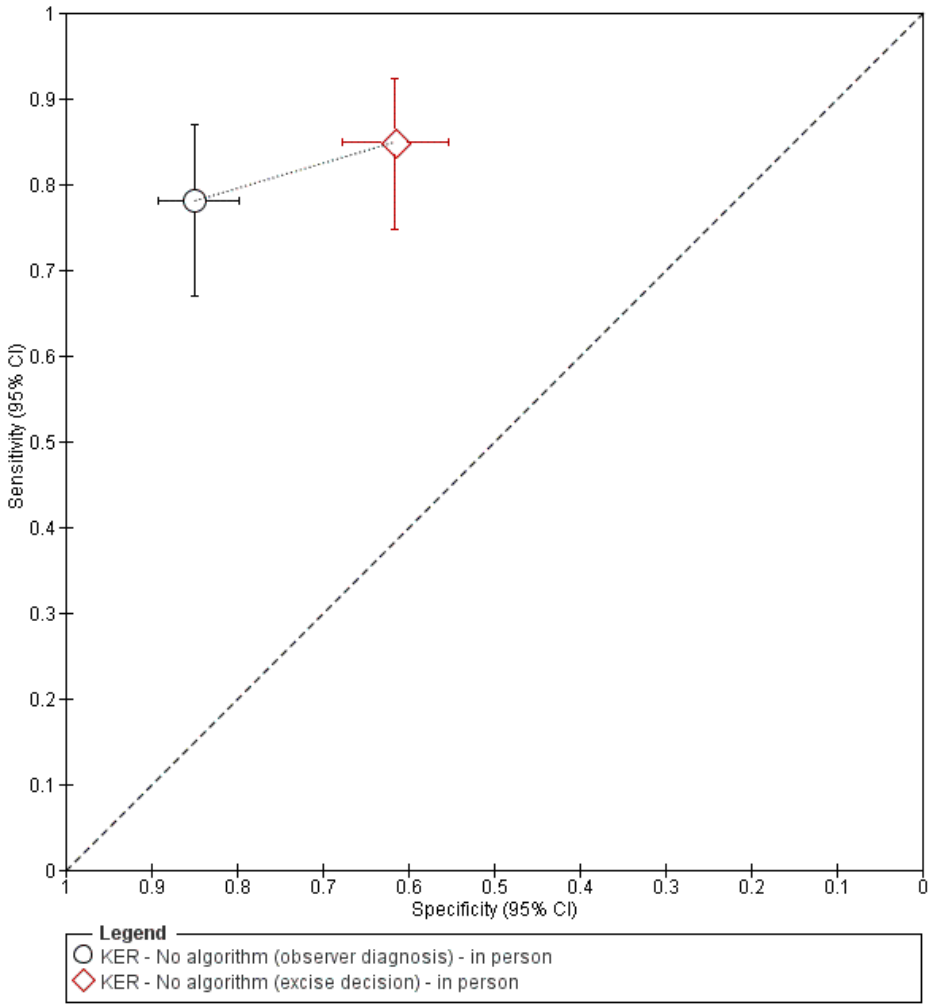
Figure 14 (Analysis 16)



Caption

ROC plot of tests: RCM for the detection of any skin cancer (KER) for a) correct diagnosis of each malignancy and b) decision to excise a lesion (image-based evaluations)

Figure 15 (Analysis 17)



Caption

ROC plot of tests: RCM for the detection of any skin cancer (KER) for a) correct diagnosis of each malignancy and b) decision to excise a lesion (in-person evaluations)

Sources of support

Internal sources

- No sources of support provided

External sources

- The National Institute for Health Research (NIHR), UK
  The NIHR, UK, is the largest single funder of the Cochrane Skin Group
- NIHR Systematic Review Programme, UK

Feedback

Appendices

1 Table of acronyms used in review text

Acronym	Definition
BCC	basal cell carcinoma
BPC	between-person comparative (study)
CAD	computer-assisted diagnosis
cSCC	cutaneous squamous cell carcinoma
DEJ	dermo epidermal junction
DTA	diagnostic test accuracy
ENT	ear, nose, and throat
FP	false positive
GP	general practitioner
KER	any keratinocyte skin cancer
LM	lentigo maligna
MEL	invasive melanoma or melanoma in situ
MM	malignant melanoma
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NML	non melanocytic lesion
OCT	optical coherence tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCM	Reflectance confocal microscopy
RCT	randomised controlled trial
RDEB	recessive dystrophic epidermolysis bullosa
ROC	receiver operating characteristic
SCC	squamous cell carcinoma
TN	true negative
UK	United Kingdom

2 Current content and structure of the Programme Grant

List of reviews	Estimated number of studies
Diagnosis of melanoma	
1. Visual inspection versus visual inspection plus dermoscopy	120
2. Teledermatology	12
3. Mobile phone applications	2
4. Computer-aided diagnosis: dermoscopy based and spectroscopy based techniques	37
5. Reflectance confocal microscopy	19
6. High frequency ultrasound	3
7. Overview: comparing the accuracy of tests for which sufficient evidence was identified either alone or in combination	–
Diagnosis of keratinocyte skin cancer (basal cell carcinoma and cutaneous squamous cell carcinoma)	
8. Visual inspection ± dermoscopy	22
9. Computer aided diagnosis: dermoscopy based and spectroscopy based techniques	3
10. Optical coherence tomography	6
11. Reflectance confocal microscopy	9
12. High frequency ultrasound	1
13. Exfoliative cytology	5
14. Overview: comparing the accuracy of tests for which sufficient evidence was identified either alone or in combination	–
Staging of melanoma	
15. Ultrasound	25 to 30
16. Computer tomography	5 to 10
17. Positron emission tomography or positron emission tomography-computer tomography	20 to 25
18. Magnetic resonance imaging	5
19. Sentinel lymph node biopsy ± high frequency ultrasound	70
20. Overview: comparing the accuracy of tests for which sufficient evidence was identified either alone or in combination	–
Staging of cutaneous squamous cell carcinoma	
21. Imaging tests review	10 to 15
22. Sentinel lymph node biopsy ± high frequency ultrasound	15 to 20

3 Proposed sources of heterogeneity

i. Population characteristics

- general versus higher risk populations
- patient population: Primary /secondary / specialist unit
- lesion suspicion: general suspicion/atypical/equivocal/NR
- lesion type: any pigmented; melanocytic
- inclusion of multiple lesions per participant
- ethnicity

ii. Index test characteristics

- type of test or algorithm used for test interpretation within each ‘group’ of tests
- the nature of and definition of criteria for test positivity
- diagnosis in person versus image-based diagnosis
- observer experience with the index test
- approaches to lesion preparation (e.g. the use of oil or antiseptic gel for dermoscopy)

iii. Reference standard characteristics

- reference standard used

- whether histology-reporting meets pathology-reporting guidelines
- use of excisional versus diagnostic biopsy
- whether two independent dermatopathologists reviewed histological diagnosis

iv. Study quality

- consecutive or random sample of participants recruited
- index test interpreted blinded to the reference standard result
- index test interpreted blinded to the result of any other index test
- presence of partial or differential verification bias (whereby only a sample of those subject to the index test are verified by the reference test or by the same reference test with selection dependent on the index test result)
- use of an adequate reference standard
- overall risk of bias

4 Final search strategies

Melanoma search strategies to August 2016

Database: Ovid MEDLINE(R) 1946 to August week 3 2016

Search strategy:

1 exp melanoma/

2 exp skin cancer/

3 exp basal cell carcinoma/

4 basalioma$1.ti,ab.

5 ((basal cell or skin) adj2 (cancer$1 or carcinoma$1 or mass or masses or tumour$1 or tumor$1 or neoplasm$1 or adenoma$1 or epithelioma$1 or lesion$1 or malignan$ or nodule$1)).ti,ab.

6 (pigmented adj2 (lesion$1 or mole$ or nevus or nevi or naevus or naevi or skin)).ti,ab.

7 (melanom$1 or nonmelanoma$1 or non-melanoma$1 or melanocyt$ or non-melanocyt$ or nonmelanocyt$ or keratinocyt$).ti,ab.

8 nmsc.ti,ab.

9 (squamous cell adj2 (cancer$1 or carcinoma$1 or mass or masses or tumor$1 or tumour$1 or neoplasm$1 or adenoma$1 or epithelioma$1 or epithelial or lesion$1 or malignan$ or nodule$1) adj2 (skin or epiderm$ or cutaneous)).ti,ab.

10 (BCC or CSCC or NMSC).ti,ab.

11 keratinocy$.ti,ab.

12 Keratinocytes/

13 or/1-12

14 dermoscop$.ti,ab.

15 dermatoscop$.ti,ab.

16 photomicrograph$.ti,ab.

17 exp epiluminescence microscopy/

18 (epiluminescence adj2 microscop$).ti,ab.

19 (confocal adj2 microscop$).ti,ab.

20 (incident light adj2 microscop$).ti,ab.

21 (surface adj2 microscop$).ti,ab.

22 (visual adj (inspect$ or examin$)).ti,ab.

23 ((clinical or physical) adj examin$).ti,ab.

24 3 point.ti,ab.

25 three point.ti,ab.

26 pattern analys$.ti,ab.

27 ABCD$.ti,ab.

28 menzies.ti,ab.

29 7 point.ti,ab.

30 seven point.ti,ab.

31 (digital adj2 (dermoscop$ or dermatoscop$)).ti,ab.

32 artificial intelligence.ti,ab.

33 AI.ti,ab.

34 computer assisted.ti,ab.

35 computer aided.ti,ab.

36 neural network$.ti,ab.

37 exp diagnosis, computer-assisted/

38 MoleMax.ti,ab.

39 image process$.ti,ab.

40 automatic classif$.ti,ab.

41 image analysis.ti,ab.

42 SIAscop$.ti,ab.

43 Aura.ti,ab.

44 (optical adj2 scan$).ti,ab.

45 MelaFind.ti,ab.

46 SIMSYS.ti,ab.

47 MoleMate.ti,ab.

48 SolarScan.ti,ab.

49 VivaScope.ti,ab.

50 (high adj3 ultraso$).ti,ab.

51 (canine adj2 detect$).ti,ab.

52 ((mobile or cell or cellular or smart) adj ((phone$1 adj2 app$1) or application$1)).ti,ab.

53 smartphone$.ti,ab.

54 (DermoScan or SkinVision or DermLink or SpotCheck).ti,ab.

55 Mole Detective.ti,ab.

56 Spot Check.ti,ab.

57 (mole$1 adj2 map$).ti,ab.

58 (total adj2 body).ti,ab.

59 exfoliative cytolog$.ti,ab.

60 digital analys$.ti,ab.

61 (image$1 adj3 software).ti,ab.

62 (teledermatolog$ or tele-dermatolog$ or telederm or tele-derm or teledermoscop$ or tele-dermoscop$ or teledermatoscop$ or tele-dermatoscop$).ti,ab.

63 (optical coherence adj (technolog$ or tomog$)).ti,ab.

64 (computer adj2 diagnos$).ti,ab.

65 exp sentinel lymph node biopsy/

66 (sentinel adj2 node).ti,ab.

67 nevisense.mp. or HFUS.ti,ab.

68 electrical impedance spectroscopy.ti,ab.

69 history taking.ti,ab.

70 patient history.ti,ab.

71 (naked eye adj (exam$ or assess$)).ti,ab.

72 (skin adj exam$).ti,ab.

73 physical examination/

74 ugly duckling.mp. or UD.ti,ab.

75 ((physician$ or clinical or physical) adj (exam$ or triage or recog$)).ti,ab.

76 ABCDE.mp. or VOC.ti,ab.

77 clinical accuracy.ti,ab.

78 Family Practice/ or Physicians, Family/ or clinical competence/

79 (confocal adj2 microscop$).ti,ab.

80 diagnostic algorithm$1.ti,ab.

81 checklist$.ti,ab.

82 virtual imag$1.ti,ab.

83 volatile organic compound$1.ti,ab.

84 dog$1.ti,ab.

85 gene expression analy$.ti,ab.

86 reflex transmission imag$.ti,ab.

87 thermal imaging.ti,ab.

88 elastography.ti,ab.

89 or/14-88

90 (CT or PET).ti,ab.

91 PET-CT.ti,ab.

92 (FDG or F18 or Fluorodeoxyglucose or radiopharmaceutical$).ti,ab.

93 exp Deoxyglucose/

94 deoxy-glucose.ti,ab.

95 deoxyglucose.ti,ab.

96 CATSCAN.ti,ab.

97 exp Tomography, Emission-Computed/

98 exp Tomography, X-ray computed/

99 positron emission tomograph$.ti,ab.

100 exp magnetic resonance imaging/

101 (MRI or fMRI or NMRI or scintigraph$).ti,ab.

102 exp echography/

103 Doppler echography.ti,ab.

104 sonograph$.ti,ab.

105 ultraso$.ti,ab.

106 doppler.ti,ab.

107 magnetic resonance imag$.ti,ab.

108 or/90-107

109 (stage$ or staging or metasta$ or recurrence or sensitivity or specificity or false negative$ or thickness$).ti,ab.

110 "Sensitivity and Specificity"/

111 exp cancer staging/

112 or/109-111

113 108 and 112

114 89 or 113

115 13 and 114

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 29 August 2016

Search strategy:

1 basalioma$1.ti,ab.

2 ((basal cell or skin) adj2 (cancer$1 or carcinoma$1 or mass or masses or tumour$1 or tumor$1 or neoplasm$1 or adenoma$1 or epithelioma$1 or lesion$1 or malignan$ or nodule$1)).ti,ab.

3 (pigmented adj2 (lesion$1 or mole$ or nevus or nevi or naevus or naevi or skin)).ti,ab.

4 (melanom$1 or nonmelanoma$1 or non-melanoma$1 or melanocyt$ or non-melanocyt$ or nonmelanocyt$ or keratinocyt$).ti,ab.

5 nmsc.ti,ab.

6 (squamous cell adj2 (cancer$1 or carcinoma$1 or mass or masses or tumor$1 or tumour$1 or neoplasm$1 or adenoma$1 or epithelioma$1 or epithelial or lesion$1 or malignan$ or nodule$1) adj2 (skin or epiderm$ or cutaneous)).ti,ab.

7 (BCC or CSCC or NMSC).ti,ab.
8 keratinocy$.ti,ab.
9 or/1-8
10 dermoscop$.ti,ab.
11 dermatoscop$.ti,ab.
12 photomicrograph$.ti,ab.
13 (epiluminescence adj2 microscop$).ti,ab.
14 (confocal adj2 microscop$).ti,ab.
15 (incident light adj2 microscop$).ti,ab.
16 (surface adj2 microscop$).ti,ab.
17 (visual adj (inspect$ or examin$)).ti,ab.
18 ((clinical or physical) adj examin$).ti,ab.
19 3 point.ti,ab.
20 three point.ti,ab.
21 pattern analys$.ti,ab.
22 ABCD$.ti,ab.
23 menzies.ti,ab.
24 7 point.ti,ab.
25 seven point.ti,ab.
26 (digital adj2 (dermoscop$ or dermatoscop$)).ti,ab.
27 artificial intelligence.ti,ab.
28 AI.ti,ab.
29 computer assisted.ti,ab.
30 computer aided.ti,ab.
31 neural network$.ti,ab.
32 MoleMax.ti,ab.
33 image process$.ti,ab.
34 automatic classif$.ti,ab.
35 image analysis.ti,ab.
36 SIAscop$.ti,ab.
37 Aura.ti,ab.
38 (optical adj2 scan$).ti,ab.
39 MelaFind.ti,ab.
40 SIMSYS.ti,ab.
41 MoleMate.ti,ab.
42 SolarScan.ti,ab.
43 VivaScope.ti,ab.
44 (high adj3 ultraso$).ti,ab.
45 (canine adj2 detect$).ti,ab.
46 ((mobile or cell or cellular or smart) adj ((phone$1 adj2 app$1) or application$1)).ti,ab.
47 smartphone$.ti,ab.
48 (DermoScan or SkinVision or DermLink or SpotCheck).ti,ab.
49 Mole Detective.ti,ab.
50 Spot Check.ti,ab.
51 (mole$1 adj2 map$).ti,ab.
52 (total adj2 body).ti,ab.

53 exfoliative cytolog$.ti,ab.

54 digital analys$.ti,ab.

55 (image$1 adj3 software).ti,ab.

56 (teledermatolog$ or tele-dermatolog$ or telederm or tele-derm or teledermoscop$ or tele-dermoscop$ or teledermatoscop$ or tele-dermatoscop$).ti,ab.

57 (optical coherence adj (technolog$ or tomog$)).ti,ab.

58 (computer adj2 diagnos$).ti,ab.

59 (sentinel adj2 node).ti,ab.

60 nevisense.mp. or HFUS.ti,ab.

61 electrical impedance spectroscopy.ti,ab.

62 history taking.ti,ab.

63 patient history.ti,ab.

64 (naked eye adj (exam$ or assess$)).ti,ab.

65 (skin adj exam$).ti,ab.

66 ugly duckling.mp. or UD.ti,ab.

67 ((physician$ or clinical or physical) adj (exam$ or triage or recog$)).ti,ab.

68 ABCDE.mp. or VOC.ti,ab.

69 clinical accuracy.ti,ab.

70 (Family adj (Practice or Physicians)).ti,ab.

71 (confocal adj2 microscop$).ti,ab.

72 clinical competence.ti,ab.

73 diagnostic algorithm$1.ti,ab.

74 checklist$.ti,ab.

75 virtual imag$1.ti,ab.

76 volatile organic compound$1.ti,ab.

77 dog$1.ti,ab.

78 gene expression analy$.ti,ab.

79 reflex transmission imag$.ti,ab.

80 thermal imaging.ti,ab.

81 elastography.ti,ab.

82 or/10-81

83 (CT or PET).ti,ab.

84 PET-CT.ti,ab.

85 (FDG or F18 or Fluorodeoxyglucose or radiopharmaceutical$).ti,ab.

86 deoxy-glucose.ti,ab.

87 deoxyglucose.ti,ab.

88 CATSCAN.ti,ab.

89 positron emission tomograph$.ti,ab.

90 (MRI or fMRI or NMRI or scintigraph$).ti,ab.

91 Doppler echography.ti,ab.

92 sonograph$.ti,ab.

93 ultraso$.ti,ab.

94 doppler.ti,ab.

95 magnetic resonance imag$.ti,ab.

96 or/83-95

97 (stage$ or staging or metasta$ or recurrence or sensitivity or specificity or false negative$ or thickness$).ti,ab.

98 96 and 97

99 82 or 98

100 9 and 99

Database: Embase 1974 to 29 August 2016

Search strategy:

1 *melanoma/

2 *skin cancer/

3 *basal cell carcinoma/

4 basalioma$.ti,ab.

5 ((basal cell or skin) adj2 (cancer$1 or carcinoma$1 or mass or masses or tumour$1 or tumor$1 or neoplasm$ or adenoma$ or epithelioma$ or lesion$ or malignan$ or nodule$)).ti,ab.

6 (pigmented adj2 (lesion$1 or mole$ or nevus or nevi or naevus or naevi or skin)).ti,ab.

7 (melanom$1 or nonmelanoma$1 or non-melanoma$1 or melanocyt$ or non-melanocyt$ or nonmelanocyt$ or keratinocyt$).ti,ab.

8 nmsc.ti,ab.

9 (squamous cell adj2 (cancer$1 or carcinoma$1 or mass or tumor$1 or tumour$1 or neoplasm$1 or adenoma$1 or epithelioma$1 or epithelial or lesion$1 or malignan$ or nodule$1) adj2 (skin or epiderm$ or cutaneous)).ti,ab.

10 (BCC or cscc).mp. or NMSC.ti,ab.

11 keratinocyte.ti,ab.

12 keratinocy$.ti,ab.

13 or/1-12

14 dermoscop$.ti,ab.

15 dermatoscop$.ti,ab.

16 photomicrograph$.ti,ab.

17 *epiluminescence microscopy/

18 (epiluminescence adj2 microscop$).ti,ab.

19 (confocal adj2 microscop$).ti,ab.

20 (incident light adj2 microscop$).ti,ab.

21 (surface adj2 microscop$).ti,ab.

22 (visual adj (inspect$ or examin$)).ti,ab.

23 ((clinical or physical) adj examin$).ti,ab.

24 3 point.ti,ab.

25 three point.ti,ab.

26 pattern analys$.ti,ab.

27 ABCD$.ti,ab.

28 menzies.ti,ab.

29 7 point.ti,ab.

30 seven point.ti,ab.

31 (digital adj2 (dermoscop$ or dermatoscop$)).ti,ab.

32 artificial intelligence.ti,ab.

33 AI.ti,ab.

34 computer assisted.ti,ab.

35 computer aided.ti,ab.

36 neural network$.ti,ab.

37 MoleMax.ti,ab.

38 exp diagnosis, computer-assisted/

39 image process$.ti,ab.

40 automatic classif$.ti,ab.

41 image analysis.ti,ab.

42 SIAscop$.ti,ab.

43 (optical adj2 scan$).ti,ab.

44 Aura.ti,ab.

45 MelaFind.ti,ab.

46 SIMSYS.ti,ab.

47 MoleMate.ti,ab.

48 SolarScan.ti,ab.

49 VivaScope.ti,ab.

50 confocal microscop$.ti,ab.

51 (high adj3 ultraso$).ti,ab.

52 (canine adj2 detect$).ti,ab.

53 ((mobile or cell$ or cellular or smart) adj ((phone$1 adj2 app$1) or application$1)).ti,ab.

54 smartphone$.ti,ab.

55 (DermoScan or SkinVision or DermLink or SpotCheck).ti,ab.

56 Spot Check.ti,ab.

57 Mole Detective.ti,ab.

58 (mole$1 adj2 map$).ti,ab.

59 (total adj2 body).ti,ab.

60 exfoliative cytolog$.ti,ab.

61 digital analys$.ti,ab.

62 (image$1 adj3 software).ti,ab.

63 (optical coherence adj (technolog$ or tomog$)).ti,ab.

64 (teledermatolog$ or tele-dermatolog$ or telederm or tele-derm or teledermoscop$ or tele-dermoscop$ or teledermatoscop$).mp. or tele-dermatoscop$.ti,ab.

65 (computer adj2 diagnos$).ti,ab.

66 *sentinel lymph node biopsy/

67 (sentinel adj2 node).ti,ab.

68 nevisense.ti,ab.

69 HFUS.ti,ab.

70 electrical impedance spectroscopy.ti,ab.

71 history taking.ti,ab.

72 patient history.ti,ab.

73 (naked eye adj (exam$ or assess$)).ti,ab.

74 (skin adj exam$).ti,ab.

75 *physical examination/

76 ugly duckling.ti,ab.

77 UD sign$.ti,ab.

78 ((physician$ or clinical or physical) adj (exam$ or recog$ or triage)).ti,ab.

79 ABCDE.ti,ab.

80 clinical accuracy.ti,ab.

81 *general practice/

82 (confocal adj2 microscop$).ti,ab.

83 clinical competence/

84 diagnostic algorithm$.ti,ab.

85 checklist$1.ti,ab.

86 virtual image$1.ti,ab.

87 volatile organic compound$1.ti,ab.

88 VOC.ti,ab.

89 dog$1.ti,ab.

90 gene expression analys$.ti,ab.

91 reflex transmission imaging.ti,ab.

92 thermal imaging.ti,ab.

93 elastography.ti,ab.

94 dog$1.ti,ab.

95 gene expression analys$.ti,ab.

96 reflex transmission imaging.ti,ab.

97 thermal imaging.ti,ab.

98 elastography.ti,ab.

99 or/14-93

100 PET-CT.ti,ab.

101 (CT or PET).ti,ab.

102 (FDG or F18 or Fluorodeoxyglucose or radiopharmaceutical$).ti,ab.

103 exp Deoxyglucose/

104 CATSCAN.ti,ab.

105 deoxyglucose.ti,ab.

106 deoxy-glucose.ti,ab.

107 *positron emission tomography/

108 *computer assisted tomography/

109 positron emission tomograph$.ti,ab.

110 *nuclear magnetic resonance imaging/

111 (MRI or fMRI or NMRI or scintigraph$).ti,ab.

112 *echography/

113 Doppler.ti,ab.

114 sonograph$.ti,ab.

115 ultraso$.ti,ab.

116 magnetic resonance imag$.ti,ab.

117 or/100-116

118 (stage$ or staging or metasta$ or recurrence or sensitivity or specificity or false negative$ or thickness$).ti,ab.

119 "Sensitivity and Specificity"/

120 *cancer staging/

121 or/118-120

122 117 and 121

123 99 or 122

124 13 and 123

Database: Cochrane Library (Wiley) 2016 searched 30 August 2016 CDSR Issue 8 of 12 2016 CENTRAL Issue 7 of 12 2016 HTA Issue 3 of 4 July 2016 DARE Issue 3 of 4 2015

Search strategy:

#1 melanoma* or nonmelanoma* or non-melanoma* or melanocyt* or non-melanocyt* or nonmelanocyt* or keratinocyte*

#2 MeSH descriptor: [Melanoma] explode all trees

#3 "skin cancer*"

#4 MeSH descriptor: [Skin Neoplasms] explode all trees

#5 skin near/2 (cancer* or carcinoma* or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*)

#6 nmsc

#7 "squamous cell" near/2 (cancer* or carcinoma* or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*) near/2 (skin or epiderm* or cutaneous)

#8 "basal cell" near/2 (cancer* or carcinoma* or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*)

#9 pigmented near/2 (lesion* or nevus or mole* or naevi or naevus or nevi or skin)

#10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9

#11 dermoscop*

#12 dermatoscop*

#13 Photomicrograph*

#14 MeSH descriptor: [Dermoscopy] explode all trees

#15 confocal near/2 microscop*

#16 epiluminescence near/2 microscop*

#17 incident next light near/2 microscop*

#18 surface near/2 microscop*

#19 "visual inspect*"

#20 "visual exam*"

#21 (clinical or physical) next (exam*)

#22 "3 point"

#23 "three point"

#24 "pattern analys*"

#25 ABDC

#26 menzies

#27 "7 point"

#28 "seven point"

#29 digital near/2 (dermoscop* or dermatoscop*)

#30 "artificial intelligence"

#31 "AI"

#32 "computer assisted"

#33 "computer aided"

#34 AI

#35 "neural network*"

#36 MoleMax

#37 "computer diagnosis"

#38 "image process*"

#39 "automatic classif*"

#40 SIAscope

#41 "image analysis"

#42 "optical near/2 scan*"

#43 Aura

#44 MelaFind

#45 SIMSYS

#46 MoleMate

#47 SolarScan

#48 Vivascope

#49 "confocal microscopy"

#50 high near/3 ultraso*

#51 canine near/2 detect*

#52 Mole* near/2 map*

#53 total near/2 body

#54 mobile* or smart near/2 phone*

#55 cell next phone*

#56 smartphone*

#57 "mitotic index"

#58 DermoScan or SkinVision or DermLink or SpotCheck

#59 "Mole Detective"

#60 "Spot Check"

#61 mole* near/2 map*

#62 total near/2 body

#63 "exfoliative cytolog*"

#64 "digital analys*"

#65 image near/3 software

#66 teledermatolog* or tele-dermatolog* or telederm or tele-derm or teledermoscop* or tele-dermoscop* or teledermatoscop* or tele-dermatolog*

#67 "optical coherence" next (technolog* or tomog*)

#68 computer near/2 diagnos*

#69 sentinel near/2 node*

#70 #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69

#71 ultraso*

#72 sonograph*

#73 MeSH descriptor: [Ultrasonography] explode all trees

#74 Doppler

#75 CT or PET or PET-CT

#76 "CAT SCAN" or "CATSCAN"

#77 MeSH descriptor: [Positron-Emission Tomography] explode all trees

#78 MeSH descriptor: [Tomography, X-Ray Computed] explode all trees

#79 MRI

#80 MeSH descriptor: [Magnetic Resonance Imaging] explode all trees

#81 MRI or fMRI or NMRI or scintigraph*

#82 "magnetic resonance imag*"

#83 MeSH descriptor: [Deoxyglucose] explode all trees

#84 deoxyglucose or deoxy-glucose

#85 "positron emission tomograph*"

#86 #71 or #72 or #73 or #74 or #75 or #76 or #77 or #78 or #79 or #80 or #81 or #82 or #83 or #84 or #85

#87 stage* or staging or metasta* or recurrence or sensitivity or specificity or "false negative*" or thickness*

#88 MeSH descriptor: [Neoplasm Staging] explode all trees

#89 #87 or #88

#90 #89 and #86

#91 #70 or #90

#92 #10 and #91

#93 BCC or CSCC or NMCS

#94 keratinocy*

#95 #93 or #94

#96 #10 or #95

#97 nevisense

#98 HFUS

#99 "electrical impedance spectroscopy"

#100 "history taking"

#101 "patient history"

#102 naked next eye near/1 (exam* or assess*)

#103 skin next exam*

#104 "ugly duckling" or (UD sign*)

#105 MeSH descriptor: [Physical Examination] explode all trees

#106 (physician* or clinical or physical) near/1 (exam* or recog* or triage*)

#107 ABCDE

#108 "clinical accuracy"

#109 MeSH descriptor: [General Practice] explode all trees

#110 confocal near microscop*

#111 "diagnostic algorithm*"

#112 MeSH descriptor: [Clinical Competence] explode all trees

#113 checklist*

#114 "virtual image*"

#115 "volatile organic compound*"

#116 dog or dogs

#117 VOC

#118 "gene expression analys*"

#119 "reflex transmission imaging"

#120 "thermal imaging"

#121 elastography

#122 #97 or #98 or #99 or #100 or #101 or #102 or #103 or #104 or #105 or #106 or #107 or #108 or #109 or #110 or #111 or #112 or #113 or #114 or #115 or #116 or #117 or #118 or #119 or #120 or #121

#123 #70 or #122

#124 #96 and #123

#125 #96 and #90

#126 #125 or #124

#127 #10 and #126

Database: CINAHL Plus (EBSCO) 1937 to 30 August 2016

Search strategy:

S1 (MH "Melanoma") OR (MH "Nevi and Melanomas+")

S2 (MH "Skin Neoplasms+")

S3 (MH "Carcinoma, Basal Cell+")

S4 basalioma*

S5 (basal cell) N2 (cancer* or carcinoma* or mass or masses or tumor* or tumour* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*)

S6 (pigmented) N2 (lesion* or mole* or nevus or nevi or naevus or naevi or skin)

S7 melanom* or nonmelanoma* or non-melanoma* or melanocyt* or non-melanocyt* or nonmelanocyt*

S8 nmsc

S9 TX BCC or cscc or NMSC

S10 (MH "Keratinocytes")

S11 keratinocyt*

S12 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11

S13 dermoscop* or dermatoscop* or photomicrograph* or (3 point) or (three point) or ABCD* or menzies or (7 point) or (seven point) or AI or Molemax or SIASCOP* or Aura or MelaFind or SIMSYS or MoleMate or SolarScan or smartphone* or DermoScan or SkinVision or DermLink or SpotCheck

S14 (epiluminescence or confocal or incident or surface) N2 (microscop*)

S15 visual N1 (inspect* or examin*)

S16 (clinical or physical) N1 (examin*)

S17 pattern analys*

S18 (digital) N2 (dermoscop* or dermatoscop*)

S19 (artificial intelligence)

S20 (computer) N2 (assisted or aided)

S21 (neural network*)

S22 (MH "Diagnosis, Computer Assisted+")

S23 (image process*)

S24 (automatic classif*)

S25 (image analysis)

S26 SIAScop*

S27 (optical) N2 (scan*)

S28 (high) N3 (ultraso*)

S29 elastography

S30 (mobile or cell or cellular or smart) N2 (phone*) N2 (app or application*)

S31 (mole*) N2 (map*)

S32 total N2 body

S33 exfoliative cytolog*

S34 digital analys*

S35 image N3 software

S36 teledermatolog* or tele-dermatolog* or telederm or tele-derm or teledermoscop* or tele-dermoscop* or teledermatoscop* or tele-dermatoscop* teledermatolog* or tele-dermatolog* or telederm or tele-derm or teledermoscop*

S37 (optical coherence) N1 (technolog* or tomog*)

S38 computer N2 diagnos*

S39 sentinel N2 node

S40 (MH "Sentinel Lymph Node Biopsy")

S41 nevisense or HFUS or checklist* or VOC or dog*

S42 electrical impedance spectroscopy

S43 history taking

S44 "Patient history"

S45 naked eye

S46 skin exam*

S47 physical exam*

S48 ugly duckling

S49 UD sign*

S50 (physician* or clinical or physical) N1 (exam*)

S51 clinical accuracy

S52 general practice

S53 (physician* or clinical or physical) N1 (recog* or triage)

S54 confocal microscop*

S55 clinical competence

S56 diagnostic algorithm*

S57 checklist*

S58 virtual image*

S59 volatile organic compound*

S60 gene expression analys*

S61 reflex transmission imag*

S62 thermal imaging

S63 S13 or S14 or S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62

S64 CT or PET

S65 PET-CT

S66 FDG or F18 or Fluorodeoxyglucose or radiopharmaceutical*

S67 (MH "Deoxyglucose+")

S68 deoxy-glucose or deoxyglucose

S69 CATSCAN

S70 CAT-SCAN

S71 (MH "Deoxyglucose+")

S72 (MH "Tomography, Emission-Computed+")

S73 (MH "Tomography, X-Ray Computed")

S74 positron emission tomograph*

S75 (MH "Magnetic Resonance Imaging+")

S76 MRI or fMRI or NMRI or scintigraph*

S77 echography

S78 doppler

S79 sonograph*

S80 ultraso*

S81 magnetic resonance imag*

S82 S64 OR S65 OR S66 OR S67 OR S68 OR S69 OR S70 OR S71 OR S72 OR S73 OR S74 OR S75 OR S76 OR S77 OR S78 OR S79 OR S80 OR S81

S83 stage* or staging or metasta* or recurrence or sensitivity or specificity or (false negative*) or thickness

S84 (MH "Neoplasm Staging")

S85 S83 OR S84

S86 S82 AND S85

S87 S63 OR S86

S88 S12 AND S87

Database: Science Citation Index SCI Expanded (Web of Science) 1900 to 30 August 2016

Conference Proceedings Citation Index (Web of Science) 1900 to 1 September 2016

Search strategy:

#1 (melanom* or nonmelanom* or non-melanoma* or melanocyt* or non-melanocyt* or nonmelanocyt* or keratinocyt*)

#2 (basalioma*)

#3 ((skin) near/2 (cancer* or carcinoma or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*))

#4 ((basal) near/2 (cancer* or carcinoma* or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*))

#5 ((pigmented) near/2 (lesion* or mole* or nevus or nevi or naevus or naevi or skin))

#6 (nmsc or BCC or NMSC or keratinocy*)

#7 ((squamous cell (cancer* or carcinoma* or mass or masses or tumour* or tumor* or neoplasm* or adenoma* or epithelioma* or lesion* or malignan* or nodule*))

#8 (skin or epiderm* or cutaneous)

#9 #8 AND #7

#10 #9 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1

#11 ((dermoscop* or dermatoscop* or photomicrograph* or epiluminescence or confocal or "incident light" or "surface microscop*" or "visual inspect*" or "physical exam*" or 3 point or three point or pattern analy* or ABCDE or menzies or 7 point or seven point or dermoscop* or dermatoscop* or AI or artificial or computer aided or computer assisted or neural network* or Molemax or image process* or automatic classif* or image analysis or siascope or optical scan* or Aura or melafind or simsys or molemate or solarscan or vivascope or confocal microscop* or high ultraso* or canine detect* or cellphone* or mobile* or phone* or smartphone or dermoscan or skinvision or dermlink or spotcheck or spot check or mole detective or mole map* or total body or exfoliative psychology or digital or image software or optical coherence or teledermatology or telederm* or teledermoscop* or teledermatoscop* or computer diagnos* or sentinel))

#12 ((nevisense or HFUS or impedance spectroscopy or history taking or patient history or naked eye or skin exam* or physical exam* or ugly duckling or UD sign* or physician* exam* or physical exam* or ABCDE or clinical accuracy or general practice or confocal microscop* or clinical competence or diagnostic algorithm* or checklist* or virtual image* or volatile organic or VOC or dog* or gene expression or reflex transmission or thermal imag* or elastography))

#13 #11 or #12

#14 ((PET or CT or FDG or deoxyglucose or deoxy-glucose or fluorodeoxy* or radiopharma* or CATSCAN or positron emission or computer assisted or nuclear magnetic or MRI or FMRI or NMRI or scintigraph* or echograph* or Doppler or sonograph* or ultraso* or magnetic reson*))

#15 ((stage* or staging or metast* or recurrence or sensitivity or specificity or false negative* or thickness*))

#16 #14 AND #15

#17 #16 OR #13

#18 #10 AND #17

Refined by: DOCUMENT TYPES: (MEETING ABSTRACT OR PROCEEDINGS PAPER)

5 Full text inclusion criteria

Criterion	Inclusion	Exclusion

Study design	**For diagnostic and staging reviews**
	• Any study for which a 2×2 contingency table can be extracted, e.g.
	  ◦ diagnostic case control studies
	  ◦ 'cross-sectional' test accuracy study with retrospective or prospective data collection
	  ◦ studies where estimation of test accuracy was not the primary objective but test results for both index and reference standard were available
	  ◦ RCTs of tests or testing strategies where participants were randomised between index tests and all undergo a reference standard (i.e. accuracy RCTs)
		• < 5 melanoma cases (diagnosis reviews)
		• < 10 participants (staging reviews)
		• Studies developing new criteria for diagnosis unless a separate 'test set' of images were used to evaluate the criteria (mainly digital dermoscopy)
		• Studies using 'normal' skin as controls
		• Letters, editorials, comment papers, narrative reviews
		• Insufficient data to construct a 2×2 table

Target condition	• Melanoma
	• Keratinocyte skin cancer (or non-melanoma skin cancer)
	  ◦ BCC or epithelioma
	  ◦ cSCC
		• Studies exclusively conducted in children
		• Studies of non-cutaneous melanoma or SCC

Population	**For diagnostic reviews** • Adults with a skin lesion suspicious for melanoma, BCC, or cSCC (other terms include pigmented skin lesion/nevi, melanocytic, keratinocyte, etc.) • Adults at high risk of developing melanoma skin cancer, BCC, or cSCC **For staging reviews** • Adults with a diagnosis of melanoma or cSCC undergoing tests for staging of lymph nodes or distant metastases or both	• People suspected of other forms of skin cancer • Studies conducted exclusively in children

Criterion	Inclusion	Exclusion
Index tests	For diagnosis • Visual inspection/clinical examination • Dermoscopy/dermatoscopy • Teledermoscpoy • Smartphone/mobile phone applications • Digital dermoscopy/artificial intelligence • Confocal microscopy • Ocular coherence tomography • Exfoliative cytology • High frequency ultrasound • Canine odour detection • DNA expression analysis/gene chip analysis • Other For staging • CT • PET • PET-CT • MRI • Ultrasound +/fine needle aspiration cytology FNAC • SLNB +/high frequency ultrasound • Other Any test combination and in any order Any test positivity threshold Any variation in testing procedure (e.g. radioisotope used)	• Sentinel lymph biopsy for therapeutic rather than staging purposes • Tests to determine melanoma thickness • Tests to determine surgical margins/lesion borders • Tests to improve histopathology diagnose • LND
Reference standard	For diagnostic studies • Histopathology of the excised lesion • Clinical follow-up of non-excised/benign appearing lesions with later histopathology if suspicious • Expert diagnosis (studies should not be included if expert diagnosis is the sole reference standard) For studies of imaging tests for staging • Histopathology (via LND or SLMB) • Clinical/radiological follow-up • A combination of the above For studies of SLNB accuracy for staging • LND of both SLN+ and SLn participants to identify all diseased nodes • LND of SLN+ participants and follow-up of SLN participants to identify a subsequent nodal recurrence in a *previously investigated* nodal basin	For diagnostic studies • Exclude if any disease positive participants have diagnosis unconfirmed by histology • Exclude if > 50% of disease negative participants have diagnosis confirmed by expert opinion with no histology or follow-up • Exclude studies of referral accuracy, i.e. comparing referral decision with expert diagnosis, unless evaluations of teledermatology or mobile phone applications
BCC: basal cell carcinoma; cSCC: cutaneous squamous cell carcinoma; CT: computed tomography; FNAC: fine needle aspiration cytology; LND: lymph node dissection; MRI: magnetic resonance imaging; PET: positron emission tomography; PET-CT: positron emission tomography computed tomography; RCT: randomised controlled trial; SCC: squamous cell carcinoma; SLN+: positive sentinel lymph node; SLn: negative sentinel lymph node; SLNB: sentinel lymph node biopsy.		

6 Quality assessment (based on QUADAS-2)

The QUADAS-2 checklist (Whiting 2011) was tailored to the review topic as follows below.

Patient selection domain (1)

Selective recruitment of study participants can be a key influence on test accuracy. In general terms, all participants eligible to undergo a test should be included in a study, allowing for the intended use of that test within the context of the study. We considered studies that separately sampled malignant and benign lesions to have used a case-control design; and those that supplemented a series of suspicious lesions with additional malignant or benign lesions to be at unclear risk of bias

In terms of exclusions, we considered studies that excluded particular lesion types particular lesion sites, or that excluded lesions on the basis of image quality or lack of observer agreement (e.g. on histopathology) to be at high risk of bias.

In judging the applicability of patient populations to the review question, we considered restriction to particular lesion

populations, such as melanocytic, nodular, high risk or restrictions by size to be of high concern for applicability.

Given that diagnosis of skin cancer is primarily lesion-based, there is the potential for study participants with multiple lesions to contribute disproportionately to estimates of test accuracy, especially if they are at particular risk of having skin cancer. We considered studies that include a high number of lesions in relation to the number of study to be less representative than studies conducted in a more general population participants (i.e. if the difference between the number of included lesions and number of included participants is greater than 5%).

Index test domain (2)

Given the potential for subjective differences in test interpretation, the interpretation of the index test blinded to the result of the reference standard is a key means of reducing bias. For prospective studies and retrospective studies that used the original index test interpretation, the diagnosis will by nature be interpreted and recorded before the result of the reference standard is known; however, studies using previously acquired images could be particularly susceptible to information bias. For these studies to be at low risk of bias, we required a clear indication that observers were unaware of the reference standard diagnosis at time of test interpretation. An item was also added to assess the presence of blinding between interpretations of different algorithms, however this item was not included in the overall assessment of risk of bias.

Pre-specification of the index test threshold was considered present if the study clearly reported that the threshold used was not data driven, i.e. was not based on study results. Studies that did not clearly describe the threshold used but that required clinicians to record a diagnosis or management decision for a lesion were considered to be unclear on this criterion. Studies reporting accuracy for multiple numeric thresholds, where ROC analysis was used to select the threshold, or that reported accuracy for the presence of independently significant lesion characteristics with no separate test set of lesions were considered at high risk of bias.

In terms of applicability of the index test to the review question, we required the test to be applied and interpreted as it would be in a clinical practice setting, i.e. in person or face-to-face with the patient, and by a single observer as opposed to a consensus decision or average across multiple observers. Image-based studies were considered to be high concern, although RCM image interpretations where the observer was also supplied with a clinical or dermoscopic image of the lesion along with some patient characteristics were considered ‘unclear’.

Despite the often subjective nature of test interpretation, it is also important for study authors to outline the particular lesion characteristics that were considered to be indicative for skin cancer, particularly where established algorithms or checklists were not used. Studies were considered of low concern if the threshold used was established in a prior study or sufficient threshold details were presented to allow replication.

The experience of the examiner will also impact on the applicability of study results. We required studies to describe the test interpreter as ‘experienced’ or ‘expert’ in RCM to have low concern about applicability.

Reference standard domain (3)

In an ideal study, consecutively recruited participants should all undergo incisional or excisional biopsy of the skin lesion regardless of level of clinical suspicion. In reality, both partial and differential verification bias are likely. Partial verification bias may occur where histology is the only reference standard used, and only those participants with a certain degree of suspicion of malignancy based on the result of the index test undergo verification, the others either being excluded from the study or defined as being disease-negative without further assessment or follow-up, as discussed above.

Differential verification bias will be present where other reference standards are used in addition to histological verification of suspicious lesions. A typical example of verification bias in skin cancer occurs when investigators do not biopsy people with benign-appearing lesions but instead follow them up for a period of time to determine whether any malignancy subsequently develops (these would be false-negatives on the index test). We defined an 'adequate' reference standard as: all disease-positive individuals having a histological reference standard either at the time of application of the index test or after a period of clinical follow-up; and at least 80% of disease-negative participants have received a histological diagnosis, with up to 20% undergoing at least three months' follow-up of benign-appearing lesions.

A further challenge is the potential for incorporation bias, i.e. where the result of the index test is used to help determine the reference standard diagnosis. It is normal practice for the clinical diagnosis (usually by visual inspection or dermoscopy) to be included on pathology request forms and for the histopathologist to use this diagnosis to help with the pathology interpretation. Although inclusion of such clinical information on the histopathology request form is theoretically a form of incorporation bias, blinded interpretation of the histopathology reference standard is not normal practice, and enforcement of such conditions would significantly limit the generalisability of the study results. For studies evaluating RCM, this item was divided into two questions, firstly whether the reference standard was blinded to the index test result (RCM), and secondly whether it was blinded to the clinical diagnosis. Only the response to the first part (i.e. blinding to RCM) was included in our overall assessment of risk of bias for the reference standard domain.

In judging the applicability of the reference standard to our review question, scored studies as high concern around applicability if they used expert diagnosis (with no follow-up) as a reference standard in any patient, or did not report histology interpretation by a dermatopathologist.

Flow and timing domain (4)

In the ideal study, the diagnosis based on the index test and reference standard should be made consecutively or as near to each other in time as possible to avoid changes in lesion over time. For lesions with a histological reference standard, we have defined a one-month period as an appropriate interval between application of the index test and the reference standard.

For studies using clinical follow-up, a minimum three-month follow-up period has been defined as at low risk of bias for detecting false-negatives.

In assessing whether all patients were included in the analysis, we considered studies at high risk of bias if participants were excluded following recruitment.

Comparative domain

A comparative domain was added to the QUADAS-2 checklist for studies comparing the accuracy of RCM and dermoscopy. Items were included to assess the presence blinding of interpretation between tests, and to specify a maximum of one month interval between application of index tests, as intervals greater than these may be accompanied by changes in tumour characteristics. As it would not be normal practice for RCM to be interpreted blinded to the clinical or dermoscopic diagnosis, the scoring of this item did not contribute to our overall assessment of risk of bias. We also considered whether both tests were applied and interpreted in a clinically applicable manner.

The following tables use text that was originally published in the QUADAS-2 tool by Whiting and colleagues (Whiting 2011 *).*

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
1) Was a consecutive or random sample of participants or images enrolled?	**Yes** – if paper states consecutive or random **No** – if paper describes other method of sampling **Unclear** – if participant sampling not described
2) Was a case-control design avoided?	**Yes** – if consecutive or random or case-control design clearly not used **No** – if study described as case-control or describes sampling specific numbers of participants with particular diagnoses **Unclear** – if not described
3) Did the study avoid inappropriate exclusions, e.g. • 'difficult to diagnose' lesions not excluded • lesions not excluded on basis of disagreement between evaluators	**Yes** if inappropriate exclusions were avoided **No** – if lesions were excluded that might affect test accuracy, e.g. 'difficult to diagnose' lesions, or where disagreement between evaluators was observed **Unclear** – if not clearly reported but there is suspicion that difficult to diagnose lesions may have been excluded
4) For between-person comparative studies only (i.e. allocating different tests to different study participants): • **A)** were the same participant selection criteria used for those allocated to each test? • **B)** was the potential for biased allocation between tests avoided through adequate generation of a randomised sequence? • **C)** was the potential for biased allocation between tests avoided through concealment of allocation prior to assignment?	**For A)** • **Yes** – if same selection criteria were used for each index test, **No** – if different selection criteria were used for each index test, **Unclear** – if selection criteria per test were not described, **N/A** – if only 1 index test was evaluated or all participants received all tests **For B)** • **Yes** – if adequate randomisation procedures are described, **No** – if inadequate randomisation procedures are described, **Unclear** – if the method of allocation to groups is not described (a description of 'random' or 'randomised' is insufficient), **N/A** – if only 1 index test was evaluated or all participants received all tests **For C)** • **Yes** – if appropriate methods of allocation concealment are described, **No** – if appropriate methods of allocation concealment are not described, **Unclear** – if the method of allocation concealment is not described (sufficient detail to allow a definite judgement is required), **N/A** – if only 1 index test was evaluated

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
Could the selection of participants have introduced bias? **For non-comparative and within person-comparative studies** 1. If answers to all of questions 1), 2), and 3) 'Yes': 2. If answers to any 1 of questions 1), 2), or 3) 'No': 3. If answers to any 1 of questions 1), 2), or 3) 'Unclear': **For between-person comparative studies** 1. If answers to all of questions 1), 2), 3), and 4) 'Yes': 2. If answers to any 1 of questions 1), 2), 3), or 4) 'No': 3. If answers to any 1 of questions 1), 2), 3), or 4) 'Unclear':	**For non-comparative and within person-comparative studies** 1. Risk is low 2. Risk is high 3. Risk unclear **For between-person comparative studies** 1. Risk is low 2. Risk is high 3. Risk unclear
PARTICIPANT SELECTION (1) CONCERNS REGARDING APPLICABILITY	
1) Are the included participants and chosen study setting appropriate to answer the review question, i.e. are the study results generalisable? • This item is not asking whether exclusion of certain participant groups might bias the study's results (as in Risk of Bias above), but is asking whether the chosen study participants and setting are appropriate to answer our review question. Because we are looking to establish test accuracy in both primary presentation and referred participants, a study could be appropriate for 1 setting and not for the other, or it could be unclear as to whether the study can appropriately answer either question • For each study assessed, please consider whether it is more relevant for A) participants with a primary presentation of a skin lesion or B) referred participants, and respond to the questions in either A) or B) accordingly. If the study gives insufficient details, please respond **Unclear** to both parts of the question	**A) For studies that will contribute to the analysis of participants with a primary presentation of a skin lesion (i.e. test naive)** **Yes** – if participants included in the study appear to be generally representative of those who might present in a usual practice setting **No** – if study participants appear to be unrepresentative of usual practice, e.g. in terms of severity of disease, demographic features, presence of differential diagnosis or comorbidity, setting of the study, and previous testing protocols **Unclear –** if insufficient details are provided to determine the generalisability of study participants **B) For studies that will contribute to the analysis of referred participants (i.e. who have already undergone some form of testing)** **Yes** – if study participants appear to be representative of those who might be referred for further investigation. If the study focuses only on those with equivocal lesions, for example, we would suggest that this is not representative of the wider referred population **No** – if study participants appear to be unrepresentative of usual practice, e.g. if a particularly high proportion of participants have been self-referred or referred for cosmetic reasons. Other factors to consider include severity of disease, demographic features, presence of differential diagnosis or comorbidity, setting of the study, and previous testing protocols **Unclear –** if insufficient details are provided to determine the generalisability of study participants
2) Did the study **avoid including** participants with multiple lesions?	**Yes** – if the difference between the number of included lesions and number of included participants is less than 5% **No** – if the difference between the number of included lesions and number of included participants is greater than 5% **Unclear –** if it is not possible to assess
Is there concern that the included participants do not match the review question? 1. If the answer to question 1) or 2) 'Yes': 2. If the answer to question 1) or 2) 'No': 3. If the answer to question 1) or 2) 'Unclear':	1. Concern is low 2. Concern is high 3. Concern is unclear

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
INDEX TEST (2) RISK OF BIAS (to be completed per test evaluated)	
1) Was the index test or testing strategy result interpreted without knowledge of the results of the reference standard?	**Yes** – if index test described as interpreted without knowledge of reference standard result or, for prospective studies, if index test is always conducted and interpreted prior to the reference standard **No** – if index test described as interpreted in knowledge of reference standard result **Unclear** – if index test blinding is not described
2) Was the diagnostic threshold at which the test was considered positive prespecified?	**Yes** – if threshold was prespecified (i.e. prior to analysing study results) **No** – if threshold was not prespecified **Unclear** – if not possible to tell whether or not diagnostic threshold was prespecified
3) For within-person comparisons of index tests or testing strategies (i.e. > 1 index test applied per participant): was each index test result interpreted without knowledge of the results of other index tests or testing strategies?	**Yes** – if all index tests were described as interpreted without knowledge of the results of the others **No** – if the index tests were described as interpreted in the knowledge of the results of the others **Unclear** – if it is not possible to tell whether knowledge of other index tests could have influenced test interpretation **N/A** – if only 1 index test was evaluated
Could the conduct or interpretation of the index test have introduced bias? **For non-comparative and between-person comparison studies** 1. If answers to questions 1) and 2) 'Yes': 2. If answers to either questions 1) or 2) 'No': 3. If answers to either questions 1) or 2) 'Unclear': **For within-person comparative studies** 1. If answers to all questions 1), 2), for any index test and 3) 'Yes': 2. If answers to any 1 of questions 1) or 2) for any index test or 3) 'No': 3. If answers to any 1 of questions 1) or 2) for any index test or 3) 'Unclear':	**For non-comparative and between-person comparison studies** 1. Risk is low 2. Risk is high 3. Risk is unclear **For within-person comparative studies** 1. Risk is low 2. Risk is high 3. Risk is unclear
INDEX TEST (2) CONCERN ABOUT APPLICABILITY	
1) Was the diagnostic threshold to determine presence or absence of disease established in a previously published study? E.g. previously evaluated/established • algorithm/checklist used • lesion characteristics • objective (usually numerical) threshold used	**Yes** – if a previously evaluated/established tool to aid diagnosis was used or if the diagnostic threshold used was established in a previously published study **No** – if an unfamiliar/new tool to aid diagnosis was used, if no particular algorithm was used, or if the objective threshold reported was chosen based on results in the current study **Unclear** – if insufficient information was reported
2) Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication? Study results can only be reproduced if the diagnostic threshold is described in sufficient detail. This item applies equally to studies using pattern recognition and those using checklists or algorithms to aid test interpretation	**Yes** – If the criteria for diagnosis were reported in sufficient detail to allow replication **No** – if the criteria for diagnosis were not reported in sufficient detail to allow replication **Unclear** – If some but not sufficient information on criteria for diagnosis to allow replication were provided

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
3) Was the test interpretation carried out by an experienced examiner?	**Yes** – if the test was interpreted by 1 or more speciality-accredited dermatologists, or by examiners of any clinical background with special interest in dermatology and with any formal training in the use of the test **No** – if the test was not interpreted by an experienced examiner (see above) **Unclear** – if the experience of the examiner(s) was not reported in sufficient detail to judge or if examiners were described as 'Expert' with no further detail given **N/A** – if system-based diagnosis, i.e. no observer interpretation
Is there concern that the index test, its conduct, or interpretation differ from the review question? 1. If answers to questions 1), 2), and 3) 'Yes': 2. If answers to questions 1), 2), or 3) 'No': 3. If answers to questions 1), 2), or 3) 'Unclear':	1. Concern is low 2. Concern is high 3. Concern is unclear
REFERENCE STANDARD (3) RISK OF BIAS	
1) Is the reference standard likely to correctly classify the target condition? A) Disease-positive – 1 or more of the following: • histological confirmation of malignancy following biopsy or lesion excision • clinical follow-up of benign-appearing lesions for at least 3 months following the application of the index test, leading to a histological diagnosis of skin cancer B) Disease-negative – 1 or more of the following: • histological confirmation of absence of malignancy following biopsy or lesion excision in at least 80% of disease-negative participants • clinical follow-up of benign-appearing lesions for a minimum of 3 months following the index test in up to 20% of disease-negative participants	A) Disease-positive **Yes** – if all participants with a final diagnosis of malignancy underwent 1 of the listed reference standards **No** – If a final diagnosis of malignancy for any participant was reached without histopathology **Unclear** – if the method of final diagnosis was not reported for any participant with a final diagnosis of malignancy or if the length of clinical follow-up used was not clear or if a clinical follow-up reference standard was reported in combination with a participant-based analysis and it was not possible to determine whether the detection of a malignant lesion during follow-up is the same lesion that originally tested negative on the index test B) Disease-negative **Yes** – If at least 80% of benign diagnoses were reached by histology and up to 20% were reached by clinical follow-up for a minimum of 3 months following the index test **No** – if more than 20% of benign diagnoses were reached by clinical follow-up for a minimum of 3 months following the index test or if clinical follow-up period was less than 3 months **Unclear** – if the method of final diagnosis was not reported for any participant with benign or non-melanoma diagnosis
2) Were the reference standard results interpreted without knowledge of the results of the index test? Please score this item for all studies even though histopathology interpretation is usually conducted with knowledge of the clinical diagnosis (from visual inspection or dermoscopy or both). We will deal with this by not including the response to this item in the 'Risk of bias' assessment for these tests. For reviews of all other tests, this item will be retained	**Yes** – if the reference standard diagnosis was reached blinded to the index test result **No** – if the reference standard diagnosis was reached with knowledge of the index test result **Unclear** – if blinded reference test interpretation was not clearly reported

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
Could the reference standard, its conduct, or its interpretation have introduced bias? **For visual inspection/dermoscopy evaluations** 1. If answer to question 1) 'Yes': 2. If answer to question 1) 'No': 3. If answer to question 1) 'Unclear': **For all other tests** 1. If answers to questions 1) and 2) 'Yes': 2. If answers to questions 1) or 2) 'No': 3. If answers to questions 1) or 2) 'Unclear':	**For visual inspection/dermoscopy evaluations** 1. Risk is low 2. Risk is high 3. Risk is unclear **For all other tests** 1. Risk is low 2. Risk is high 3. Risk is unclear
REFERENCE STANDARD (3) CONCERN ABOUT APPLICABILITY	
1) Expert opinion (with no histological confirmation) was not used as a reference standard 'Expert opinion' means diagnosis based on the standard clinical examination, with no histology or lesion follow-up ***do not complete this item for teledermatology studies	**Yes** – if expert opinion was not used as a reference standard for any participant **No** – if expert opinion was used as a reference standard for any participant **Unclear** – if not clearly reported
2) Was histology interpretation carried out by an experienced histopathologist or by a dermatopathologist?	**Yes –** if histology interpretation was reported to be carried out by an experienced histopathologist or dermatopathologist **No –** if histology interpretation was reported to be carried out by a less experienced histopathologist **Unclear** – if the experience/qualifications of the pathologist were not reported
Is there concern that the target condition as defined by the reference standard does not match the review question? 1. If answers to both questions 1), 2), 'Yes': 2. If answers to any 1 of questions 1), 2), 'No': 3. If answers to any 1 of questions 1), 2), 'Unclear': ***For teledermatology studies only** 1. If answers to all questions 1) and 3) 'Yes': 2. If answers to questions 1) or 3) 'No': 3. If answers to questions 1) or 3) 'Unclear':	1. Concern is low 2. Concern is high 3. Concern is unclear ***For teledermatology studies only** 1. Concern is low 2. Concern is high 3. Concern is unclear
FLOW AND TIMING (4): RISK OF BIAS	
1) Was there an appropriate interval between index test and reference standard? **A)** For histopathological reference standard, was the interval between index test and reference standard ≤ 1 month? **B)** If the reference standard includes clinical follow-up of borderline/benign-appearing lesions, was there at least 3 months' follow-up following application of index test(s)?	**A)** **Yes** – if study reports ≤ 1 month between index and reference standard **No** – if study reports > 1 month between index and reference standard **Unclear** – if study does not report interval between index and reference standard **B)** **Yes** – if study reports ≥ 3 months' follow-up **No** – if study reports < 3 months' follow-up **Unclear** – if study does not report the length of clinical follow-up

Item	Response (delete as required)
PARTICIPANT SELECTION (1) RISK OF BIAS	
2) Did all participants receive the same reference standard?	**Yes** – if all participants underwent the same reference standard **No** – if more than 1 reference standard was used **Unclear** – if not clearly reported
3) Were all participants included in the analysis?	**Yes** – if all participants were included in the analysis **No** – if some participants were excluded from the analysis **Unclear**– if not clearly reported
4) **For within-person comparisons of index tests** Was the interval between application of index tests ≤ 1 month?	**Yes** – if study reports ≤ 1 month between index tests **No** – if study reports > 1 month between index tests **Unclear** – if study does not report the interval between index tests
Could the participant flow have introduced bias? **For non-comparative and between-person comparison studies** 1. If answers to questions 1), 2), and 3) 'Yes': 2. If answers to any 1 of questions 1), 2), or 3) 'No': 3. If answers to any 1 of questions 1), 2), or 3) 'Unclear': **For within-person comparative studies** 1. If answers to all questions 1), 2), 3), and 4) 'Yes': 2. If answers to any 1 of questions 1), 2), 3), or 4) 'No': 3. If answers to any 1 of questions 1), 2), 3), or 4) 'Unclear':	**For non-comparative and between-person comparison studies** 1. Risk is low 2. Risk is high 3. Risk is unclear **For within-person comparative studies** 1. Risk is low 2. Risk is high 3. Risk is unclear
BCC = basal cell carcinoma; cSCC = cutaneous squamous cell carcinoma.	

7 Details of RCM algorithms and diagnostic thresholds for diagnosis

Castro 2015 BCC	Guitera 2012 (two step algorithm to id BCC then MM)	Longo 2013 BCC (and MM)	Nori 2004 (based on Gonzalez 2002) BCC
Previously-published RCM criteria assessed (cites Agero 2006; Nori 2004; Guitera 2012) Selected characteristics chosen; ≥3 RCM criteria present, including either presence of § 'dark silhouettes' or § 'bright tumor islands' Additional criteria assessed: § 'streaming' polarization of nuclei in neoplastic aggregates along the same axis of orientation; § 'peripheral palisading' of nuclei at the tumor islands' periphery; § dark 'peritumoral clefts' around the tumor islands; § fibrotic stroma with 'thickened collagen bundles'; § dilated and tortuous 'linear blood vessels' and 'coiled blood vessels'; § 'bright dendritic structures' within tumor islands; and § 'bright round cells' in the stroma.	Evaluated 47 RCM features (referenced to a number of prior studies) and conducted multivariate analysis on the training set of lesions to identify independently significant features for MM and for BCC; assume presence of any one indicated T+] Correct id as MM or BCC (based on independently significant features as id from training set) For BCC: § Polarized in the honeycomb § Linear telangiectasia-like horizontal vessels § Basaloid cord or nodule § Epidermal shadow § Convoluted glomerular-like vessels § Non-visible papillae § Cerebriform nests § Disarray of the epidermal layer	47 RCM features recorded; multivariate analysis id 4 positive independent significant features for BCC § tumour islands (dark silhouettes or tightly packed basaloid islands); § cauliflower architecture; § bright filaments within the tumour islands; and § presence of bright collagen.	Selected 5 criteria from a number of morphologic characteristics previously investigated by the same group, on the basis that they were 'easily and unambiguously detected by non-dermatopathologists and a novice reviewer' Data presented for >=2, >=3, >=4, =5 chars present § elongated monomorphic basaloid nuclei; § polarization of these nuclei along the same axis of orientation; § prominent inflammatory infiltrate; § increased dermal vasculature; § pleomorphism of the overlying epidermis indicative of actinic changes.
Incel 2015 BCC/SCC	Results based on 'observer diagnosis'		

Castro 2015 BCC	Guitera 2012 (two step algorithm to id BCC then MM)	Longo 2013 BCC (and MM)	Nori 2004 (based on Gonzalez 2002) BCC
Selected characteristics from to assist correct diagnosis of different lesion types (cites Malvehy 2012; Eichert 2010; Ahlgrimm-Siess 2010; Röwert-Huber 2007; Ahlgrimm-Siess 2011) Characteristics listed for BCC included: § Dark silhouettes in dermis, § Bright tumour islands at DEJ and in the dermis; § Cleft-like dark areas; § Dendritic cells, § Bright round cells, § Canalicular vessels. Characteristics listed for SCC included: Refractile squam/crust in stratum corneum and nucleated cells with dark center (parakeratotic)cells; § Atypical honeycomb pattern, disarranged pattern at stratum granulosum layer; § Large, round, nucleated cells at the granular layer (dyskeratotic cells); § Dendritic cells at the granular layer and small edged papillae at DEJ;	Curchin 2011: applied RCM score: >=3 for suspected melanomas and LM score for suspected lentigo maligna of the face (Guitera 2010); reports observer correct diagnosis of BCC; No further details presented Farnetani 2015: Evaluators completed a 'pattern description' (presence/absence of a number of RCM features) and gave an overall diagnosis of malignant (melanoma or BCC) or benign Discriminant analysis also used to id features independently associated with malignancy (and with MM and BCC separately) 3 more frequently observed in BCC were: § basaloid cord–like structures, § presence of ulceration, § a specific DEJ pattern Pellacani 2014: presented Rao 2013: Observers gave diagnosis (MM/BCC/SCC) and excise decision (no further details) Witkowski 2016: Report correct diagnosis (MM/BCC/SCC)and excise decision (no further details)		

8 Summary study details by lesion population

Study author Outcomes reported	Study type Country Setting	Inclusion criteria	No. patients / lesions	Index tests (algorithm) Diagnostic approach	Observer qual. (n); experience Additional data available	Reference standard Final diagnoses	Exclusions
Any suspicious lesion							

Study author Outcomes reported	Study type Country Setting	Inclusion criteria	No. patients / lesions	Index tests (algorithm) Diagnostic approach	Observer qual. (n); experience Additional data available	Reference standard Final diagnoses	Exclusions
Any suspicious lesion							
Curchin 2011 BCC KER (MEL)	NC P-CS Australia Secondary	Patient scheduled for minor excision	42 / 50	VivaScope 1500; Observer diagnosis (RCM score for suspected melanomas) In-person (Single observer)	Observer qual NR; (n=NR) Described as no ce to RCM analysis after completing a RCM analysis course in Modena, Italy. Dermoscopic and RCM images were aligned over the top of each other	Histology alone MM 12; MiS 1; BCC: 9; cSCC: 6 (includes SK or AK, or both) 'Benign' diagnoses: 23	Reported correct diagnosis of all 6 SCC or precursors (not disaggregated)
Guitera 2012 BCC (MEL)	WPC NR-CS Australia/ Italy Specialist clinic/Secondary	Patients with suspicious lesions, including those on face and neck suspicious for LM, and requiring histology to rule out an epithelial tumour or an MM; predominantly melanocytic or suspicious for BCC	663 / 710 356 lesions randomised to 'test set' included here	VivaScope 1500; No algorithm (independently significant features for BCC) Image-based (Single observer)	Dermatologist (n=2); described as expert observers RCM guided by dermoscopic findings but interpretation blinded to all but lesion location and patient age	Histology alone MEL 105; BCC: 52; cSCC: 9 BN 132; SN 16; AK 8; 31 benign macule of the face and 3 DF	BCC: 2MM and 2 SCC were FP
Pellacani 2014b (doc) BCC (MEL)	NC P-CS Italy Specialist clinic	Patients requesting a mole check or with suspicion of melanoma who were referred to PLC clinic. Documentation group (doc) includes lesions with consistent clinical or dermoscopic criteria, or both, for melanoma diagnosis	171 / 184 (1/184 did not undergo RCM)	VivaScope 1500; Observer diagnosis (assumed RCM score for suspected melanomas) In-person (Single obs)	Observer qual NR (n=1); diagnosis made at 'confocal unit' RCM reader was aware that lesions were dermoscopically atypical but blinded to 'RCM documentation' or 'RCM consultation'	Histology alone (documentation group) MM 13; MiS 9; BCC: 19; 1 mel mets; BN 121; SN 8; SK or other keratotic 7; other benign 5	

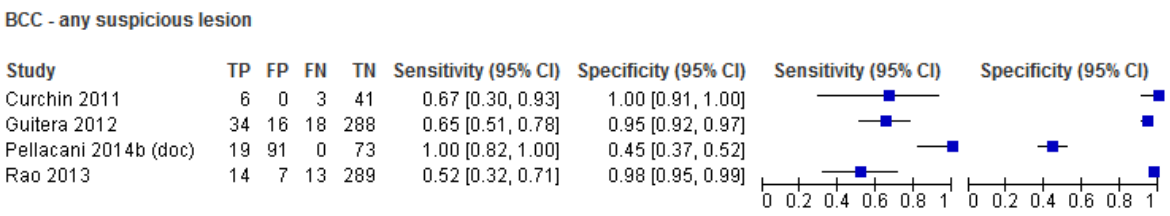
Study author Outcomes reported	Study type Country Setting	Inclusion criteria	No. patients / lesions	Index tests (algorithm) Diagnostic approach	Observer qual. (n); experience Additional data available	Reference standard Final diagnoses	Exclusions
Any suspicious lesion							
Rao 2013 BCC SCC KER (MEL)	NC NR-CS US Secondary	All lesions removed for cosmetic or medical reasons that were imaged using a confocal scanning microscope	NR / 334 Reader 1 (novice) evaluated 318 lesions; Reader 2 (expert) evaluated 323 lesions; 284 were examined by both readers	VivaScope 1500; No algorithm (correct dx of each malignancy; overall observer diagnosis of malignancy) Image-based (Single observer)	Observer qual NR (n=2); Reader 1 had 1 year RCM experience at the start of the study; Reader 2 had > 9 years' experience with RCM. Diagnosis was based on the dermoscopic image and confocal microscopy evaluation	Histology alone MM 8; Melanoma (in situ); 1; BCC: 27; cSCC: 42 BN 176; SK 22; AK 24; 23 other	BCC: 4 SCC were FP SCC: 9 BCCs picked up as SCCs were considered TN as per Methods
Equivocal lesion studies							
Farnetani 2015 BCC (MEL)	NC R-CS Italy Secondary	Diagnostically equivocal lesions excised due to clinical or dermoscopic suspicion of melanoma, where a specific clinical and dermoscopic diagnosis could not be rendered with certainty	NR / 100	VivaScope 1500; No algorithm (observer dx - pattern description and diagnostic judgment) Image-based (Single observer)	Dermatologist (n=9); 6 experienced (>=3 years RCM experience) and 3 'recent' RCM users. Experienced reader randomly selected for primary analysis. Dermoscopic image provided; no additional clinical information (eg, age and melanoma or lesion history)	Histology alone MEL 20; BCC: 15 SK 7; BN 55; AK 3	BCC: 14 MM FP

Study author Outcomes reported	Study type Country Setting	Inclusion criteria	No. patients / lesions	Index tests (algorithm) Diagnostic approach	Observer qual. (n); experience Additional data available	Reference standard Final diagnoses	Exclusions
Any suspicious lesion							
Pellacani 2014a (cons) BCC (MEL)	NC P-CS Italy Specialist clinic	Patients requesting a mole check or with suspicion of melanoma who were referred to PLC clinic. Consultation group (cons) includes lesions requiring an outcome decision from RCM (dx could not be reached on clinical or dermoscopic criteria, or both)	252 / 309 1/308 did not undergo RCM	VivaScope 1500; Observer diagnosis (assumed RCM score for suspected melanomas) In-person (Single observer)	Observer qual NR (n=1); diagnosis made at 'confocal unit' RCM reader was aware that lesions were dermoscopically atypical but blinded to 'RCM documentation' or 'RCM consultation'	Histology plus FU (cons group); 227/308 referred for sequential digital FU; 28 later excised MM 2; MiS 4; BCC: 19; BN 71; SN 5; benign NML 8; 199 benign on FU	
Witkowski 2016 BCC SCC MM	WPC-tests R-CS Italy Secondary	Clinically equivocal 'pink' cutaneous lesions with absent pigmentation or containing less than 10% pigment and absence of pigment network.	NR / 260	VivaScope 1500; No algorithm (correct dx of each malignancy) Dermoscopy Image-based (Single observer)	Dermatologist (assumed; n=2, 1 dermoscopy 1 RCM); experience NR No additional information provided	Histology alone MEL 12; BCC: 114; cSCC: 13; 1 syringoid eccrine carcinoma Benign keratotic 25; BN 47; SN 6; DF 18; other benign 24	BCC: 1 SCC FP
Other lesion populations							
Castro 2015 BCC	WPC NR-CS Brazil and USA Specialist clinic	Patients with one or more skin lesions deemed suspicious for BCC based on clinical and dermoscopic examination.	NR/54	Vivascope 1500 Vs Vivascope 3000 (No algorithm; >= 3 characteristics present) Unclear if image-based; consensus of 2	Dermatologist (n=1); experienced with RCM examination and supervised by skin cancer expert Clinical, dermoscopic and RCM imaging performed by same dermatologist	Histology BCC: 45; 'Benign' diagnoses: 9	38 of original 92 lesions excluded as only accessible to Vivascope 3000 (mostly facial).

Study author Outcomes reported	Study type Country Setting	Inclusion criteria	No. patients / lesions	Index tests (algorithm) Diagnostic approach	Observer qual. (n); experience Additional data available	Reference standard Final diagnoses	Exclusions
Any suspicious lesion							
Incel 2015 BCC SCC	NC P-CS Turkey Secondary	Patients with nonpigmented suspected tumoral lesions or proliferative skin lesions and with a vascular structure on dermoscopic examination	114/122	Vivascope 3000 (No algorithm; selected characteristics; correct dx of BCC/SCC) Unclear if image-based; unclear if single observer)	Observer qual NR (n=NR); states “First 60 lesions subjected to blinded evaluation by 2 observers”, no further details provided	Histology BCC: 56; cSCC: 9 KA 3; SK 11; AK 8; BD 7; and 22 other benign nonpigmented tumours	BCC: All SCCs considered test negative
Longo 2013 BCC SCC KER (MEL)	NC R-CS Italy Specialist clinic/Secondary	Clinically nodular lesions (defined as cutaneous palpable/superficial seated lesions and not subcutaneous ones) that underwent excision	140/140	Vivascope 1500 (NR but assumed to be used; correct diagnosis of BCC/SCC) Image-based (Single observer)	Dermatologist (n=1); 5 years’ experience in RCM Blinded to dermoscopy	Histology NM 23; BCC: 28; cSCC: 6; Other malignant: 9 mel mets SK 14; BN 32 including 7 SN); 5 vascular; 6 other benign	Non evaluable and non specific results excluded (n=11); including 1 BCC and 1 SCC
Nori 2004 BCC	WPC CCS US and Spain Secondary/Private clinic	Biopsy confirmed BCC and convenience sample of non-BCC with 'range of common diagnoses'	145/152 105 had VI diagnosis	Vivascope 1000 (No algorithm; selected characteristics; >= 3 present) VI (clinical photographs; high/medium/low probability BCC) Image-based (Single observer)	Observer qual NR (n=1); 'Novice confocal reviewer' Blinded interpretation	Histology or expert diagnosis; 15/65 benign had clinical dx BCC: 83 (58 in VI analysis); cSCC: 4 'Benign' diagnoses: 65	BCC: Cannot disaggregate SCC result (n=4) from rest of D- group

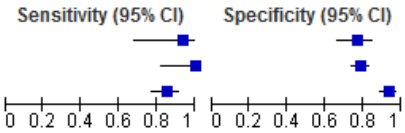
NR – not reported; PLC – pigmented lesion clinic; MM – malignant melanoma; MiS – melanoma *in situ* (or lentigo maligna); BCC – basal cell carcinoma; cSCC – cutaneous squamous cell carcinoma; LS – lentigo simplex; SK – seborrheic keratosis; SN – Spitz nevi; AK – actinic keratosis; BN – benign naevi; BD – Bowen’s disease; DF – dermatofibroma; FU – follow-up; R –retrospective; P – prospective; CS – case series; CCS – case control study; WPC – within person comparison (of tests); BPC – between person comparison (of tests); NC – non comparative; RCM – reflectance confocal microscopy; Cons - consensus diagnosis; exp - experience; VI - visual inspection; dx - diagnosis

Graphs



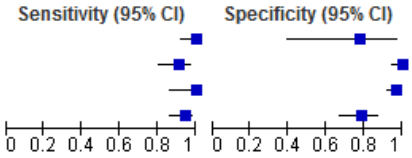
BCC - equivocal lesions

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	14	20	1	65	0.93 [0.68, 1.00]	0.76 [0.66, 0.85]
Pellacani 2014a (cons)	19	62	0	227	1.00 [0.82, 1.00]	0.79 [0.73, 0.83]
Witkowski 2016	97	9	17	137	0.85 [0.77, 0.91]	0.94 [0.89, 0.97]



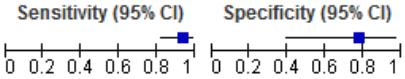
BCC - other lesion populations

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Castro 2015	45	2	0	7	1.00 [0.92, 1.00]	0.78 [0.40, 0.97]
Incel 2015	51	0	5	66	0.91 [0.80, 0.97]	1.00 [0.95, 1.00]
Longo 2013	25	3	0	101	1.00 [0.86, 1.00]	0.97 [0.92, 0.99]
Nori 2004	78	15	5	54	0.94 [0.86, 0.98]	0.78 [0.67, 0.87]



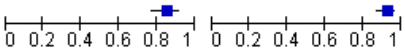
BCC - RCM - other - Vivascope 3000

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Castro 2015	42	2	3	7	0.93 [0.82, 0.99]	0.78 [0.40, 0.97]



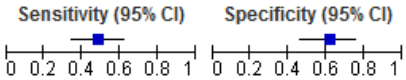
BCC - Dermoscopy - equivocal lesions

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	97	11	17	135	0.85 [0.77, 0.91]	0.92 [0.87, 0.96]



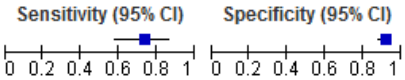
BCC - Visual inspection - other lesion populations

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nori 2004	28	18	30	29	0.48 [0.35, 0.62]	0.62 [0.46, 0.75]



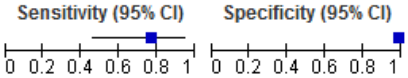
SCC - RCM - all comer

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	31	23	11	258	0.74 [0.58, 0.86]	0.92 [0.88, 0.95]



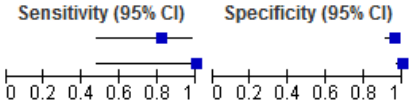
SCC - RCM - equivocal

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	10	4	3	243	0.77 [0.46, 0.95]	0.98 [0.96, 1.00]



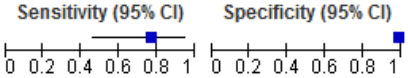
SCC - RCM - other

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Incel 2015	9	4	2	107	0.82 [0.48, 0.98]	0.96 [0.91, 0.99]
Longo 2013	5	0	0	124	1.00 [0.48, 1.00]	1.00 [0.97, 1.00]



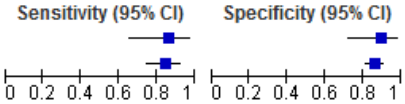
SCC - Dermoscopy - equivocal

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	10	3	3	244	0.77 [0.46, 0.95]	0.99 [0.96, 1.00]



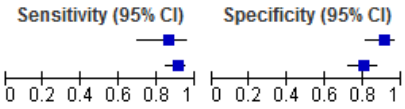
KER - RCM - all comer

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Curchin 2011	19	3	3	25	0.86 [0.65, 0.97]	0.89 [0.72, 0.98]
Rao 2013	66	34	12	211	0.85 [0.75, 0.92]	0.86 [0.81, 0.90]



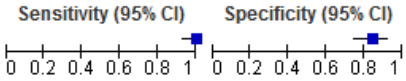
KER - RCM - equivocal

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	30	6	5	59	0.86 [0.70, 0.95]	0.91 [0.81, 0.97]		
Witkowski 2016	127	24	13	96	0.91 [0.85, 0.95]	0.80 [0.72, 0.87]		



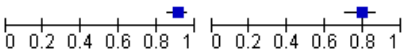
KER - RCM - other

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Longo 2013	51	12	0	66	1.00 [0.93, 1.00]	0.85 [0.75, 0.92]		



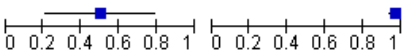
KER - Dermoscopy - equivocal

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	128	25	12	95	0.91 [0.86, 0.95]	0.79 [0.71, 0.86]		



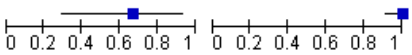
MM2 - RCM - equivocal (non-pigmented) not in melanoma review

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	6	8	6	240	0.50 [0.21, 0.79]	0.97 [0.94, 0.99]		



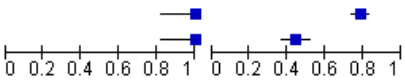
BCC - RCM score at >=3 - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Curchin 2011	6	0	3	41	0.67 [0.30, 0.93]	1.00 [0.91, 1.00]		



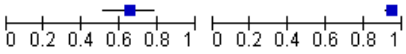
BCC - RCM score at NR (likely >=3) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Pellacani 2014a (cons)	19	62	0	227	1.00 [0.82, 1.00]	0.79 [0.73, 0.83]		
Pellacani 2014b (doc)	19	91	0	73	1.00 [0.82, 1.00]	0.45 [0.37, 0.52]		



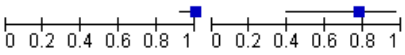
BCC - Guitera Two-step alg (significant chars for BCC) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Guitera 2012	34	16	18	288	0.65 [0.51, 0.78]	0.95 [0.92, 0.97]		



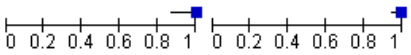
BCC - No algorithm (significant characteristics) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Castro 2015	45	2	0	7	1.00 [0.92, 1.00]	0.78 [0.40, 0.97]		



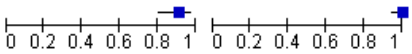
BCC - No algorithm (significant characteristics) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Longo 2013	27	1	0	102	1.00 [0.87, 1.00]	0.99 [0.95, 1.00]		



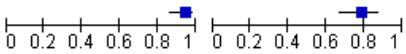
BCC - No algorithm (selected characteristics) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Incel 2015	51	0	5	66	0.91 [0.80, 0.97]	1.00 [0.95, 1.00]		



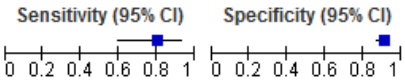
BCC - No algorithm (selected characteristics) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Nori 2004	78	15	5	54	0.94 [0.86, 0.98]	0.78 [0.67, 0.87]		



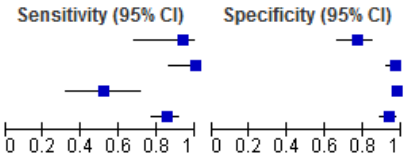
BCC - No algorithm (observer diagnosis) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	20	26	5	267	0.80 [0.59, 0.93]	0.91 [0.87, 0.94]



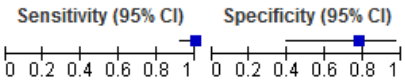
BCC - No algorithm (observer diagnosis) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	14	20	1	65	0.93 [0.68, 1.00]	0.76 [0.66, 0.85]
Longo 2013	25	3	0	101	1.00 [0.86, 1.00]	0.97 [0.92, 0.99]
Rao 2013	14	7	13	289	0.52 [0.32, 0.71]	0.98 [0.95, 0.99]
Witkowski 2016	97	9	17	137	0.85 [0.77, 0.91]	0.94 [0.89, 0.97]



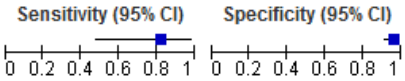
BCC - Handheld RCM - No algorithm (significant characteristics)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Castro 2015	45	2	0	7	1.00 [0.92, 1.00]	0.78 [0.40, 0.97]



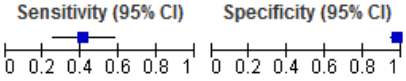
SCC - No algorithm (selected characteristics) in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Incel 2015	9	4	2	107	0.82 [0.48, 0.98]	0.96 [0.91, 0.99]



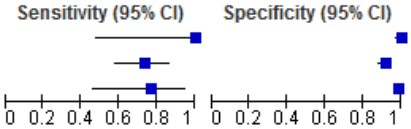
SCC - No algorithm (observer diagnosis) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	16	7	23	272	0.41 [0.26, 0.58]	0.97 [0.95, 0.99]



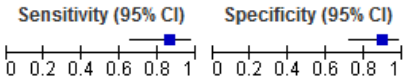
SCC - No algorithm (observer diagnosis) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Longo 2013	5	0	0	124	1.00 [0.48, 1.00]	1.00 [0.97, 1.00]
Rao 2013	31	23	11	258	0.74 [0.58, 0.86]	0.92 [0.88, 0.95]
Witkowski 2016	10	4	3	243	0.77 [0.46, 0.95]	0.98 [0.96, 1.00]



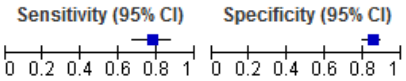
KER - RCM at >=3 - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Curchin 2011	19	3	3	25	0.86 [0.65, 0.97]	0.89 [0.72, 0.98]



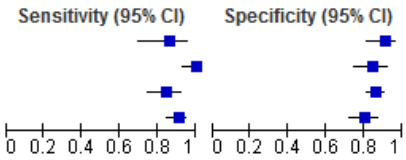
KER - No algorithm (observer diagnosis) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	57	37	16	208	0.78 [0.67, 0.87]	0.85 [0.80, 0.89]



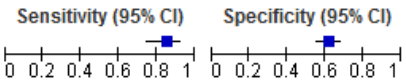
KER - No algorithm (observer diagnosis) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	30	6	5	59	0.86 [0.70, 0.95]	0.91 [0.81, 0.97]
Longo 2013	51	12	0	66	1.00 [0.93, 1.00]	0.85 [0.75, 0.92]
Rao 2013	66	34	12	211	0.85 [0.75, 0.92]	0.86 [0.81, 0.90]
Witkowski 2016	127	24	13	96	0.91 [0.85, 0.95]	0.80 [0.72, 0.87]



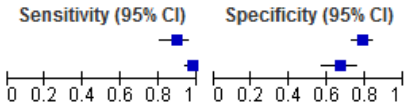
KER - No algorithm (excise decision) - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	62	94	11	151	0.85 [0.75, 0.92]	0.62 [0.55, 0.68]



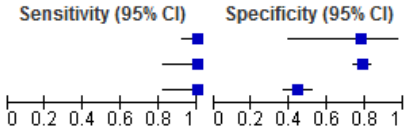
KER - No algorithm (excise decision) - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	70	52	8	193	0.90 [0.81, 0.95]	0.79 [0.73, 0.84]
Witkowski 2016	137	40	3	80	0.98 [0.94, 1.00]	0.67 [0.57, 0.75]



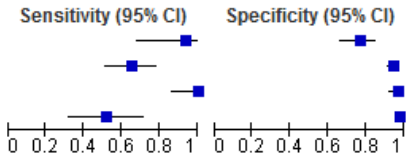
BCC - by observer - high - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Castro 2015	45	2	0	7	1.00 [0.92, 1.00]	0.78 [0.40, 0.97]
Pellacani 2014a (cons)	19	62	0	227	1.00 [0.82, 1.00]	0.79 [0.73, 0.83]
Pellacani 2014b (doc)	19	91	0	73	1.00 [0.82, 1.00]	0.45 [0.37, 0.52]



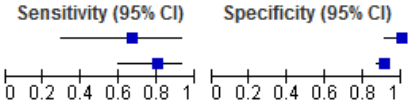
BCC - by observer - high - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	14	20	1	65	0.93 [0.68, 1.00]	0.76 [0.66, 0.85]
Guitera 2012	34	16	18	288	0.65 [0.51, 0.78]	0.95 [0.92, 0.97]
Longo 2013	25	3	0	101	1.00 [0.86, 1.00]	0.97 [0.92, 0.99]
Rao 2013	14	7	13	289	0.52 [0.32, 0.71]	0.98 [0.95, 0.99]



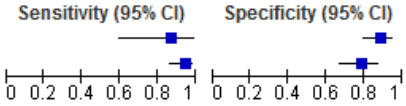
BCC - by observer - low - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Curchin 2011	6	0	3	41	0.67 [0.30, 0.93]	1.00 [0.91, 1.00]
Rao 2013	20	26	5	267	0.80 [0.59, 0.93]	0.91 [0.87, 0.94]



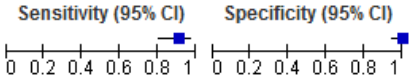
BCC - by observer - low - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	13	10	2	75	0.87 [0.60, 0.98]	0.88 [0.79, 0.94]
Nori 2004	78	15	5	54	0.94 [0.86, 0.98]	0.78 [0.67, 0.87]



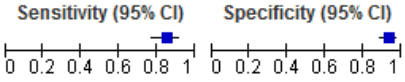
BCC - by observer - NR - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Incel 2015	51	0	5	66	0.91 [0.80, 0.97]	1.00 [0.95, 1.00]



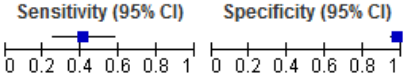
BCC - by observer - NR - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	97	9	17	137	0.85 [0.77, 0.91]	0.94 [0.89, 0.97]



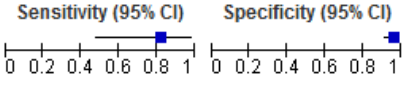
SCC - by observer - low - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rao 2013	16	7	23	272	0.41 [0.26, 0.58]	0.97 [0.95, 0.99]



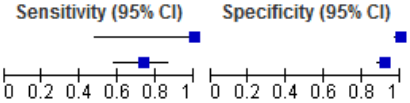
SCC - by observer - NR - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Incel 2015	9	4	2	107	0.82 [0.48, 0.98]	0.96 [0.91, 0.99]



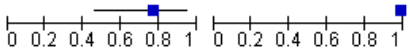
SCC - by observer - high - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Longo 2013	5	0	0	124	1.00 [0.48, 1.00]	1.00 [0.97, 1.00]
Rao 2013	31	23	11	258	0.74 [0.58, 0.86]	0.92 [0.88, 0.95]



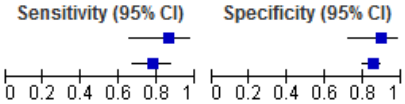
SCC - by observer - NR - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	10	4	3	243	0.77 [0.46, 0.95]	0.98 [0.96, 1.00]		



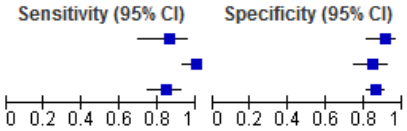
KER - by observer - low - in person

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Curchin 2011	19	3	3	25	0.86 [0.65, 0.97]	0.89 [0.72, 0.98]		
Rao 2013	57	37	16	208	0.78 [0.67, 0.87]	0.85 [0.80, 0.89]		



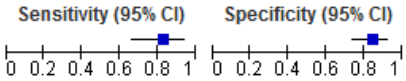
KER - by observer - high - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	30	6	5	59	0.86 [0.70, 0.95]	0.91 [0.81, 0.97]		
Longo 2013	51	12	0	66	1.00 [0.93, 1.00]	0.85 [0.75, 0.92]		
Rao 2013	66	34	12	211	0.85 [0.75, 0.92]	0.86 [0.81, 0.90]		



KER - by observer - low - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Farnetani 2015	29	10	6	55	0.83 [0.66, 0.93]	0.85 [0.74, 0.92]		



KER - by observer - NR - image-based

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Witkowski 2016	127	24	13	96	0.91 [0.85, 0.95]	0.80 [0.72, 0.87]