Article:
Stevens, H.P., Pellacani, G., Angus, C. et al. (2025) The use of Reflectance Confocal Microscopy to diagnose Malignant Melanoma and Lentigo Maligna in the United Kingdom: a prospective observational trial at a single centre. British Journal of Dermatology, 192 (1). pp. 27-35. ISSN: 0007-0963

https://doi.org/10.1093/bjd/ljae354

The use of Reflectance Confocal Microscopy to diagnose Malignant Melanoma and Lentigo Maligna in the United Kingdom: A Prospective Observational Trial at a Single Centre

Running head: Diagnostic Accuracy of Reflectance Confocal Microscopy for diagnosis of Melanoma

Howard P. Stevens,[1] Giovanni Pellacani,[2] Colin Angus[3] and Joseph N. El-Jabbour[4]

[1] The Skin Care Network, Barnet, London, UK
[2] University La Sapienza, Rome, Italy
[3] School of Medicine and Population Health, University of Sheffield, UK
[4] Cellular Pathology Services, Watford, Hertfordshire, UK

Corresponding Author: Dr Howard Stevens
Email: howard@stevens.co.uk

Funding sources: No external funding. Mavig GmbH (now VivaScope GmbH) provided technical and manuscript support at no cost.
Conflicts of interest: None to declare.
Data availability: The data underlying this article will be shared on reasonable request to the corresponding author.
Ethics statement: Protocol number: HS-MAV-002. IRAS project ID: 218996. The study has been reviewed and approved by the sub-committee.
Patient consent: Written patient consent for publication was obtained.

What is known about this topic?

- Dermoscopy increases the diagnostic accuracy of Malignant Melanoma (MM) but this has not been quantified in UK population.
- Reflectance Confocal Microscopy (RCM) has been shown in studies to have potential utility to further improve non-invasive diagnosis of Malignant Melanoma and Lentigo Maligna (LM).
- RCM plus dermoscopy has potential for higher diagnostic accuracy than dermoscopy alone.

What does this study add?

- First UK prospective study quantifying diagnostic accuracy of RCM in MM.
- Support RCM use in UK for MM diagnosis as it increases diagnostic accuracy.
- High diagnostic accuracy of MM with RCM is rapidly achieved.

Abstract

Background
Previous work with Reflectance Confocal Microscopy (RCM) imaging has shown high sensitivity and specificity for Malignant Melanoma (MM), but to date there have been no studies on a UK cohort.

Objectives
The study hypothesised that RCM could be used prospectively to accurately diagnose MM and lentigo maligna (LM) in a private UK secondary care, single clinician setting. The study assessed the potential for RCM to be used as a routine screening procedure.

Methods
597 patients were recruited consecutively where MM or LM featured in the differential diagnosis after clinical examination. A sequential record was made of the clinical, dermoscopic, and RCM findings by a single dermatologist [HS] prior to biopsy. Imaging used the arm-mounted confocal microscope unless access was restricted and required the handheld probe. The likelihood of MM was scored for each modality, each diagnosis building on the last. Histology was assessed by a single blinded histopathologist [JJ].

Results
734 lesions were included in the analysis, including 86 MM and LM with a median diameter of 7.0 mm. The benign to malignant ratio was 3 to 1 (non-melanocytic malignancies included) and 8.3 to 1 for MM and LM only. The sensitivity and specificity for MM and LM was 62.8% (95% CI 51.70% to 72.98%) and 63.2% (59.27% to 66.84%) for clinical examination; 91.9% (83.95% to 96.66%) and 42.1% (38.14% to 45.88%) for dermoscopy; 94.2% (86.95% to 98.09%) and 83.2% (79.91% to 85.84%) for RCM. For RCM, PPV was 42.4% (38.13% to 46.81%) and NPV was 99.1% (97.87% to 99.60%).

Conclusion
This study demonstrates that RCM can reliably diagnose MM and is fast enough to be integrated into UK pigmented lesion clinics by dermatologists trained in RCM. “Number needed to treat” dropped from 3.9 with clinical examination to 3.0 with dermoscopy to 1.3 with RCM.

Clinical Trial Registration: NCT03508297

Introduction

The incidence of Malignant Melanoma (MM) is increasing Worldwide in all countries with fair skinned populations, Australia excepted.[1–5] As a result, the cost of treating MM has become a significant financial burden to healthcare providers to fair-skinned populations worldwide, rising steadily in line with incidence.[6,7]

A recent systematic review examined 29 articles including a total of 398,549 biopsies/excisions. The overall number needed to treat (NNT) was 9.71 (95% CI, 7.72-12.29): 22.62 (12.95-40.10) for primary care, 9.60 (6.97-13.41) for dermatology, and 5.85 (4.24-8.27) for pigmented lesion specialists.[8] The heterogeneity in this data demonstrates that the setting, and the expertise of the clinician, can have a large influence on the NNT. In addition, the incidence of MM, and the average skin type, varies widely amongst clinical settings.

Reducing the NNT is important to reduce the economic burden of MM, especially as a number of studies have shown that compliance with follow up is poor and drops off with time.[9,10] The unnecessary removal of multiple benign lesions, combined with the pain/discomfort of multiple biopsies, may further increase this drop off rate.

Reflectance Confocal Microscopy (RCM) has been incorporated into the European Melanoma Guidelines as it "...increases diagnostic specificity in equivocal dermoscopic melanocytic lesions both in prospective studies, and in a recent meta-analysis conducted by the Cochrane Collaboration".[11] Currently 320 centres in the EU are using RCM for clinical dermatology, plus another 60 cosmetic centres, however only 4 centres in the UK use RCM clinically and RCM is not currently recommended for use in the UK to assess skin cancer.[12]

RCM has been shown to improve diagnostic specificity in equivocal dermoscopic melanocytic lesions in prospective studies.[13–17] A Cochrane meta-analysis demonstrated an improvement, concluding that "RCM may have a potential role in clinical practice, particularly for the assessment of lesions that are difficult to diagnose using visual inspection and dermoscopy alone, where the evidence suggests that RCM may be both more sensitive and specific in comparison to dermoscopy".[18] RCM is used to allow excision of clear-cut lesions, discharge of benign lesions, and further investigation of indeterminate lesions. A recent randomized clinical trial showed a higher predictive positive value and a lower number needed to excise compared with standard therapeutic care, demonstrating efficacy and safety in a prospective interventional setting.[19]

NICE diagnostics guidance in 2015 recommended further collection of data on the impact of RCM on the workflow of MM assessment.[20] There are few studies on diagnosis of MM in UK populations, but the European consensus-based interdisciplinary guidelines for melanoma recommend that "Confocal laser microscopy can be used for further evaluation of clinically/dermoscopic equivocal skin lesions".[11]

RCM has been shown to improve patient care internationally[21–25], applicability to the UK population remains unproven.[20] The UK has greater reliance on screening of skin disease

in primary care rather than by Dermatologist, therefore NICE recommended further UK research.[20]

This study assessed:

1. The diagnostic performance of RCM to diagnose MM in the UK population, using the histological assessment of the surgically excised lesion as the gold standard;

2. The diagnostic performance of dermoscopy alone for MM using histopathology as the gold standard;

Patients and methods

Design

The trial ran from March 2017 to August 2020 with patients recruited from a single private clinic setting (The Skin Care Network, Barnet, UK). The lack of UK centres precluded a multi-centre approach, so the study was designed so that results could be read in conjunction with international data.

Training

The principal investigator (HS) and the nurses assisting in the image acquisition attended an RCM teaching program at the University of Modena and Reggio Emilia, Modena, Italy, consisting of a 3-day basic training course on the use of RCM, followed by 3 days of practice and the controlled evaluation of 100 cases on an on-line platform.

Ethics and Governance

The protocol obtained Ethical Committee approval (Wales REC 7 17/WA/0044) and was registered on clinicaltrials.gov (NCT03508297).

All data was anonymised. Information regarding study patients was managed in accordance with the General Data Protection Regulation, Caldicott Guardian requirements, the Research Governance Framework for Health and Social Care, and had Research Ethics Committee approval.

Study data is stored at The Skin Care Network and within the Bluespier (Droitwich, Worcestershire, WR9 7ER) electronic patient record (EPR) system under normal arrangements for patient confidentiality. Only authorised members of the study team were given access to the study data

Sample size calculation

A pre-study evaluation of the power of the study was undertaken by Quantics Biostatistics (Edinburgh. EH3 8EG. Report Number: 0087, HS-MAV-002). Both specificity (the proportion of negative lesions correctly identified as negative) and sensitivity (the

proportion of positive lesions correctly identified as positive) will be estimated along with lower 95% confidence limits. The specificity is of primary concern and is the basis for the sample size calculation.

The sample size assumes that lesions within a patient are independent. The numbers of lesions required to provide a lower confidence limit for the specificity, which is no more than 3% lower than the estimate of specificity, was calculated for a range of assumptions about the true specificity, based on the meta-analysis referred to above (Table S1 included in supplementary material). However, the result of biopsy for each lesion was not known until after the lesion had been included in the study and the clinical, dermoscopic and RCM assessment had been undertaken. Therefore, lesions were added to the study until the required number of true negatives (true negatives plus false positives) of 654 had been recruited. The resulting true positives were then used to estimate the sensitivity as a secondary end point.

Study population

Patients 18 years or older with lesions suspected of MM or LM, or where that diagnosis was in the differential diagnosis before dermoscopy, were recruited prospectively and sequentially during the period March 2017 - August 2020. The patient cohort was predominantly referrals from GPs. Approximately a third of the patients were reviews following previous diagnoses of melanoma or non-melanoma skin cancer.

781 lesions were recruited. 48 lesions were subsequently excluded, 2 didn't undergo confocal imaging, 2 were missing dermoscopy, 18 were missing one or more scores, and 26 had no histology, 1 electively choosing to be referred back to the NHS where access to histology was not possible.

Study Workflow

The study protocol is shown in Supplementary Figure 1. Clinical, dermoscopic and RCM examinations were conducted sequentially by the principal investigator (HS), each technique adding to the diagnostic information available. The lesion was initially examined clinically and the diagnostic likelihood of MM recorded with a score of 1-3 (1 = possibly malignant, 2 = probably malignant, 3 = malignant). A score of zero was not possible as a differential diagnosis of MM was an inclusion criteria for the trial. A clinical photograph was obtained with an iPad 3 (8-megapixel camera, Apple, California, USA). Once a patient was included in the trial exclusion only occurred where it was not possible to obtain histology or RCM due to patient not attending for follow-up.

Next, dermoscopy was performed and a diagnosis rendered using the two-step algorithm method.[13,26,27] Examination was undertaken using both polarised and non-polarised immersion (using alcohol) contact dermoscopy with a handheld dermatoscope (DermLite DL4, 3Gen, San Juan Capistrano, CA, USA) with 20-fold magnification, and with an iPad 3 fitted with a 3Gen iPad adaptor. Each lesion was scored between 0-3 (0 = not malignant, 1 = possibly malignant, 2 = probably malignant, 3 = malignant) using clinical and dermoscopic information.

Finally, the lesion was imaged with RCM (VivaScope 1500 or 3000, Gen.3 and then Gen 4, VivaScope GmbH, Munich, Germany). The handheld 3000 system was used only rarely where access was an issue, for example on the nose. A minimum of three mosaics were obtained: at the superficial epidermis, dermo-epidermal junction (DEJ), and papillary dermis. Mosaic was the full size of lesion or maximum capture size available in the case of large lesions. The method has previously been described.[13,15,21,28–31] If that was not possible, two stacks of 4 blocks were taken at 30-micron intervals. The RCM images were taken during the running of a normal clinic or minor operation appointment with an additional 10 minutes being allocated for the RCM examination.

The principal investigator (HS) read the RCM images using the numerical score methodology outlined in Pellacani[28], and graded the likelihood of diagnosis of MM using the combination of clinical, dermoscopic and RCM information (0 = not malignant, 1 = possibly malignant, 2 = malignant)." Pellacani's method was abbreviated into a score of 0-2 for the purposes of this paper's analysis

When RCM imaging was complete, the lesions were either excised or a diagnostic biopsy was performed by the principal investigator (HS). The histopathological diagnosis was made by a dermatopathologist (JEJ) using conventional haematoxylin and eosin-stained sections and where necessary immunohistochemistry staining was utilised.

Statistical Analysis

Statistical analysis was performed using Microsoft Excel, MedCalc's online statistics calculator,[32] and easyROC.[33] Histopathological diagnosis was represented as non-MM = 0 and MM = 1.

Based on an outcome of MM or LM vs. other diagnosis, sensitivity, specificity, PPV, NPV, accuracy, and likelihood ratios were all calculated.

NNT, sometimes described as "number needed to diagnose" (NND) in a diagnostic study, is widely used to present the success of diagnostic tests.[34,35] For this trial, it would be defined as the average number of patients biopsied in order to find one MM: 1/(sensitivity + specificity – 1).[36,37] "Number needed to predict" (NNP), defined as 1/(PPV + NPV – 1), is dependent on prevalence so is a better descriptor of diagnostic tests in patient populations with different prevalence of disease.[38] For diagnostic tests, low values of NNT and NNP are desirable.

NNT is closely related to PPV, which represents the same underlying data as a proportion. PPV represents the proportion of biopsied lesions that are, in fact, MM. PPV = TP / (TP + FP).[35]

Results

All melanomas diagnosed during this period were recruited: 733 lesions from 597 patients. Median age was 59.3 years (range 18-99 years) and the male:female ratio was

1.95:1. 500 patients had 1 lesion, 73 had 2 lesions, 15 had 3 lesions, 7 had 4 lesions, 1 had 7 lesions, and 1 had 8 lesions.

648 lesions were true negatives. 654 true negatives predetermined the trial's end-point, but 48 lesions were excluded.

Histologically, the lesions were: 326 naevi, 102 seborrheic keratoses and solar lentigos, 93 BCC (mostly pigmented), 86 MM or LM (including 2 completely amelanotic MM, 33 LM and 43 SSMM), 8 SCC (2 in situ and 6 invasive), and 119 'other', all benign except 3 sarcomas. Mean and median sizes are listed in Table 1. 346 lesions were on the trunk, 151 on head or neck, 148 on the lower limbs, and 89 on upper limbs.

Mean MM Breslow thickness was 0.33 mm (SD 0.13 mm) in 40 invasive lesions, with 46 in situ MMs and LM.

Diagnostic performance

For any diagnostic test there is a balance to be struck between sensitivity and specificity. A point on the ROC curve needs to be chosen. Given the lethal nature of MM, and given that this is a group with lots of photodamage, we have chosen a cautious route, and where there is any significant risk of MM we will excise to protect the patient. The result is very high NPV but lower PPV, albeit much improved over clinical and dermoscopy.

Malignancy thresholds used were: Clinical:2, Dermoscopic:2, RCM:1. Table 2 and Figures 1-2 detail diagnostic performance. Table S3 and Figure 4 show performance at other threshold values.

Clinical alone, clinical plus dermoscopy, and clinical plus dermoscopy and RCM respectively shows sensitivity of 62.8%, 91.9% and 94.2%, specificity of 63.1%, 42.0% and 83.0%, PPV of 18.4%, 17.4% and 42.4%, NPV of 92.7%, 97.5% and 99.1%, accuracy of 63.1%, 47.9% and 84.3%. NNT was 3.86, 2.96 and 1.30, and NNP was 8.95, 6.73 and 2.41. Full figures including 95% CI are presented in Table 2 and 4 and Figures 2 and 3.

Each additional modality showed a significant increase in Area Under the Curve (AUC), see Table 2 and 3. Clinical examination alone showed an AUC of 0.65 (95% CI 0.60 - 0.71), adding dermoscopy 0.79 (0.74 - 0.84) and adding RCM 0.92 (0.89 - 0.95). Likelihood ratio, the likelihood of TP vs FP, went from 1.70 (1.41 – 2.06) to 1.58 (1.45 – 1.73) to 5.55 (4.64 – 6.63).

The benign to malignant ratio was 2.99:1. If the 36 BCC lesions that would have been excluded by dermoscopy are removed, the benign to malignant ratio is 3.53:1. The ratio of naevus to melanoma was 3.79:1. Supplementary data shows complete removal of BCCs.

Of the 5 false negatives shown in Figure 3, one in situ melanoma was misclassed as a benign naevus and two as benign solar lentigo. Two invasive MM were misclassed, one as a benign naevus and one as a seborrheic keratosis (Breslow thickness 0.3 mm and 1 mm respectively). 4 of the 5 false negatives reversed a true positive under dermoscopy, suggesting that some caution is merited in this situation.

Of the 110 false positives, 84 were naevi (48 compound, 7 intradermal, 25 junctional, and 4 dysplastic), 7 were pigmented AK, and 2 were BCC. Full results are in Table S4.

During the study we made the following additional observations.

- 10 minutes was sufficient for the physician (HS) to both acquire and read the images.
- Patients generally liked the procedure and found "provisional results reassuring"
- Appropriately trained nurses could quickly acquire RCM images using the VivaScope 1500. In this case, physician reading took 5 minutes.
- Image acquisition using the handheld VivaScope 3000 had to be undertaken by the reading physician as interpretation was a dynamic process.

Discussion

This is the first prospective observational study in the UK of MM and LM diagnosis with RCM in a UK secondary care setting. The study showed that it is possible to reliably diagnose MM. The majority of the lesions were small and or early (i.e. thin melanomas or in situ lesions), important as the experienced dermatologist does not need help to diagnose thick or late melanomas. Two amelanotic melanomas were detected that might otherwise have been missed or been treated inappropriately.

Detection of early and/or small and/or difficult to diagnose MMs is traditionally accompanied by a high value of NNT. This study demonstrates a reduction from 3.96 to 2.96 with dermoscopy only to 1.30 with the addition of RCM.

Of the 93 BCCs recruited, only three were given a primary diagnosis of MM under clinical examination. 36 would have been excluded by dermoscopy.

The false positives amongst the junctional compound melanocytic lesions were the result of investigator concern regarding architectural atypia, particularly at the DEJ, high degrees of pigmentary incontinence, or inflammation, with dendritic Langerhans cell in the epidermis being misread as pagetoid melanocytes.[39] In addition, the reader's initial partial experience in assessing the degree of cytological atypia resulted in an over-rating of mild or moderate degrees of dysplasia. As was expected, performance improved with greater experience reader (data not shown).[40]

The distinction between pigmented actinic keratosis, lichenoid keratosis and lentigo maligna was sometimes extremely difficult and had potential to cause diagnostic confusion. The ease of identifying obvious dendritic cells could lead to an overreliance on this feature, but it is important to carefully consider other supporting features to distinguish between these three lesions.

80% of False Negatives (n=4) contradicted a Dermoscopic True Positive diagnosis, suggesting that this situation should be treated with caution. In contrast, RCM correctly ruled out 281 Dermoscopy false positives.

The introduction of RCM in a mole diagnostic workflow seems to be effective and practical, but the possibility of delaying a MM diagnosis should be considered. In this paper in fact 5 MMs were missed, in line with previous data.[40] However all these lesions were thin or in situ except one case that resulted 1 mm thick.

The cost of the equipment is not insubstantial. However, one paper found that a reduction in diagnostic cost from €144 to €105 was achieved, which would lead to a cost saving of €262k per 1M inhabitants.[41] A recent UK study found that just diagnostic biopsy avoidance allowed an estimated cost saving of £18,480 over a period of three months in an NHS Dermatology clinic.[42]

Conclusion

RCM was found to be valuable in the diagnosis of MM in a UK population and can be safely used to exclude MM or malignancy.
Incorporation of RCM into patient screening prior to diagnosis not only speeds up confirmatory diagnosis, but also reduces NNT.
RCM can be incorporated into the workflow of an outpatient secondary care clinic with an acceptable learning curve for clinicians.

Acknowledgements

We would like to thank Mrs Donna Turner and Mrs Teena Skakum-Dring for their help with RCM imaging. While the microscope is owned by Skin Care Network, we gratefully acknowledge the technical support offered by the microscope manufacturer VivaScope GmbH (then Mavig GmbH), and we thank their UK associate Dr Gordon McKenzie of GBF Strategy Ltd for his assistance with this manuscript.

References

1 Olsen CM, Green AC, Pandeya N, Whiteman DC. Trends in melanoma incidence rates in eight susceptible populations to 2015. *J Invest Dermatol* 2018; **139**:1392–5.

2 Garbe C, Keim U, Eigentler TK, *et al*. Time trends in incidence and mortality of cutaneous melanoma in Germany. *J Eur Acad Dermatol Venereol* 2019; **33**:1272–80.

3 Glazer AM, Winkelmann RR, Farberg AS, Rigel DS. Analysis of trends in US melanoma incidence and mortality. *JAMA Dermatol* 2017; **153**:225–6.

4 Bianconi F, Crocetti E, Grisci C, *et al*. What has changed in the epidemiology of skin melanoma in central Italy during the past 20 years? *Melanoma Res* 2018.URL https://europepmc.org/article/med/30480621.

5 Pellacani G, Lo Scocco G, Vinceti M, *et al*. Melanoma epidemic across the millennium: time trends of cutaneous melanoma in Emilia-Romagna (Italy) from 1997 to 2004. *J Eur Acad Dermatol Venereol* 2008; **22**:213–8.

6 Morris S, Cox B, Bosanquet N. Cost of skin cancer in England. *Eur J Health Econ* 2009; **10**:267–73.

7 Vallejo-Torres L, Morris S, Kinge JM, *et al*. Measuring current and future cost of skin cancer in England. *J Public Health* 2014; **36**:140–8.

8 Petty AJ, Ackerson B, Garza R, *et al*. Meta-analysis of number needed to treat for diagnosis of melanoma by clinical setting. *J Am Acad Dermatol* 2020; **82**:1158–65.

9 Kittler H, Weitzdorfer R, Pehamberger H, *et al*. Compliance with follow-up and prognosis among patients with thin melanomas. *Eur J Cancer* 2001; **37**:1504–9.

10 Kalimullah FA, Brown CW. Compliance with follow-up among patients with melanoma and non-melanoma skin cancers. *Dermatol Online J* 2014; **20**.URL https://www.ncbi.nlm.nih.gov/pubmed/24612570.

11 Garbe C, Amaral T, Peris K, *et al*. European consensus-based interdisciplinary guideline for melanoma. Part 1: Diagnostics - Update 2019. *Eur J Cancer* 2020; **126**:141–58.

12 NICE. Melanoma: assessment and management. NICE Guideline [NG14]. , NICE, 2015URL https://www.nice.org.uk/guidance/ng14 [accessed on 25 August 2020].

13 Guitera P, Menzies SW, Longo C, *et al*. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. *J Invest Dermatol* 2012; **132**:2386–94.

14 Alarcon I, Carrera C, Palou J, *et al*. Impact of in vivo reflectance confocal microscopy on the number needed to treat melanoma in doubtful lesions. *Br J Dermatol* 2014; **170**:802–8.

15 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. *Br J Dermatol* 2014; **171**:1044–51.

16 Borsari S, Pampena R, Lallas A, *et al*. Clinical Indications for Use of Reflectance Confocal Microscopy for Skin Cancer Diagnosis. *JAMA Dermatol* 2016; **152**:1093–8.

17 Cinotti E, Fiorani D, Labeille B, *et al*. The integration of dermoscopy and reflectance confocal microscopy improves the diagnosis of lentigo maligna. *J Eur Acad Dermatol Venereol* 2019; **33**:e372–4.

18 Dinnes J, Deeks JJ, Saleh D, *et al*. Reflectance confocal microscopy for diagnosing cutaneous melanoma in adults. *Cochrane Database Syst Rev* 2018. doi:10.1002/14651858.CD013190.

19 Pellacani G, Farnetani F, Ciardo S, *et al*. Effect of reflectance confocal microscopy for suspect lesions on diagnostic accuracy in melanoma: a randomized clinical trial. *JAMA Dermatol* 2022. doi:10.1001/jamadermatol.2022.1570.

20 NICE. VivaScope 1500 and 3000 imaging systems for detecting skin cancer lesions, NICE Diagnostic Guidance [DG19]. , NICE, 2015URL https://www.nice.org.uk/guidance/dg19/.

21 Pellacani G, Scope A, Gonzalez S, *et al*. Reflectance Confocal Microscopy made easy: the four must-know key features for the diagnosis of melanoma and non-melanoma skin cancers. *J Am Acad Dermatol* 2019. doi:10.1016/j.jaad.2019.03.085.

22 Ahlgrimm-Siess V, Weitzer F, Arzberger E, *et al*. Diagnostic impact of reflectance confocal microscopy as a second-level examination for facial skin lesions. *J Dtsch Dermatol Ges* 2019; **17**:266–73.

23 Kadouch DJ, Leeflang MM, Elshot YS, *et al*. Diagnostic accuracy of confocal microscopy imaging versus punch biopsy for diagnosing and subtyping basal cell carcinoma. *J Eur Acad Dermatol Venereol* 2017; **31**:1641–8.

24 Longo C, Borsari S, Pampena R, *et al*. Basal cell carcinoma: the utility of in vivo and ex vivo confocal microscopy. *J Eur Acad Dermatol Venereol* 2018; **32**:2090–6.

25 Stefanski M, Le Guern A, Visseaux L, *et al*. Real-life practice of reflectance confocal microscopy in France: A prospective multicenter study. *J Am Acad Dermatol* 2024. doi:10.1016/j.jaad.2024.01.079.

26 Argenziano G, Soyer HP, Chimenti S, *et al*. Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet. *J Am Acad Dermatol* 2003; **48**:679–93.

27 Marghoob A. Two-step algorithm [WWW Document]. dermoscopedia. 2019.URL https://dermoscopedia.org/w/index.php?title=Two-step_algorithm&oldid=15900 [accessed on 7 July 2023].

28 Pellacani G, Cesinaro AM, Seidenari S. Reflectance-mode confocal microscopy of pigmented skin lesions--improvement in melanoma diagnostic specificity. *J Am Acad Dermatol* 2005; **53**:979–85.

29 Shahriari N, Grant-Kels JM, Rabinovitz H, *et al*. Reflectance confocal microscopy features of melanomas on the body and non-glabrous chronically sun-damaged skin. *J Cutan Pathol* 2018; **45**:754–9.

30 Rajadhyaksha M, Grossman M, Esterowitz D, *et al*. In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. *J Invest Dermatol* 1995; **104**:946–52.

31 Rajadhyaksha M, González S, Zavislan JM, *et al*. In vivo confocal scanning laser microscopy of human skin II: advances in instrumentation and comparison with histology. *J Invest Dermatol* 1999; **113**:293–303.

32 MedCalc Software. MedCalc's Diagnostic test evaluation calculator [WWW Document]. 2020.URL https://www.medcalc.org/calc/diagnostic_test.php [accessed on 8 September 2020].

33 Goksuluk D, Korkmaz S, Zararsiz G, Karaagaoglu AE. easyROC: an interactive web-tool for ROC curve analysis using R language environment. *R J* 2016; **8**:213.

34 Bender R. Calculating confidence intervals for the number needed to treat. *Control Clin Trials* 2001; **22**:102–10.

35 Marchetti MA, Yu A, Nanda J, *et al*. Number needed to biopsy ratio and diagnostic accuracy for melanoma detection. *J Am Acad Dermatol* 2020; **83**:780–7.

36 Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; **3**:32–5.

37 Larner AJ. Number needed to diagnose, predict, or misdiagnose: useful metrics for non-canonical signs of cognitive status? *Dement Geriatr Cogn Dis Extra* 2018; **8**:321–7.

38 Linn S, Grunau PD. New patient-oriented summary measure of net total gain in certainty for dichotomous diagnostic tests. *Epidemiol Perspect Innov* 2006; **3**:11.

39 Hashemi P, Pulitzer MP, Scope A, *et al*. Langerhans cells and melanocytes share similar morphologic features under in vivo reflectance confocal microscopy: a challenge for melanoma diagnosis. *J Am Acad Dermatol* 2012; **66**:452–62.

40 Fraga-Braghiroli N, Grant-Kels JM, Oliviero M, *et al*. The role of reflectance confocal microscopy in differentiating melanoma in situ from dysplastic nevi with severe atypia: A cross-sectional study. *J Am Acad Dermatol* 2020; **83**:1035–43.

41 Pellacani G, Witkowski A, Cesinaro AM, *et al*. Cost–benefit of reflectance confocal microscopy in the diagnostic performance of melanoma. *J Eur Acad Dermatol Venereol* 2015. doi:10.1111/jdv.13408.

42 Baloch H, Nicholls I, Garioch J, Fadhil M. Reflectance Confocal Microscopy and Basal Cell Carcinoma: Is it time to embrace the new technology? In: *Annual Meeting of British Association of Dermatologists*. , 2024.

43 Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934.URL https://www.jstor.org/stable/2331986.

44 Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007; **26**:2170–83.

Figure legends

Figure 1: Diagnostic Accuracy

Diagnostic statistics for each modality. Error bars show 95% confidence limits for sensitivity, specificity, and accuracy[43] and predictive values.[44]

Figure 2: ROC Curve
Receiver Operator Curves for each modality. ROC plot produced using easyROC [33].

Table 1: Lesion sizes against the final histological diagnosis

Diagnosis	Count of Diagnosis	Mean Diameter (mm)	Median diameter (mm)
Atypical	50	7.6	6.7
Epithelial	43	8.0	7.0
AK	19	9.5	10.0
PAK	24	6.9	6.7
Melanocytic	7	5.1	4.0
ALP	3	6.9	5.5
AMH	1	3.0	3.0
Atypical lent	1	2.8	2.8
MELTUMP	2	4.5	4.5
Benign	501	5.3	4.5
Melanocytic	338	4.4	4.0
CN	170	4.7	4.5
IDN	47	4.0	4.0
JN	93	3.8	3.0
Other	28	5.1	5.2
Non-Melanocytic	147	7.5	6.3
LK	21	7.3	7.0
Other	16	9.2	5.2
PIH	8	9.0	8.7
SK	42	6.4	5.5
SL	60	7.7	6.0
Soft tissue	16	5.0	4.0
ALHWE	1	6.0	6.0
DF	6	5.2	4.5
Haemangiom	7	4.9	3.0
Neurofibrom	1	5.0	5.0
Schwannom	1	4.0	4.0
Malignant	179	7.8	7.0
Epithelial	100	7.5	7.0
BCC	92	6.9	6.9
Bowen's dise	5	17.7	20.0
SCC	3	9.3	9.0
Melanocytic	77	8.0	4.0
in situ	37	7.7	7.0
invasive	40	8.3	7.2
Soft tissue	2	13.7	4.0
Leiomyosarc	1	24.5	24.5
Pleomorphic	1	3.0	3.0
Post Excision	3	5.0	5.0
Scar	3	5.0	5.0
Grand Total	733	6.1	5.0

Mean and median lesion sizes per diagnosis.

Table 2: Diagnostic Results

	Clinical			Dermoscopic			RCM	
Statistic	**Value**	**95% CI**		**Value**	**95% CI**		**Value**	**95% CI**
Sensitivity	62.79%	51.70% to 72.98%		91.86%	83.95% to 96.66%		94.19%	86.95% to 98.09%
Specificity	63.12%	59.27% to 66.84%		41.98%	38.14% to 45.88%		83.02%	79.91% to 85.84%
Positive Predictive Value	18.43%	15.73% to 21.48%		17.36%	16.10% to 18.70%		42.41%	38.13% to 46.81%
Negative Predictive Value	92.74%	90.61% to 94.42%		97.49%	95.00% to 98.76%		99.08%	97.87% to 99.60%
Accuracy	63.08%	59.47% to 66.58%		47.82%	44.15% to 51.50%		84.33%	81.50% to 86.89%
Positive Likelihood Ratio	1.70	1.41 to 2.06		1.58	1.45 to 1.73		5.55	4.64 to 6.63
Negative Likelihood Ratio	0.59	0.45 to 0.78		0.19	0.09 to 0.40		0.07	0.03 to 0.16
Disease prevalence	11.7%	9.49% to 14.29%		11.7%	9.48% to 14.27%		11.7%	9.48% to 14.27%
Number Needed to Treat	3.86	9.12 to 2.51		2.96	4.53 to 2.53		1.30	1.50 to 1.19
Number needed to Predict	8.95	15.77 to 6.29		6.73	9.01 to 5.73		2.41	2.78 to 2.15

Diagnostic statistics for each modality. Cutoffs to be considered malignant were: Clinical:2; Dermoscopic:2: RCM:1.

Table 3. Counts of diagnostic likelihood score by stage

	Clinical	Dermoscopy	RCM
0	0	3	542
1	440	277	66
2	247	372	125
3	46	81	

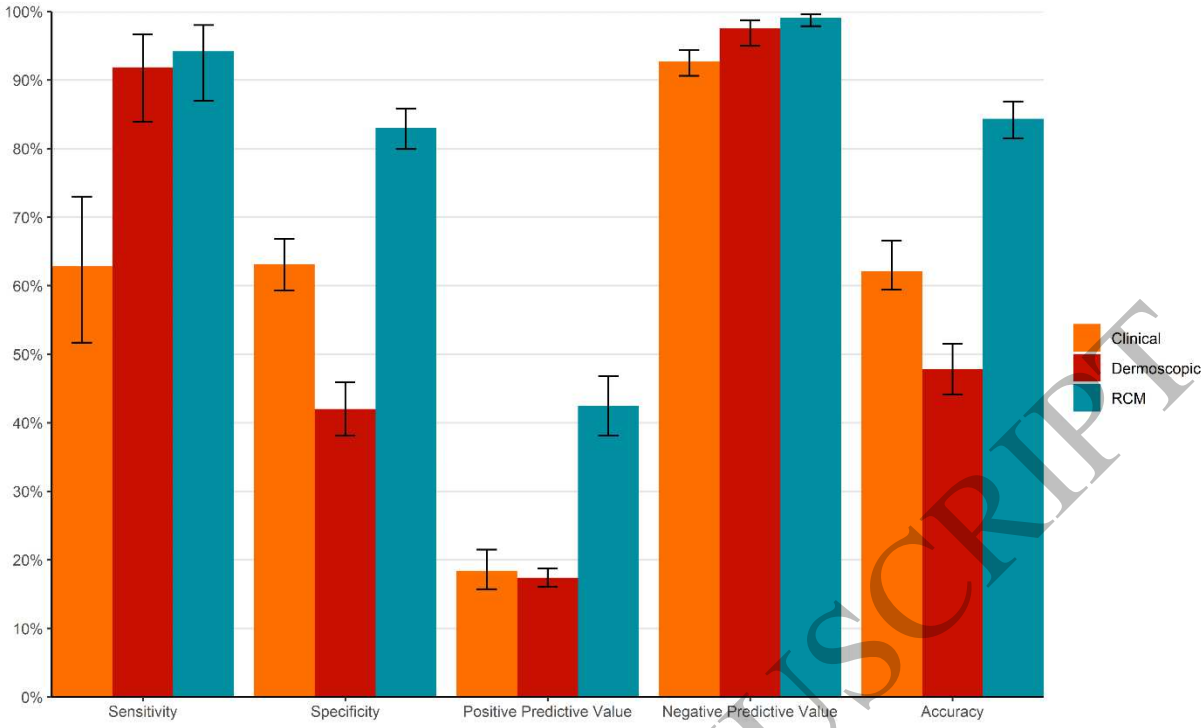

Figure 1
165x99 mm (x DPI)

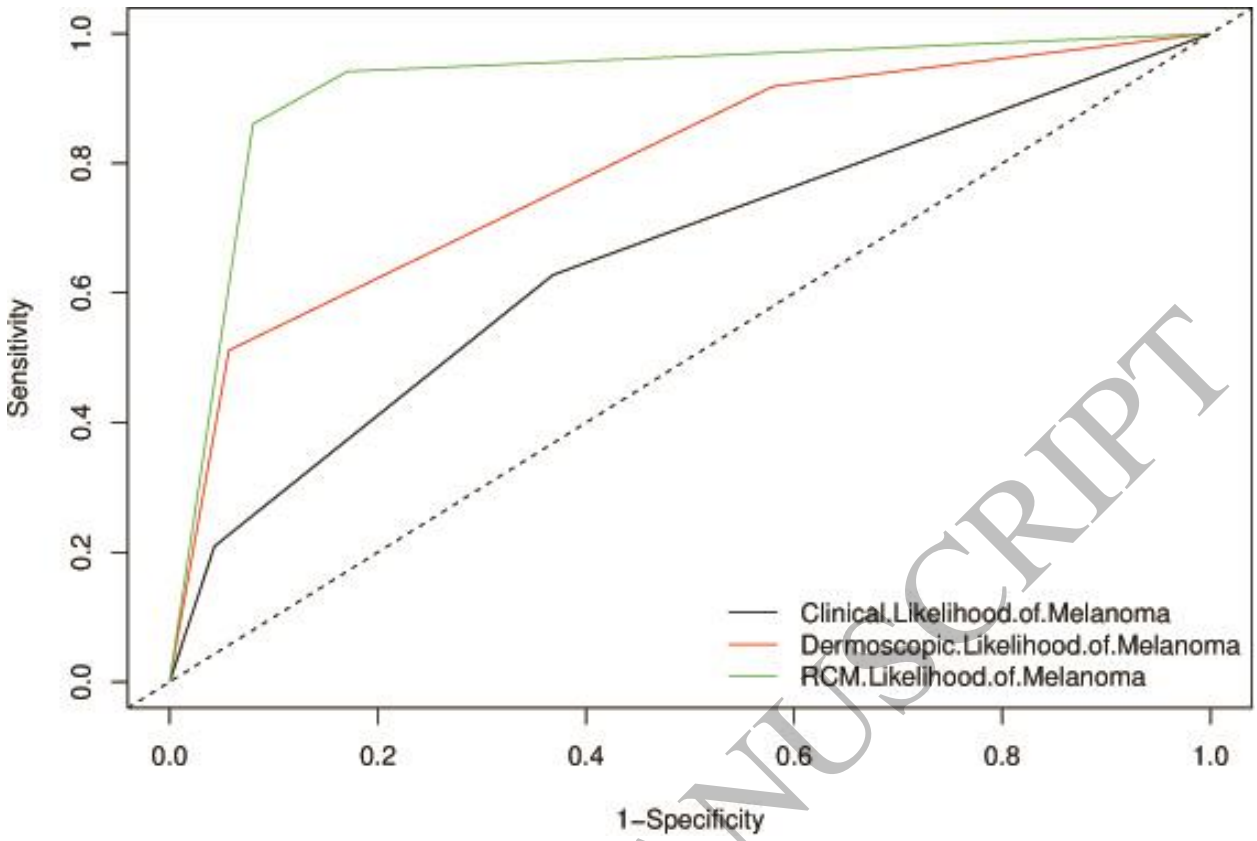

Figure 2
165x109 mm (x DPI)

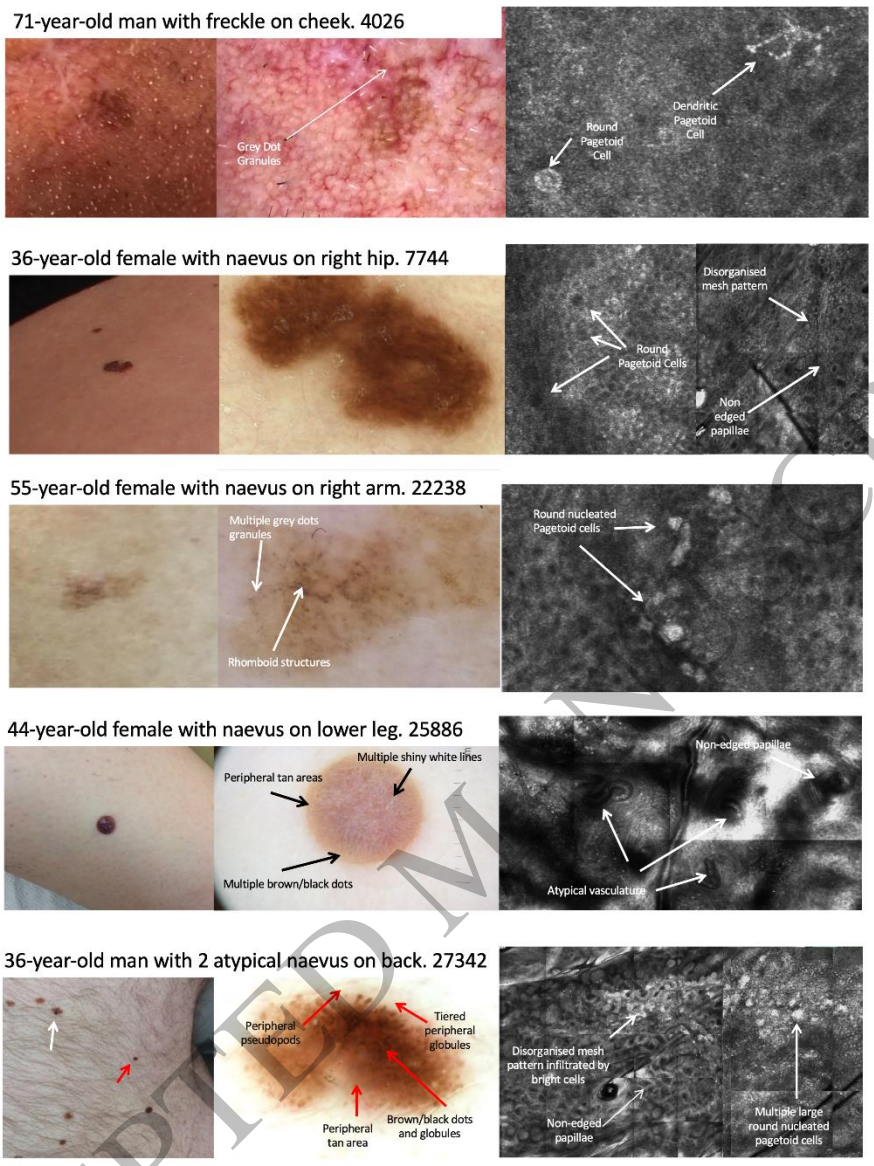

Figure 3
165x233 mm (x DPI)

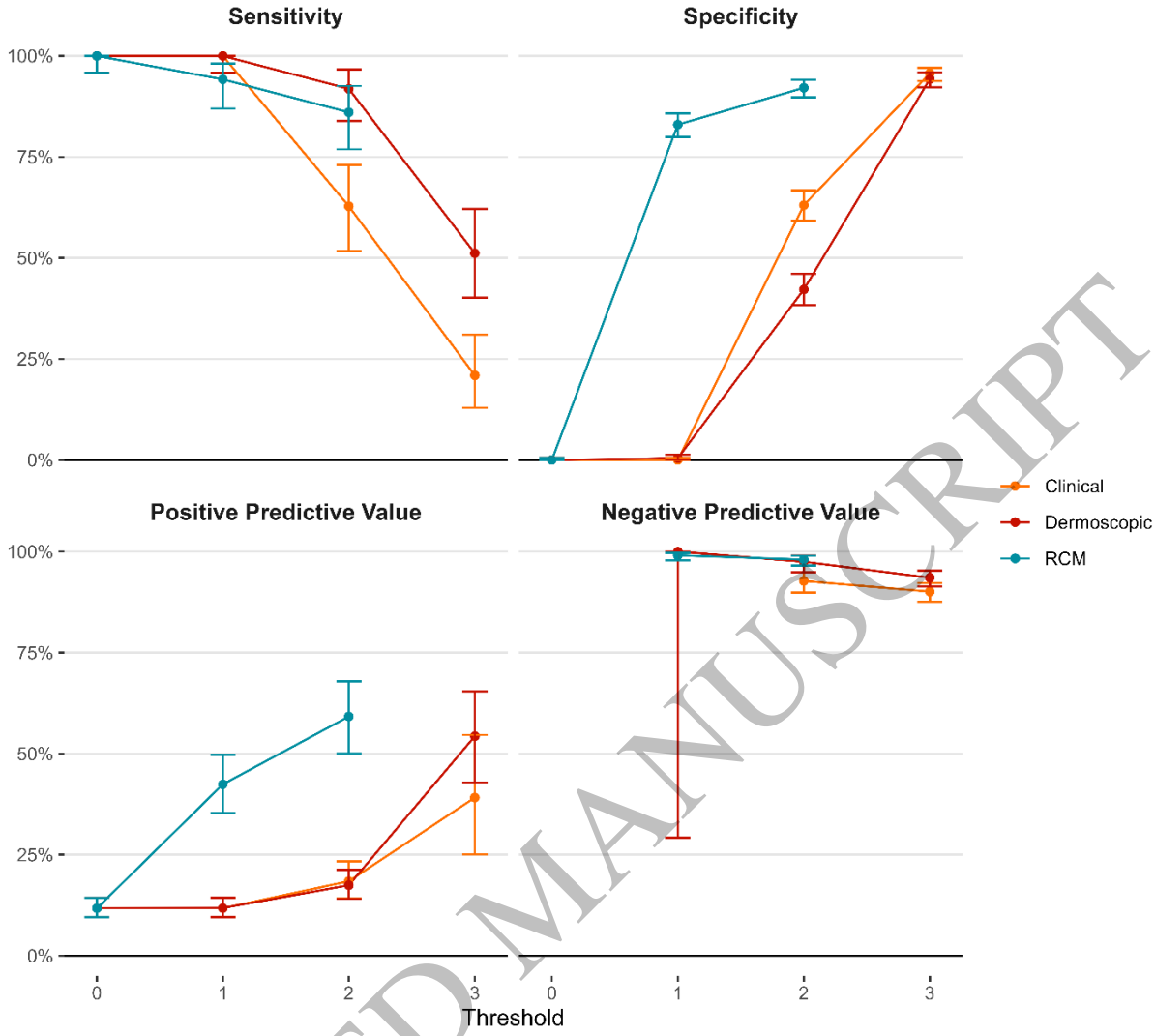

Figure 4
165x144 mm (x DPI)