Diplomarbeit

Reflectance Confocal Microscopy - A helpful tool for diagnosing benign and malignant facial lesions?

eingereicht von

Friedrich Weitzer

Geboren am 25.11.1988

zur Erlangung des akademischen Grades

Doktor der gesamten Heilkunde

(Dr. med. univ.)

an der

Medizinischen Universität Graz

ausgeführt an der

Univ. Klinik für Dermatologie und Venerologie

unter der Anleitung von

Univ. Prof. Dr. med. univ. Rainer Hofmann-Wellenhof und Priv.Doz. Dr. med. univ. Verena Ahlgrimm-Siess

Graz, am…

Unterschrift:

Eidesstattliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig und ohne andere Hilfe als die der Betreuer verfasst habe andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am… Unterschrift:

Index Page no.

Summary:

English:

Introduction: Clinical differentiation between benign and malignant lesions of the skin may be difficult even for experienced dermatologists. In clinical equivocal cases a biopsy is often required to rule out malignancy. Facial lesions are an even greater diagnostic challenge because of site specific clinical features which reflect the unique histomorphology of facial skin and the limitations for biopsy in this cosmetically and functionally sensitive area. Non-invasive measures to improve the detection rate of malignant skin lesions include the use of dermoscopy, computer image analysis and in vivo reflectance confocal laser microscopy (RCM), as a promising new imaging technique. There are only few RCM- studies in the literature which included equivocal facial skin lesions to demonstrate the benefit of using this technology in this specific body area.

Objective: The aim of the present retrospective study was to demonstrate the additional ability of RCM to differentiate between benign and malignant facial skin lesions. Furthermore, considering the unique histomorphology of facial skin, we investigated the difference between a novice observer and two expert observers using this new technology.

Patients and methods: 160 facial skin lesions were imaged in 148 patients. Patients were aged from 5 to 93 years with a mean age of 65.1 years. 42 patients (28%) were male and 106 (72%) were female. 12 lesions were labial lesions. Images were evaluated two times by a single novice observer and one time consensually by two expert observers.

Results: The first course of image evaluation done by the single novice observer, showed a low sensitivity (43%), but a high specificity (81%) for detection of malignant skin lesions when assessing clinical, dermoscopic and RCM images together. Sensitivity and specificity for evaluation of clinical and dermoscopic images alone were 65% and 79%, respectively.

In the second course of image evaluation done by the single novice observer, a good sensitivity (71%) and specificity (87%) for detection of malignant skin lesions

was obtained for evaluation of all images together. Sensitivity and specificity for evaluation of clinical and dermoscopic images was 68% and 71%, respectively.

Image evaluation done consensually by the two expert observers showed a good sensitivity (80%) and specificity (84%) for detection of malignant skin lesions when assessing clinical, dermoscopic and RCM images together and a sensitivity and specificity of 71% and 82%, respectively, for assessment of clinical and dermoscopic images alone.

Conclusion: Our study clearly shows the advantage of using RCM as adjunct to clinical and dermoscopic examination over clinical and dermoscopic examination alone. A second course of image evaluation done by the single novice observer showed a better sensitivity and specificity than for the first course, clearly representing a short learning curve. However, the two experienced observers still showed better sensitivities and specificities compared to the single novice observer.

Deutsch:

Zusammenfassung: Die klinische Unterscheidung zwischen benignen und malignen Hautläsionen ist auch für erfahrene Dermatologen schwierig. In klinisch suspekten Fällen ist oft eine Biopsie erforderlich um Malignität auszuschließen. Gesichtsläsionen stellen eine noch größere diagnostische Herausforderung dar; einerseits wegen der ortsspezifischen klinischen und auflichtmikroskopischen Kriterien, welche die einzigartige Histomorphologie der Gesichtshaut widerspiegeln und andererseits wegen der strengen Indikationsstellung für Biopsien in diesem kosmetisch und funktionell empfindlichen Gebiet. Nicht invasive Methoden um die Detektionsrate von malignen Hautläsionen zu verbessern schließen den Gebrauch von Dermatoskopie, computer-gestützter Bildanalyse und, als viel versprechende neue bildgebende Methode, die in vivo konfokale Lasermikroskopie (RCM) ein. In der Literatur sind nur wenige RCM-Studien bekannt, welche suspekte Gesichtsläsionen beinhalten, um den Nutzen dieser Technologie in dieser spezifischen Körperregion zu zeigen.

Ziel: Das Ziel der gegenwärtigen retrospektiven Studie war es, den Benefit der RCM zur Unterscheidung benigner und maligner Gesichtsläsionen zu zeigen.

Weiters untersuchten wir unter Beachtung der speziellen Histomorphologie der Gesichtshaut den Unterschied zwischen einem unerfahrenen Beobachter und zwei Experten bei Nutzung dieser neuen Technologie.

Ergebnisse: 160 Gesichtsläsionen von insgesamt 148 Patienten wurden untersucht. Das Alter der Patienten reichte von 5 bis 93 Jahren mit einem mittleren Alter von 65,1 Jahren. 42 Patienten (28%) waren männlich und 106 (72%) waren weiblich. 12 Läsionen waren labiale Läsionen. Die Bilder wurden zweimal von einem einzelnen unerfahrenen Beobachter und einmal konsensuell von zwei Experten untersucht.

Ergebnisse: Die erste Untersuchung der Bilder durch den unerfahrenen Beobachter zeigte eine niedrige Sensitivität (43%) aber eine hohe Spezifität (81%) für die Erkennung maligner Gesichtsläsionen unter Verwendung klinischer, Auflicht- und RCM Bilder zusammen. Die Sensitivität und die Spezifität für die Verwendung klinischer und Auflichtbilder alleine waren 65 % bzw. 79%.

Die zweite Untersuchung der Bilder durch den unerfahrenen Beobachter zeigte eine gute Sensitivität 71 % und Spezifität 87 % für die Erkennung maligner Hautläsionen wurde durch die Verwendung aller Bildtypen zusammen erreicht. Die Sensitivität und die Spezifität für die Verwendung klinischer und Auflichtbilder alleine waren 68 % bzw. 71%.

Die konsensuelle Bilduntersuchung durch die zwei Experten zeigte eine gute Sensitivität (80 %) und Spezifität (84 %) für die Erkennung maligner Gesichtsläsionen unter Verwendung klinischer, Auflicht- und RCM Bilder zusammen. Die Sensitivität und die Spezifität für die Verwendung klinischer und Auflichtbilder alleine waren 71 % bzw. 82 %.

Schlussfolgerung: Unsere Studie zeigt klar die Vorteile der RCM Methode in Verbindung mit klinischen und Auflichtbildern über die Untersuchung klinischer und Auflichtbilder alleine. Eine zweite Untersuchung der Bilder durch den unerfahrenen Beobachter zeigte eine bessere Sensitivität und eine gleiche Spezifität, was klar eine kurze Lernkurve darstellt. Darüber hinaus zeigten zwei Experten verglichen mit einem unerfahrenen Beobachter eine bessere Sensitivität und Spezifität.

Introduction:

The clinical diagnosis of malignant facial skin lesions, such as basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and lentigo maligna (LM) is particularly challenging because of their tendency to occur in a background of significant photodamage. Associated benign skin neoplasms (e.g. solar lentigines, seborrheic keratoses) may clinically and dermoscopically mimic malignant skin lesions and vice versa; in addition, collisions between benign and malignant skin neoplasms have been described in the literature. While BCC and SCC tend to show only local destructive growth and usually do not metastasize, early excision of LM with adequate margins is crucial since delayed recognition puts the patient at risk for death of disease once the tumor had progressed and is capable of metastasis.

Lentigo maligna (LM) is a variant of malignant melanoma which affects especially elderly patients with chronically sun-damaged skin. LM usually appears as a widespread light to dark brown macule on clinical examination. Elevation and color variegation is rare at an early stage of disease and clinical differentiation from benign skin lesions like solar lentigo (SL) and seborrheic keratosis (SK) or from pigmented actinic keratosis (in situ SCC) may be impossible. Additionally, the dermoscopic features of facial pigmented skin lesions vary greatly from those at other sites because the dermo-epidermal junction (DEJ) of sun-damaged facial skin is flattened and usually does not produce a classic pigment network, which is often relied upon to distinguish melanocytic from non-melanocytic lesions.

Reflectance confocal microscopy:

Reflectance confocal microscopy enables non-invasive imaging of the skin at cellular level resolution. Skin cytomorphology can be visualized in vivo at the bedside regardless of the lesion location or the patient´s general health condition. Like dermoscopy, RCM acquires images in the horizontal (en face) plane and allows direct correlation of RCM features with dermoscopic structures; the correlation with histopathologic features is, however, limited due to the different planes of imaging. Besides the in vivo application, this new imaging technology can also be used for ex vivo imaging of freshly excised tissue.

Basic principles of RCM:

In 1995 Rajadhyaksha et al (1) presented an entirely newly designed and constructed in vivo reflectance confocal microscope (originally called confocal scanning laser microscope; CSLM) which was built for in vivo real-time imaging of the epidermis and upper dermis at up to 100x resolution. The laser wavelength was the major depending factor of the imaging depth. A wavelength of 400-700nm was usable for scanning the epidermis and a wavelength of 800-900nm for the deeper dermis. Live imaging with this prototype also displayed the capillary blood flow in the papillary dermis.

The currently commercially available RCM Vivascope 1500® model (See figure no 1) is a slightly improved version of the original one described by Rajadhyaksha et al (2) in 1999. It uses a near-infrared, low-power laser beam (830nm diode laser, up to 35mW) as light source; the imaging depth is limited to approximately 250 µm, which correlates to the level of the papillary (body) or the reticular (face) dermis. The laser beam is focused to a certain point of interest in the tissue that should be scanned and is reflected with variable intensity by the specific skin components; melanin, keratin and collagen are highly reflective skin structures. Only the light reflected back from the focus enters the detector through a pinhole and is processed by dedicated software. By scanning the laser beam in a two-dimensional grid over the skin, a thin horizontal optical section is obtained at a preselected imaging depth and displayed as a gray scale image. The real-time imaging enables also visualization of dynamic skin processes, such as the blood flow through dermal capillaries. The laser wavelength used is the major limiting factor for imaging depth and resolution. A wavelength of 400-700nm provides optimal resolution for scanning the epidermis, but has a limited penetration depth. Wavelengths of 800nm and higher penetrate deeper into the skin, but provide lower resolution.

During live imaging a 500x500 µm grey scale image (“single image”) is displayed on the computer screen. Mosaic images (“Vivablock”) composed by up to 16x16 contiguous images may be obtained in the horizontal plane (xy-axis) by an automated stepper (8x8mm field-of-view).The imaging depth may be preselected by adjusting the focal laser length. A series of single images at various imaging

depths at the same point in the tissue (z-axis) may be automatically stacked in areas of special interest (“Vivastack”, “optical biopsy”). The lateral resolution provided is 0.5 – 1 µm and the axial resolution (thickness of the horizontal optical sections) is 3 – 5 µm, comparable to the thickness of conventional histopathology sections. However, the RCM assessment of nuclear details is inferior to histopathology and there is currently no staining available to improve the contrast of RCM images.

RCM imaging:

For RCM-imaging, a metal adaptor ring provided with a disposable polymer window is attached to the patient’s skin using a double-sided adhesive and coupled magnetically to the RCM probe to minimize motion artifacts. Water based gel and crodamol oil serve as immersion media between window and RCM probe and skin and polymer window, respectively (See figure no 2). A dermoscopic picture taken through the adaptor ring before starting RCM imaging is displayed simultaneously on the computer screen; it correlates precisely to the RCM mosaic images and serves as a gross map to guide RCM-imaging of sub-regions of the lesion.

Figure 1: Vivascope ® 1500 reflectance confocal laser microscope: the microscope probe is mounted on an arm that pivots through all axes. The lighter blue device in the middle is the so called Vivacam ® (dermoscopic camera). The smaller dark blue device on the left side of the Vivascope® is a new hand-held device (Vivascope ® 3000) especially designed for difficult to access areas (as shown in the picture on the right).

Figure 2: A metal adaptor ring with a polymer window for RCM imaging is attached to the patient's skin using a double sided adhesive.

RCM imaging of normal skin:

There are different RCM- patterns observed at various skin levels according to the different skin structures. The normal skin surface appears as bright, wrinkled sheets of stratum corneum. The dark “valleys” between the sheets correlate to the skin folds. Occasionally large polygonal cells without visible nuclei are seen within the sheets; these cells are single corneocytes. The granular and the spinous layers usually appear as so called “honeycomb pattern” in RCM, according to the microscopic skin architecture at this imaging depth. The nuclei of the keratinocytes appear dark and form the “holes” of the “honeycomb pattern”, while cytoplasm, outer membranes and intercellular junctions appear bright and correspond to the “network” of this pattern. In healthy epidermis both structures are uniform in shape and size. (See figure no. 3)

The so-called “cobblestone pattern” is a variant of the regular skin pattern observed at the granular-spinous layer and is usually found in darker skin types and in tanned skin. The “cobblestone pattern” is formed by pigmented, bright round keratinocytes aggregated back-to-back whose nuclei are obscured by melanin. In addition darker skin types and sun-exposed skin may also display a brighter stratum corneum and a higher amount of basal cells per millimeter. (3)

The basal layer is visualized at an imaging depth of approximately 100 µm in RCM. Basal cells appear brighter than other keratinocytes due to a higher amount of intercellular melanin. Supranuclear caps of melanin in basal keratinocytes are easily visualized but unlike in animals, dendritic melanocytes are usually not observed because human melanocytes transfer the melanin rapidly to the surrounding cells and do not accumulate it. UV induced melanin production can be easily displayed in a weekly RCM follow-up after sun exposure. (4)

During RCM imaging the healthy, undulating dermo-epidermal junction appears first as disk-like aggregates of bright cells corresponding to the basal keratinocytes on top of dermal papillae. Going deeper, dark round dermal papillae surrounded by a rim of bright basal keratinocytes (“edged papillae”) are seen.

The blood flow within papillary blood vessels may be additionally observed as mentioned previously.

In a flattened dermo-epidermal junction (e.g. as seen in chronically sun-damaged skin) the basal layer is visualized as a sheet of bright basal keratinocytes cells above the dermis and without appearance of bright rings indicating the presence of dermal papillae.

Within the upper dermis, bright fibrillar structures forming a web-like pattern are seen correlating to dermal collagen of the papillary dermis. Thicker bundles are observed at deeper dermal levels according to stronger collagen bundles in the reticular dermis. Appendageal skin structures like sweat glands or hair follicles are additionally displayed in RCM. Sweat gland openings form bright centrally hollow structures which are spiraling through the entire epidermis if followed at serial images of different imaging depths. Hair follicles are visualized as round to oval structures with smaller keratinocytes at the periphery and larger flat ones next to the center, which is formed by a dark ostium that contains the hair shaft. (See figure no. 3)

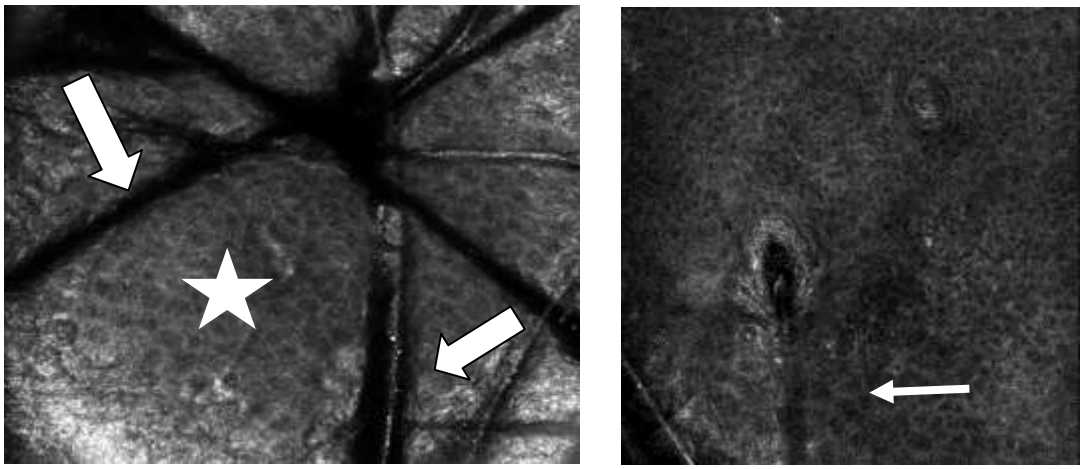

Figure 3: Single RCM images of the upper epidermis. (500x500 µm field of view) Left: The stratum corneum appears as bright reflecting wrinkled sheets (asterisk). The dark "valleys" between the sheets correspond to the skin folds (dermatoglyphs; thick arrows). Right: Typical "honeycomb pattern" at high resolution with a hair follicle in the image center characterized by a dark ostium containing a hair shaft (arrow).

RCM of non-melanoma skin cancer:

Non-melanoma skin cancers are usually found in elderly patients with chronically sun-damaged skin. The patient's skin often presents multiple basal cell carcinomas (BCC), actinic keratoses (AK, synonymous: solar keratosis, *in situ* SCC) and squamous cell carcinomas (SCC) in sun-exposed areas which reflects the causative relationship between chronic sun exposure and skin cancerogenesis ("field cancerization"). For dermatologists it is often tricky to distinguish these malignant lesions from benign lesions which are also frequently found in elderly patients (e.g. seborrheic keratoses or solar lentigines). Dermoscopy can improve diagnostic accuracy; however it is of limited value in some cases as mentioned previously. In addition, malignant facial skin neoplasms often lack pigmentation ("pink lesions") and may be therefore difficult to differentiate from other benign skin neoplasms or inflammatory skin conditions. RCM has been described as a helpful tool in the management of these clinically difficult to diagnose pink lesions. (5)

Actinic keratosis and squamous cell carcinoma:

Actinic keratoses are *in situ* squamous cell carcinomas of the skin which present clinically as red (non-pigmented variant) to brown (pigmented variant) scaly patches on sun-exposed areas. (see figure no. 4) Approximately 10% of AKs evolve into an invasive SCC. On clinical examination, SCCs are usually larger and thicker than AK reflecting their infiltrative growth. Dermoscopy of non-pigmented AK and early invasive SCC shows dotted or glomerular vessels and scaling; pigmented AK and SCC may display network-like and regression structures. The diagnosis is usually based on clinical and dermoscopic findings of an individual lesion and the presence of multiple similar lesions in chronically sun-damaged skin. Single lesions and so called "ugly ducklings" may be difficult to diagnose on clinical grounds alone and may require a biopsy to prove diagnosis and rule out other skin tumors or inflammatory skin conditions (6). With RCM, AK and SCC show cytomorphologic and architectural irregularity of the epidermis such as seen in conventional histopathology. Although similar, the morphologic RCM changes of AK are usually less pronounced and more focal than in SCC. The skin surface shows highly reflective irregularly shaped structureless skin islands with focal dark nuclei. These RCM findings correlate to the ortho- and parakeratosis described in

histopathology. The granular-spinous layers show an irregular honeycomb or cobblestone pattern with multiple atypical nuclei. (see figure no. 5 and 6) The upper dermis is characterized by clumped bundles of bright collagen and dilated round blood vessels running perpendicular to the plane of imaging corresponding to the solar elastosis observed in histopathology and the red dots (vessels) seen in dermoscopy. SCCs may even show a focal or total loss of the cobblestone or honeycomb pattern as well as single bright round cells within the granular-spinous layers, whose nuclei are often obscured. These cells correlate to dyskeratotic cells seen on histopathology. RCM has been shown to be useful in diagnosing SCC and AK, but it cannot be used to measure the invasion depth of SCC due to the limited imaging depth. Ulrich et al (7) described a disarranged pattern (irregularity or loss of the honeycomb/cobblestone pattern) and cellular pleomorphism (atypia of nuclei) as the best predictors for AK with a high sensitivity and specificity (80 to 98.6%, respectively). However, differentiation between AK and invasive SCC cannot be done with high specificity with RCM since both entities show full-thickness atypia and the depth of imaging is limited. Beside this, the hyperkeratotic stratum corneum is a limiting factor by hindering the penetration of light in RCM. (8)

Figure 4: On clinical examination, actinic keratosis appears as a red to brown macule or plaque with a rough surface.

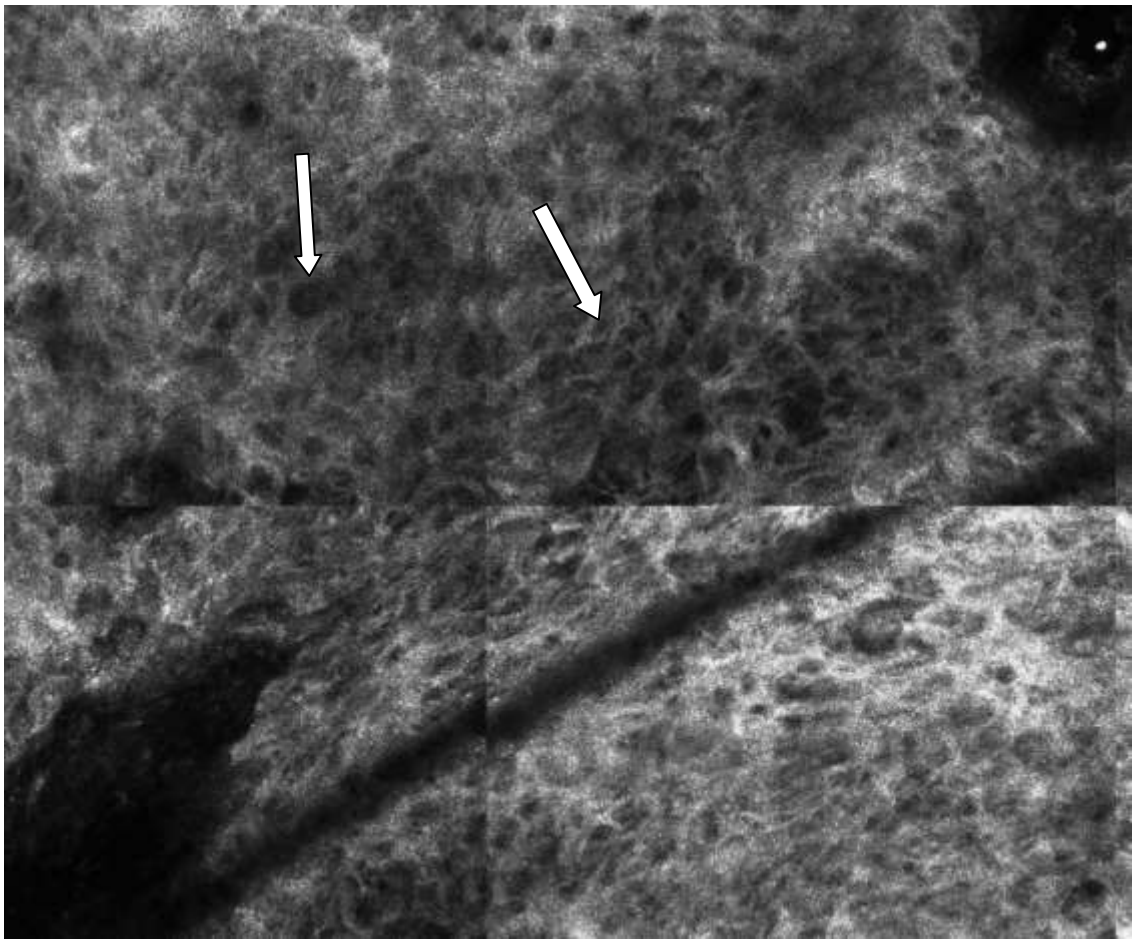

Figure 5: Actinic keratosis. At granular-spinous layer, an irregular honeycomb pattern is seen (approximately 500x500µm field-of-view) consisting of lines varying in width and brightness and holes varying in size and shape (arrows) correlating to abnormal keratinocytes.

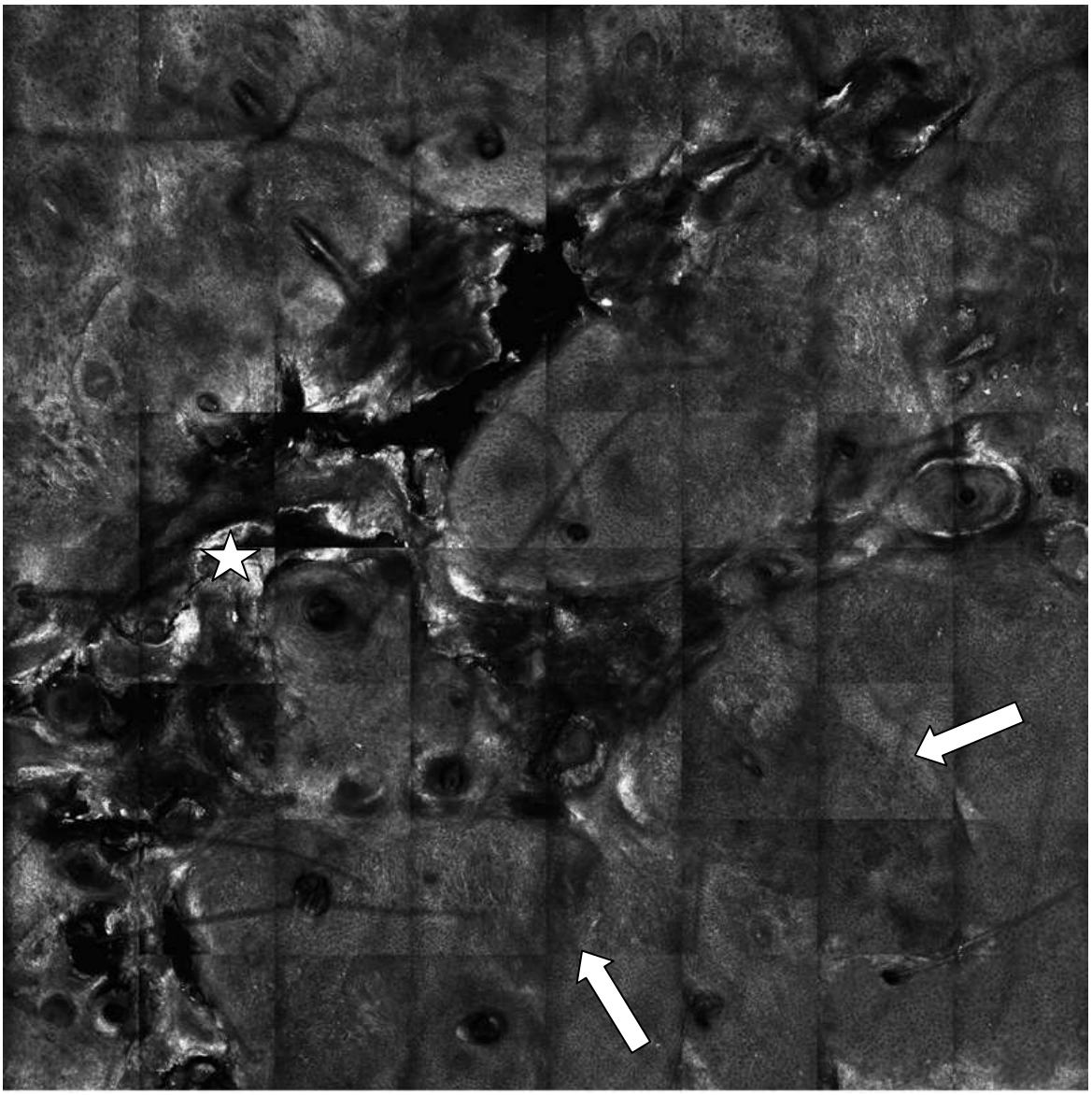

Figure 6: Actinic keratosis. The RCM mosaic of the upper epidermis (oblique section) shows the transition from normal honeycomb pattern (healthy skin) on the lower right corner to an irregular honeycomb pattern (upper left part of the image, thick arrows). In the middle, an irregular bright reflecting stratum corneum is seen (asterisk).

Basal cell carcinoma:

On clinical examination, basal cell carcinoma (BCC) usually appears as pink-white patch, plaque or nodule, correlating to the various histopathological subtypes (superficial, fibrosing and nodular type). (see figures no. 7 and no. 8) Pigmented BCCs are observed less frequently and show a focal or global tan to gray-blue coloration. Pigmented BCCs or non-pigmented superficial BCCs may be difficult to differentiate from other skin lesions (mainly melanoma), clinically and with dermoscopy. With RCM, BCC may be easily diagnosed by presence of round to oval tumor islands, which appear as hyporeflective structures (see figure no. 9) or as bright reflective areas with surrounding dark cleft-like spaces, depending on the amount of melanin accumulated within the tumor cells. Between the tumor islands, bright dendritic cells are visualized which represent either melanocytes or Langerhans cells. The RCM features enumerated correlate well with the histopathological findings. However, an interesting phenomenon which is only seen with RCM is the so-called “streaming” of distorted basaloid cells. These basaloid cells are focally elongated and strictly oriented along the same axis in the horizontal plane of RCM imaging. (see figure no. 10) The cells are supposed to either correlate to the palisading tumor cells seen in histopathology or to “stretched” keratinocytes above tumor islands. Dense bright bundles of collagen and tortuous, branching blood vessel within the upper dermis seen with RCM correspond to the fibrotic tumor stroma and ectatic blood vessels seen in histopathology. (9) The blood flow with leucocytes rolling along the walls of the dilated vessels between the tumor islands can easily be visualized in RCM, thereby providing significantly more details than dermoscopy. (10) A retrospective randomized and blinded multicenter study of 152 BCCs has shown that a correct diagnosis can be done with a high sensitivity and specificity using 5 RCM criteria and has reinforced RCM as a helpful tool to support dermoscopy and clinical examination in diagnosing BCC. (11) A pilot study has demonstrated the value of RCM for detecting residual BCC in patients undergoing Mohs surgery. (12)

Figure 7: On clinical examination, BCC usually appears as non-pigmented nodule with ectatic blood vessels as shown here.

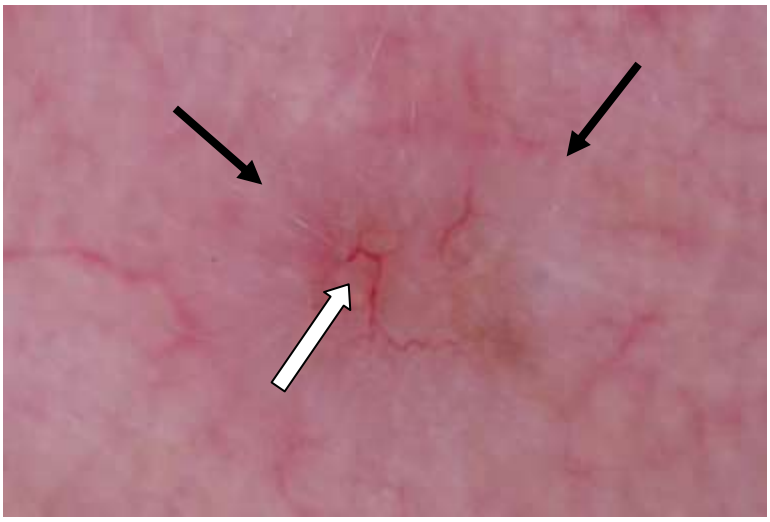

Figure 8: With dermoscopy, BCC is characterized by a pink-white structureless area (black arrows) and arborizing ectatic blood vessels (white arrow).

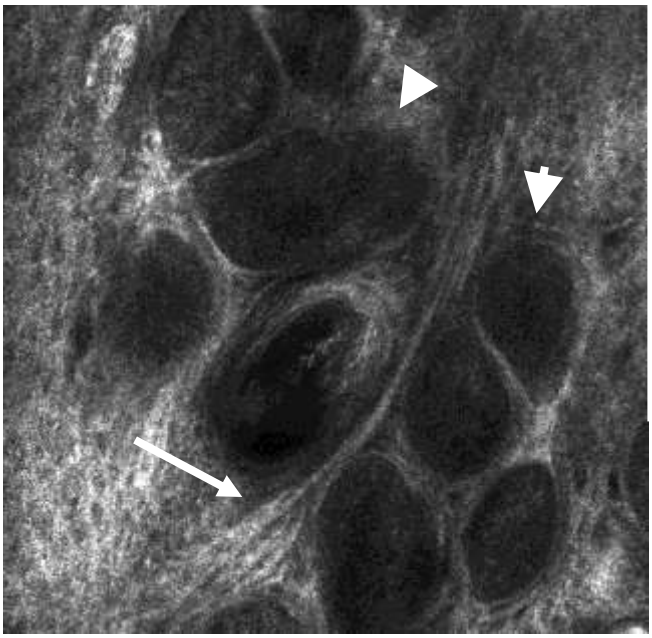

Figure 9: Hyporeflective tumor islands (arrowheads) between strong bright bundles of collagen (arrow) are displayed within the upper dermis in RCM.

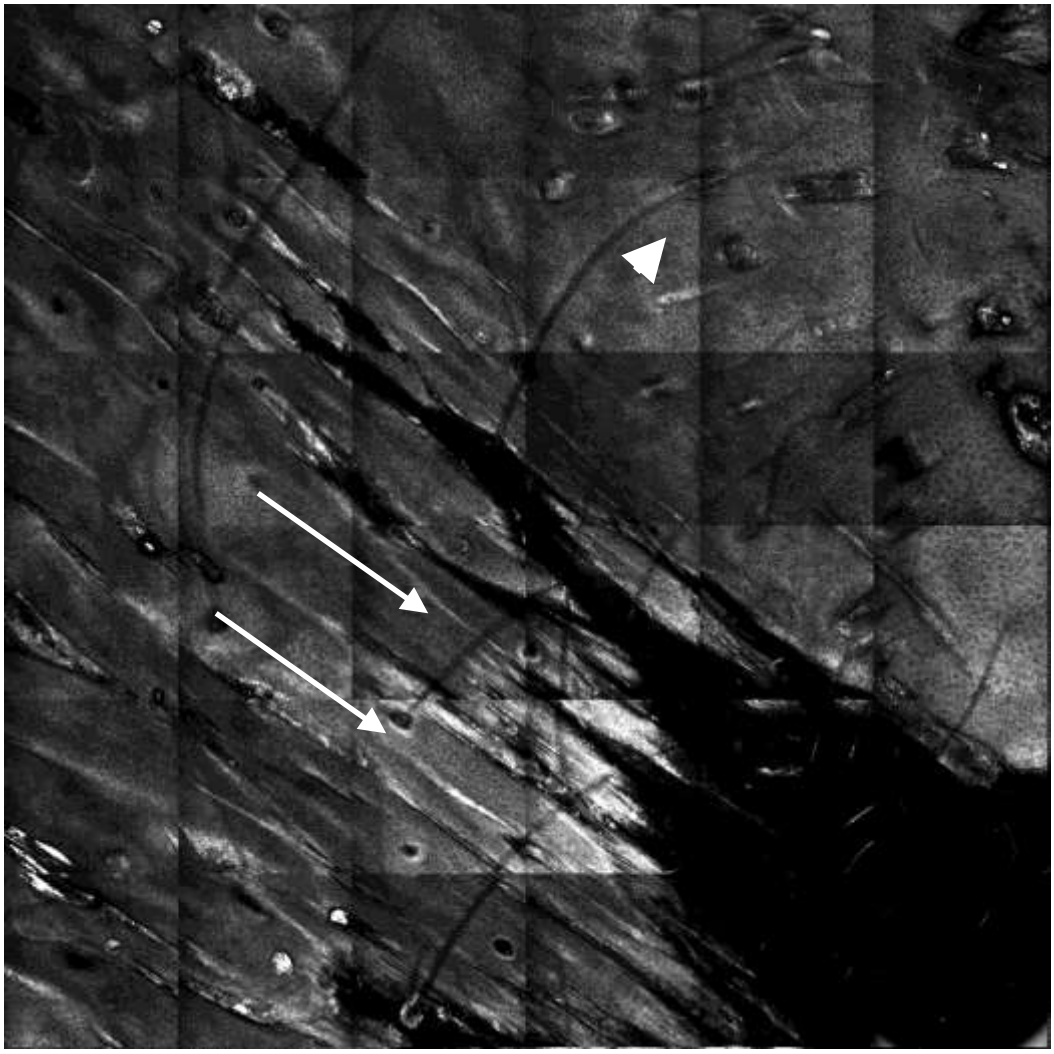

Figure 10: The so-called "Streaming" is a phenomenon which is frequently found in BCC. In the left lower corner all cells of the granular-spinous layer seem to be stretched along the same axis (arrows). In the upper right corner a normal honeycomb pattern is visualized (arrowhead).

Melanocytic skin tumors:

RCM has been described as a useful tool for differentiating melanocytic from non-melanocytic skin lesions and benign nevi from melanoma. Image evaluation is facilitated in pigmented skin lesion due to the high reflectivity of melanin in RCM. RCM images of melanocytic nevi and melanoma were first published by Langley et al in 2001. (13) Different patterns of nevi and melanoma were described by the authors: nevi showed cohesive nests of uniform round cells and a regular epidermal architecture. Melanomas were characterized by a disarranged epidermal pattern, atypical polymorphous melanocytes and pagetoid spreading of melanocytes. Later, Ahlgrimm et al (14) described three main RCM patterns in 30 common nevi according to the variable dermoscopic and histopathological nevus subtypes. In 2007 Scope et al. published a glossary of RCM criteria and terms for describing and diagnosing melanocytic skin lesions. (15) Pellacani et al (16) described cerebriform dermal clusters of melanocytes which were only found in melanoma while dense nests of melanocytes in the lower epidermis and at the dermo-epidermal junction were significant for benign nevi and were rarely found in melanoma. Sparse cells clusters were frequently seen in melanoma but also in one Spitz naevus. A sheet like distribution of single atypical cells at the dermo-epidermo-junction, corresponding to melanocytes, was also described as characteristic RCM feature of melanoma. These terms have shown a good reproducibility in a larger retrospective blinded study including 136 melanomas and 215 nevi. (17)

In 2007, Pellacani et al. (17) defined six RCM criteria for diagnosing melanoma; two major criteria (non-edged papillae, atypical cells) and four minor criteria (isolated nucleated cells within dermal papillae, pagetoid cells, widespread pagetoid infiltration and cerebriform clusters in the dermis) contributing to a scoring algorithm (each major criterion 2 points, each minor criterion 1 point), whereby a score ≥ 3 signifies melanoma. A two-step algorithm for differentiating melanoma from benign lesions has been published later by Segura et al. in 2009 (18): In the first step, melanocytic skin lesions are differentiated from non-melanocytic lesions by evaluation of four RCM criteria (presence of cobblestone pattern at epidermal layers, pagetoid spread, mesh appearance of the dermo-epidermal junction, and dermal nests), and then scored according to Pellacani's

(19) algorithm. RCM proved to be a helpful tool for diagnosing melanoma in this study; however, these results have to be evaluated in further clinical studies containing a larger number of lesions. Furthermore Lorber et al. (20) evaluated 857 CSLM images of 408 nevi and 448 melanomas by automatic image analysis (classification tree analysis). Three nodes signified nevi and six nodes indicated malignancy. This analysis tree showed a high correlation between recently published criteria and a correct diagnosis.

A study of 202 lesions including 76 melanomas and 114 nevi showed a good correlation of dermoscopic and histopathologic findings with RCM patterns. For example, black dots seen in dermoscopy referred well to clusters of pagetoid cells in melanoma visualized with both, RCM and histopathology; retiform bundles of collagen and intermingled melanophages observed with RCM and histopathology correlated well with regression structures visualized with dermoscopy. (19)

Nevi:

In the daily routine, the clinical and dermoscopic differentiation between dermal nevi and other skin neoplasms can easily be done while junctional nevi may be confounded with melanoma. In dermoscopy, junctional nevi usually show a regular pigment network and small round structures called globules as signs of benignity. (see figure no. 11)

In conventional histopathology, junctional nevi are characterized by an increase of melanocytes in single units along the DEJ and presence of small nests at the tips and sides of the rete ridges. In RCM, a ringed pattern is seen at the DEJ; the typical ringed pattern is characterized by monomorphous bright cells (melanocytes and pigmented keratinocytes) located around the dermal papillae (“edged papillae”) and by small melanocytic nests protruding into the papillary dermis. These RCM features correlate well with the dermoscopic and histopathologic findings. (see fig. no 12)

Melanoma:

Clinically, malignant melanoma (MM) is diagnosed by using the ABCDE criteria: lesion asymmetry, irregular borders, color variegation, diameter more than 5 millimeters, and distinct elevation are suspicious for malignancy. With dermoscopy, presence of ≥3 global structures, color variegation, appearance of regression structures, and atypical vessels are characteristic findings for melanoma. (see figure no. 13).

With RCM, a typical “honeycomb” pattern or a typical “cobblestone” pattern are usually not found. Instead, a so-called “disarranged honeycomb” pattern is characteristically observed at epidermal layers due to pagetoid infiltration by atypical melanocytes, singly and in clusters. In addition, a loss of the regular DEJ architecture (ringed pattern) is seen due to presence of irregular melanocytic nests or a sheet-like distribution of atypical melanocytes at the DEJ.

Pagetoid atypical melanocytes are visualized as large, round, spindle-shaped or dendritic nucleated cells which vary in size, shape and reflectivity. As known from histopathology, pagetoid cells are very rare in benign skin lesions but are commonly found in melanoma. Recognition of this pattern in RCM has been described as a useful method for diagnosing melanoma. (21) However, a pagetoid spread of melanocytes is not always present in melanoma and other RCM features have additionally to be considered. (22)

Dermoscopy and histopathology of facial LM vary from other melanoma subtypes; again, RCM shows a high correlation with histopathology and dermoscopy. (23) Adnexal invasion by atypical melanocytes, sheets of mainly dendritic melanocytes at the DEJ and/or granular-spinous layers and cord-like rete ridges at the DEJ have been recently described as characteristic RCM findings of facial lentigo maligna. These criteria showed a high reproducibility in a case series. (24) Characteristic RCM features of lentigo maligna are shown in figure no. 14.

RCM has been shown to be a useful tool in distinguishing pigmented and non-pigmented benign and malignant facial macules (e.g. nevi from lentigo maligna) by using a score which has been proposed by Guitera et al in 2010. (25) The score is based on an examination of 203 benign macules and 81 lentigo malignas.

The two major criteria were presence of non-edged papillae and pagetoid cells larger than 20 µm (each criterion scored +2 points). Three of the four minor criteria scored +1 point: three or more atypical cells at the dermo-epidermal junction in five 0.5 x 0.5 mm fields, follicular localization of atypical cells, and nucleated cells within the dermal papillae. A broadened honeycomb pattern scored -1 point. Presence of two or more points resulted in a high sensitivity of 85% and specificity of 76% for the diagnosis of lentigo maligna.

RCM of solar lentigo and seborrheic keratoses:

With RCM, solar lentigo (SL) displays a regular honeycomb pattern of the epidermis and densely packed, round to polymorphous, well-circumscribed dermal papillae as well as cord-like rete ridges at the DEJ. (26) As compared to melanoma there are no irregularly shaped cells or melanocytic nests found in benign lentigines. (27) (see fig. no 15) Seborrheic keratoses show features of SL and additional RCM findings mainly due to increased thickness of the epidermis, such as cerebriform surface structures (epidermal projections and keratin-filled invaginations of lesion surface).

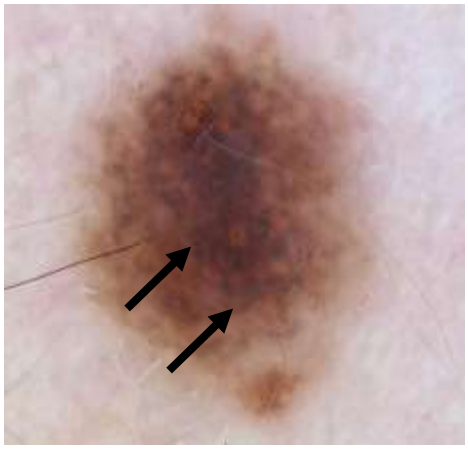

Figure 11: Dermoscopic image of a typical junctional nevus displaying network structures and dark brown to black globules (arrows).

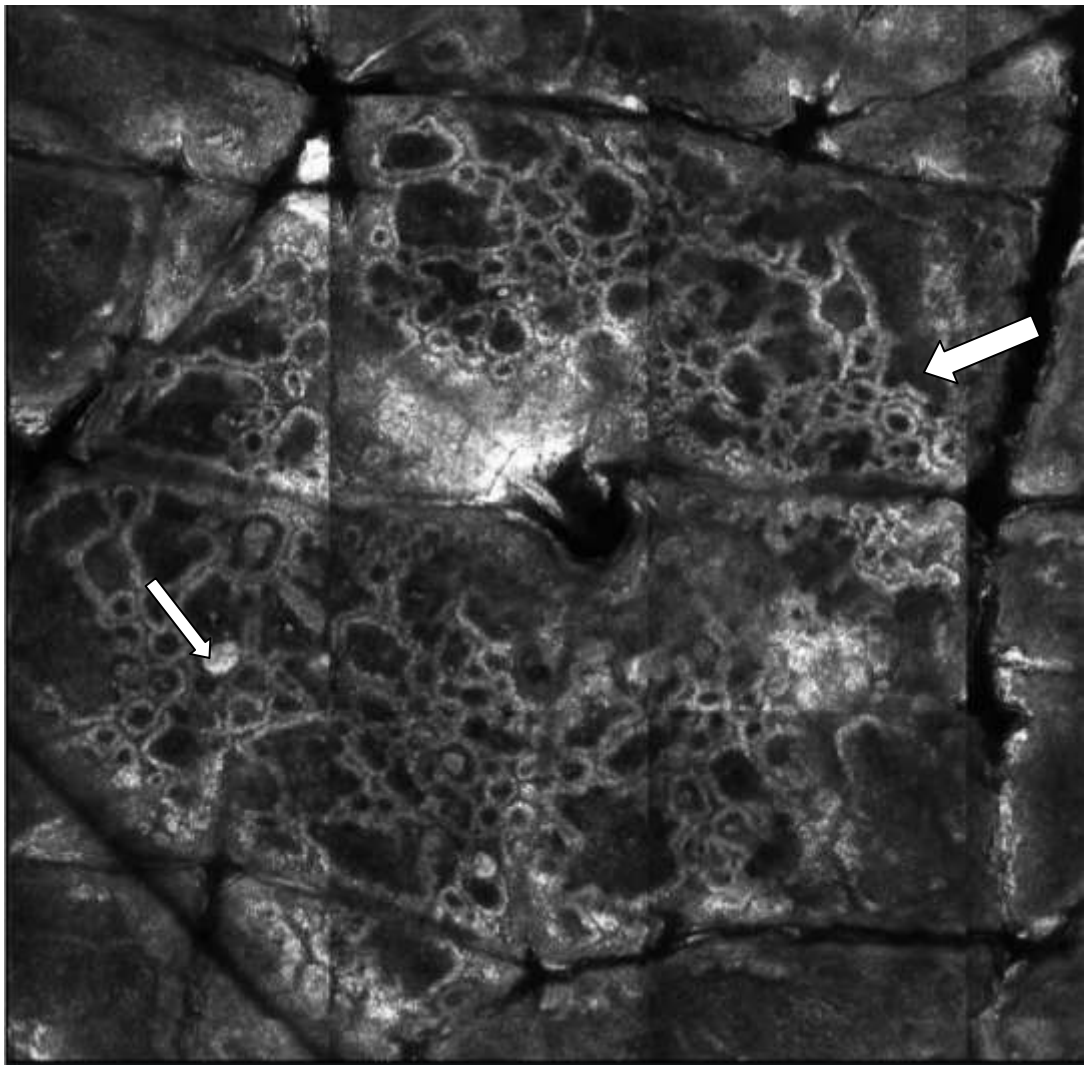

Figure 12: Junctional nevus. A typical ringed pattern (thick arrow) and small junctional melanocytic nests (thin arrow) are seen at the DEJ.

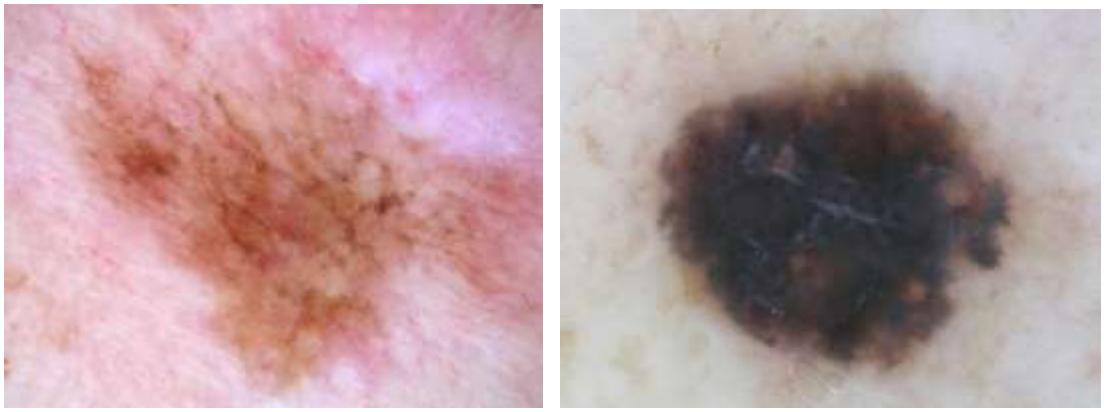

Figure 13: Dermoscopic images of malignant melanomas. On the left picture, a lentigo maligna with a so-called annular-granular pattern is seen. The granular structures correlate to melanocytic nests seen with histopathology and RCM. The lesion on the right shows a homogenous center with peripheral pseudopods. Almost the entire lesion shows a blue-white hue, which is frequently seen in melanoma; blue-white areas indicate either tumor regression or tumor invasion into the dermis, such as in this case.

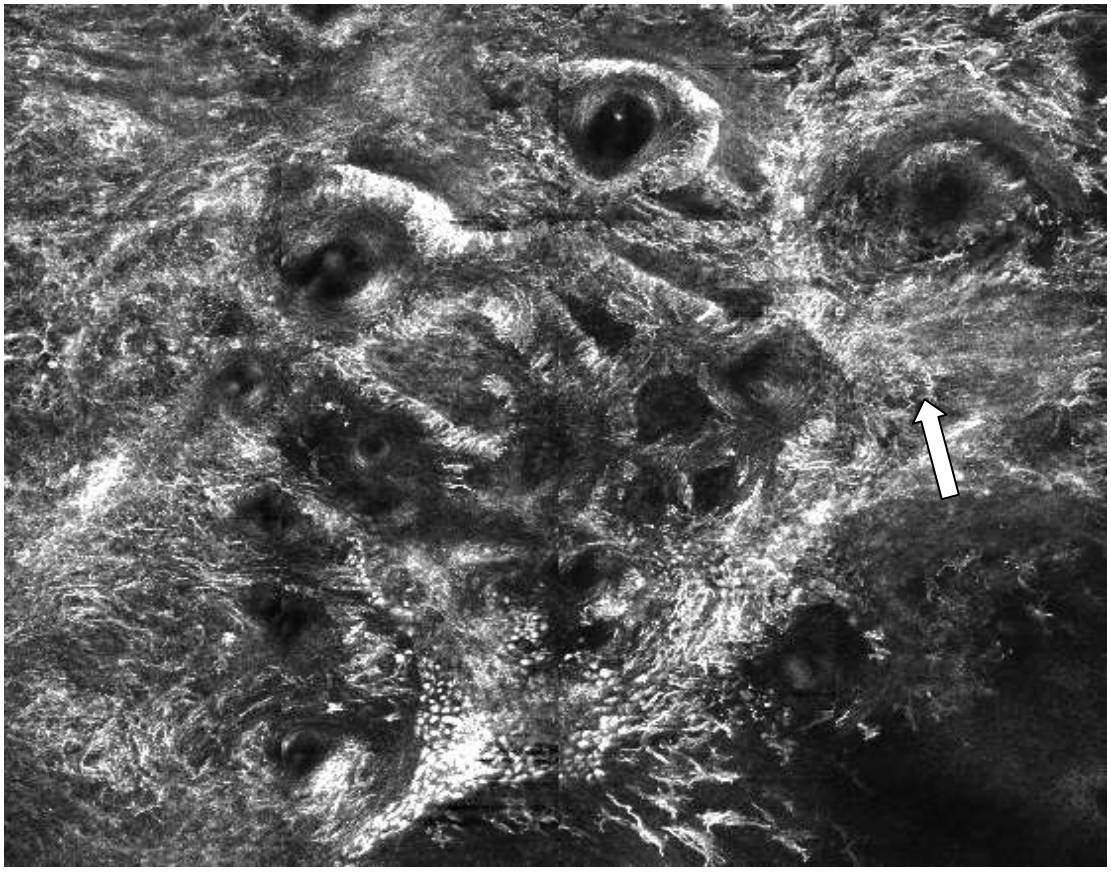

Figure 14: Lentigo maligna (mosaic 2x2mm). Sheets of mainly dendritic atypical melanocytes (asterisk) and loss of the regular DEJ architecture are observed. A descent of atypical melanocytes along adnexal structures (arrow) and cord-like rete ridges are additionally seen.

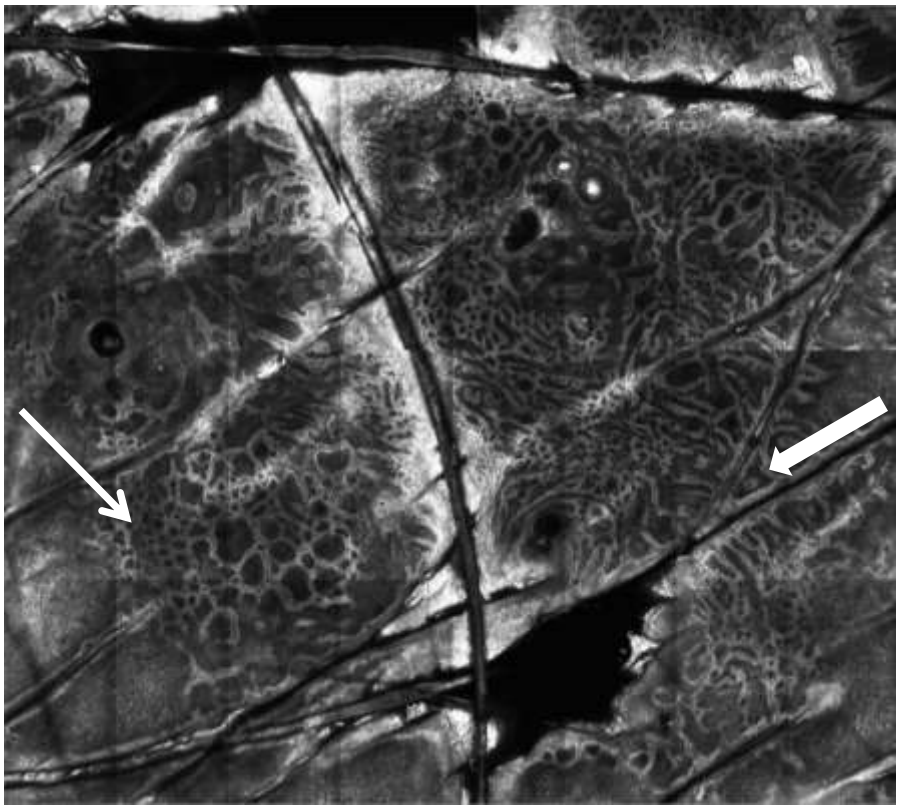

Figure 15: Solar lentigo (mosaic, 4x4mm). Densely packed, round to polymorphous papillae surrounded by a rim of bright monomorphous cells (thin arrow) and cord-like rete-ridges (thick arrow) are seen at the DEJ.

RCM-study:

Objective:

The aim of the present study was to evaluate the impact of RCM as adjunct to clinical and dermoscopic examination in differentiating benign from malignant facial skin lesions. The value of RCM in the diagnosis of melanoma has been shown in several studies; one study included also macular facial skin lesions. However, the unique histomorphology of facial skin was not considered in these studies and classic melanoma criteria were applied to diagnose lentigo maligna. In addition, there is no greater scientific evidence of the sensitivity and specifity of the proposed diagnosing criteria. Many authors of these articles agree that blinded studies are required to validate the results of these studies in a large number of cases. There are also many pilot studies and case reports which confirm above mentioned criteria, but studies that show the interobserver reproducibility of these criteria are still lacking. However, RCM is considered an adequate method to avoid unnecessary biopsies in several expert reviews.

Patients:

Patients were recruited consecutively between June 2008 and November 2010 from the outpatient clinic of the Department of Dermatology, University Hospital Graz, Austria. Acquired data contained dermoscopic images taken with the Vivacam®, conventional clinical and dermoscopic pictures taken by a digital camera (Nikon D200) and the RCM images acquired with the Vivascope 1500® reflectance confocal microscope.

A total of 160 facial skin lesions were imaged in 148 patients. Patients were aged from 5 to 93 years with a mean age of 65.1 years. 42 patients (28%) were male and 106 (72%) were female. 12 lesions (8%) were labial lesions. Patients´ skin types ranged from skin phototype I to skin phototype IV according to the major skin phototypes found in middle Europe.

A biopsy followed by histopathological examination was taken in 128 cases (80%) based on clinical and dermoscopic examination, the remaining skin lesions were

clinically and dermoscopically clear-cut benign lesions. Complete excision was done in 79 cases (49%) while punch biopsy was done in 36 cases (23%) and shave biopsy in only 13 cases (8%). A short term follow-up after three months was chosen in 18 of the clinically benign appearing lesions (11%) and no further action was taken in the remaining 14 cases (9%).

98 (61%) lesions were malignant (biopsy proven) and 62 (39%) were benign (biopsy proven or clinically clear-cut). 87 (54%) skin lesions were melanocytic and 73 (46%) were non-melanocytic in origin. Corresponding diagrams are seen in figure no. 16. The melanocytic lesions included 71 (44%) melanomas, including 5 recurrent melanomas and 2 melanoma metastases. The 16 (10%) nevi included 5 atypical naevi and one Spitz nevus. The 73 (46%) non-melanocytic skin tumors included 15 (9%) basal cell carcinomas, 16 (10%) solar lentigines, 8 (5%) seborrheic keratoses and 7 (4%) cases of actinic keratoses and squamous cell carcinomas. Furthermore there were 5 (3%) lymphomas and 22 (14%) cases of miscellaneous benign lesions including scars and unspecific pigmentations. The corresponding diagram providing absolute numbers and percentages is seen in figure no. 17

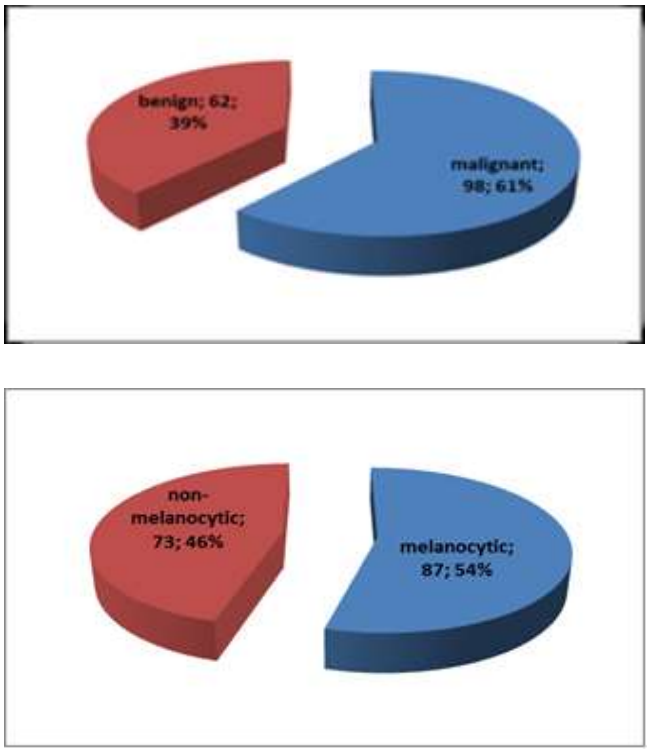

Figure 16: Benign/malignant ratio and melanocytic/non-melanocytic ratio of included skin lesions.

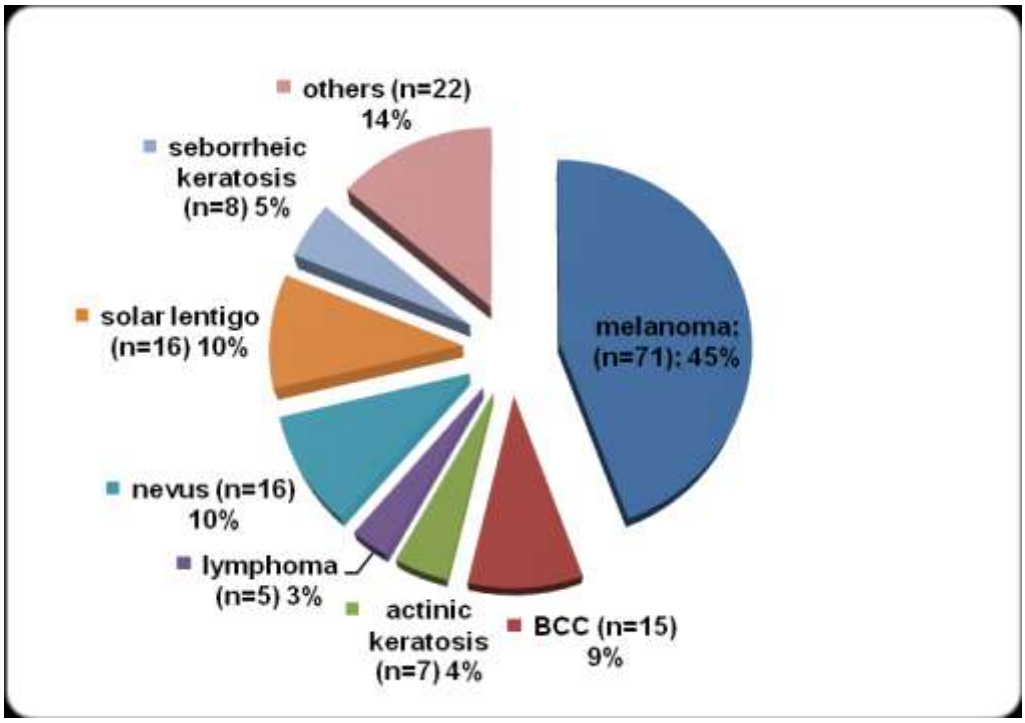

Figure 17: Absolute numbers and percentages of diagnoses for included lesions.

Image evaluation:

Three data sets were acquired in three consecutive steps; in the first step only RCM images were evaluated by the observers (two experts and one novice in RCM) who were blinded to the diagnoses of the skin lesions. In the second step the clinical and dermoscopic images were assessed together by the same blinded observers. In the third step all images were evaluated together. The single novice observer performed two courses of image evaluation (each with three steps) to evaluate the learning curve for RCM image analysis. The work flow is shown in figure no. 18.

For RCM image analysis, mosaics were first assessed concerning the overall lesion appearance (symmetry/asymmetry) and then for specific RCM criteria at various imaging depths:

- Stratum corneum /Skin surface
- Spinous-granular layer
- Dermo-epidermal junction
- Upper dermis

- Diagnostic RCM criteria applied were selected from previously published studies (22), (25). For more details see table no. 1

- After each step of image evaluation, the observers opted for one of three possible treatment strategies for each lesion: excision, follow-up or no further treatment, depending on the presumed diagnosis selected out of fourteen proposed diagnoses. Proposed diagnoses included: nevus, atypical nevus (Clark nevus), melanoma, melanoma metastasis, recurrent melanoma, basal cell carcinoma, actinic keratosis, Bowen's disease, squamous cell carcinoma, solar lentigo, seborrheic keratosis, unspecific pigmentation, lymphoma and "others".

- A frame for a more user-friendly desktop interface was created with Microsoft Access 2007®. This frame allowed the observers to see all proposed criteria and to select recognized criteria at the same time. The final diagnosis was chosen from a drop-down list. Three different databases were created for the three

consecutive evaluation steps. A screenshot of the user-friendly desktop interface is seen in Figure no. 19.

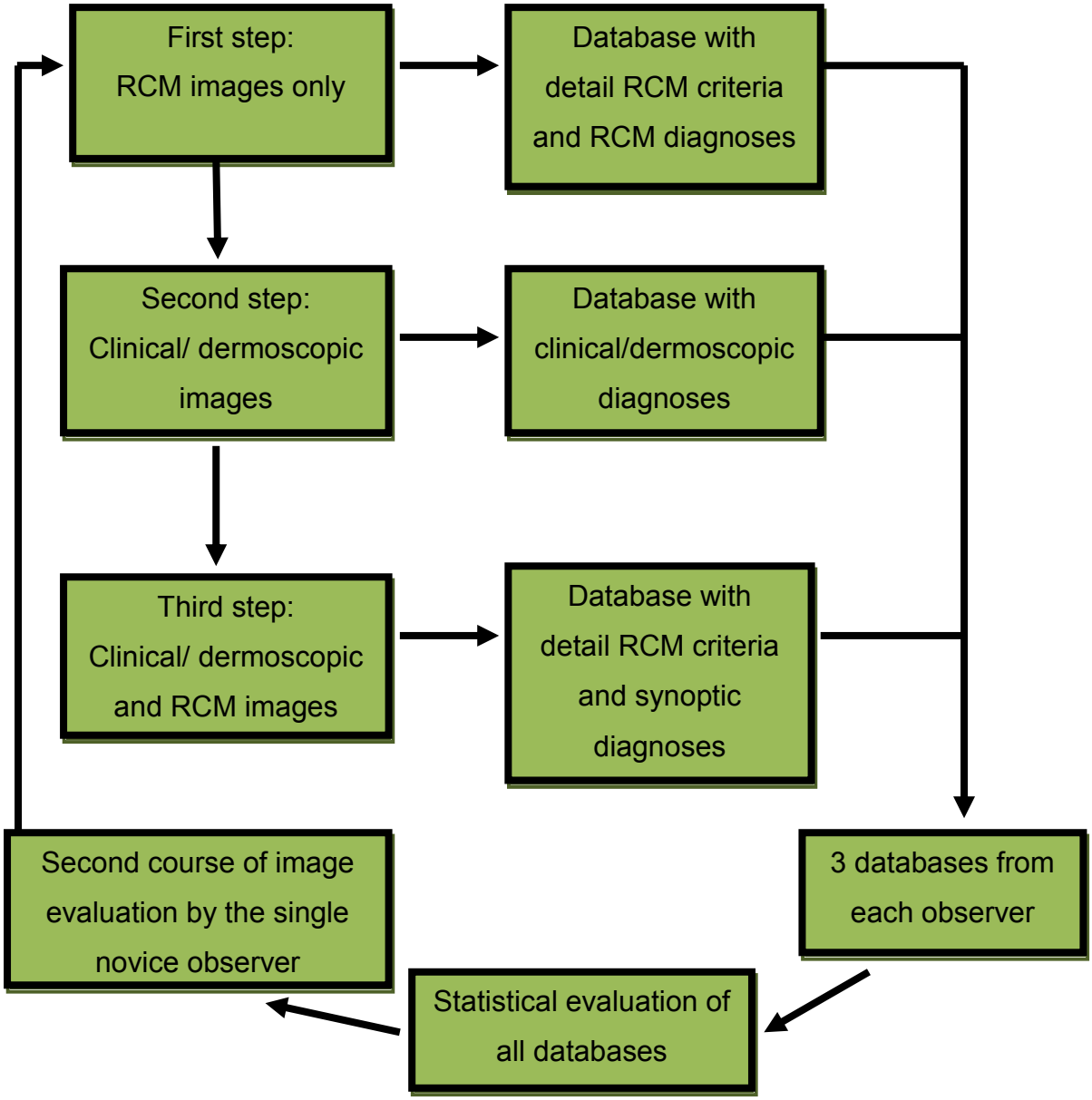

Figure 18: Diagram displaying work-flow of image evaluation.

Stratum corneum / Skin surface
Regular skin folds
Irregularly shaped, bright reflecting, structureless skin islands (scale; disruption of Stratum corneum)
Nucleated cells (parakeratosis)
Epidermal projections
Keratin-filled invaginations of the lesion surface
Spinous-granular layer
Typical honeycomb pattern
Typical cobblestone pattern
Broadened honeycomb pattern
Atypical honeycomb pattern
Atypical cobblestone pattern
Disarranged pattern (focal/complete loss)
Streaming
Large (>keratinocytes), round, nucleated reflective cells (pagetoid cells)
Dendritic, nucleated reflective cells (pagetoid cells)
Pleomorphic, nucleated reflective cells (pagetoid cells)
Sheets of bright reflecting dendrites or dendritic cells (pagetoid cells)
melanocytic nests
Small (<keratinocytes), bright cells (Inflammatory cells)
Horn cysts
Dermo-epidermal junction
Sharply demarcated (edged) papillae
Densely packed, round to polymorphous, edged papillae
Flattened DEJ (no papillae)
Ill defined (non-edged) papillae
Disarranged pattern (focal/complete loss)
Single, large, bright nucleated cells (atypical melanocytes)
Sheets of bright reflecting cells (atypical melanocytes)
Sheets of bright reflecting dendrites and dendritic cells (atypical melanocytes)
Cord-like rete ridges with bulbous projections
Round to oval, junctional dense melanocytic nests (protruding into dermal papillae)
Spindled or tubular junctional dense melanocytic nests (junctional thickenings between dermal papillae)
Round to oval sparse melanocytic nests
Cerebriform melanocytic nests
Junctional thickenings by single, bright, nucleated cells (atypical melanocytes)

Upper dermis
Fibrotic tumor stroma (thick, elongated bundles)
Solar elastosis
Bright, plump, non-nucleated cells (Melanophages)
Small (<keratinocytes), bright cells (Inflammatory cells)
Linear, tortuous and branching vessels
Dilated round blood vessels within dermal papillae ("button hole")
Linear and round vessels (mixed vascular pattern)
Isolated nucleated cells (melanocytes)
Cerebriform melanocytic nests
Dense melanocytic nests
Sparse melanocytic nests
Descent of bright, nucleated cells (melanocytes) along adnexal structures
Tumor islands (hypo-/hyperreflective)
Dendritic structures/cells within tumor islands
Cleft-like spaces around tumor islands

Table 1: Specific RCM criteria evaluated at various imaging depths.

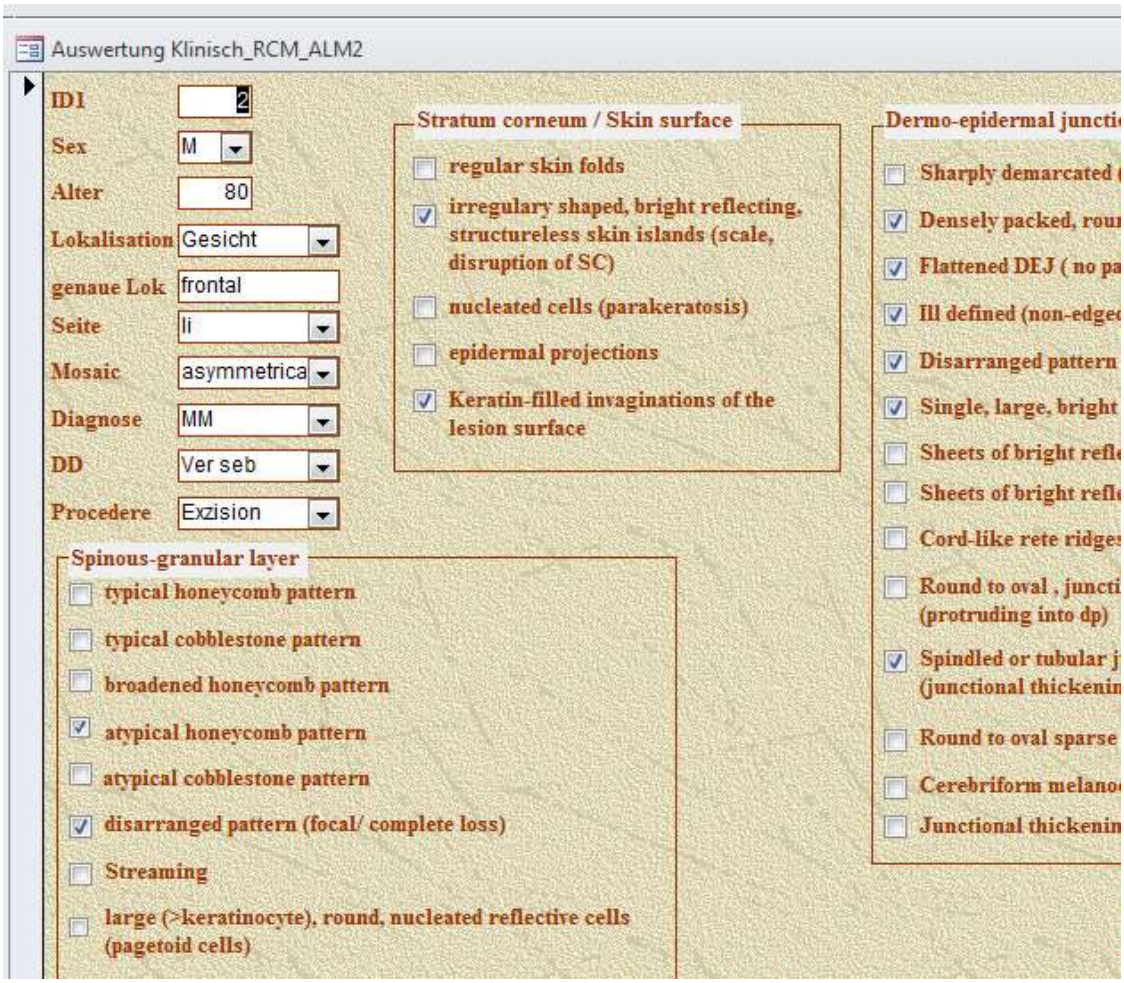

Figure 19: Screenshot of the user-friendly desktop interface created with Microsoft Access 2007®

Results:

Image evaluation and statistical analysis of data started in August 2011. The three Microsoft Access® 2007 databases for each course of image evaluation were collected from all three observers and compared with the diagnoses stated in the patient's chart or in the histopathological report, if available. Statistical analysis was done with the Microsoft Excel 2007® software. (See figure no. 18)

Single novice observer's first course of image evaluation:

In the first image evaluation done by the single novice observer, the sensitivity for all diagnoses was varying from 33 % to 75 % while the specificity was quite high varying from 56 % to 98 %. The harmonic, geometric and arithmetic mean of the sensitivity and specificity of all different questions showed a medium sensitivity varying from 50 % to 55 % but a high specificity varying from 78 % to 86 %. The highest sensitivity of 57 % was attained by using clinical, dermoscopic and RCM images together; the highest specificity of 87 % was attained by using only clinical and dermoscopic images. The sensitivity and specificity for melanoma detection based on clinical and dermoscopic images was 55 % and 88 %, respectively. RCM image analysis showed a sensitivity of 41 % and a specificity of 89 % for melanoma detection. The highest specificity (98 %) for melanoma detection was attained by evaluation of clinical, dermoscopic and RCM images together; the sensitivity, however, dropped to 35 %. The positive predictive value (PPV) was highest for evaluation of clinical, dermoscopic and RCM images together (0.93 vs. 0.74 – 0.78 for RCM or clinical and dermoscopic image evaluation alone). Negative predictive values ranged from 0.65 - 0.71.

For more details see table no. 2-3.

recommended procedure	Clinical and dermoscopic images		only RCM images		Clinical, dermoscopic, RCM images		Actually performed procedure	
excision	75	46.88%	73	45.63%	64	40.00%	128	80.00%
follow up	51	31.88%	38	23.75%	49	30.63%	18	11.25%
none	34	21.25%	49	30.63%	47	29.38%	14	8.75%

Table 2: Numbers and percentages of recommended procedures made in each step of the single novice observer´s first evaluation course.

Clinical and dermoscopic images		
	sensitivity	specificity
malignant versus benign	0.65	0.79
non MM malignant vs. benign	0.44	0.89
MM vs. non MM	0.55	0.88
BCC vs. non BCC	0.33	0.94
nevus vs. non nevus	0.75	0.83
harmonic mean	0.50	0.86
geometric mean	0.52	0.86
arithmetic mean	0.54	0.87
Only RCM images		
	sensitivity	specificity
malignant versus benign	0.48	0.82
non MM malignant vs. benign	0.52	0.92
MM vs. non MM	0.41	0.89
BCC vs. non BCC	0.60	0.98
nevus vs. non nevus	0.50	0.59
harmonic mean	0.49	0.81
geometric mean	0.50	0.83
arithmetic mean	0.50	0.84
Clinical, dermoscopic and RCM images		
	sensitivity	specificity
malignant versus benign	0.43	0.81
non MM malignant vs. benign	0.59	0.87
MM vs. non MM	0.35	0.98
BCC vs. non BCC	0.73	0.96
nevus vs. non nevus	0.75	0.56
harmonic mean	0.52	0.80
geometric mean	0.55	0.82
arithmetic mean	0.57	0.84

Table 3: Sensitivities and specificities for detection of malignant skin lesions and melanoma and for differentiation between non-melanoma skin cancer and benign lesions for each evaluation step of the single novice observer´s first evaluation course.

Clinical and dermoscopic images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	64	13	0.83	PPV
Diagnosis: benign	34	49	0.59	NPV
sensitivity	0.65	0.79	specificity	
Only RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	47	11	0.81	PPV
Diagnosis: benign	51	51	0.50	NPV
sensitivity	0.48	0.82	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	42	10	0.81	PPV
Diagnosis: benign	56	52	0.48	NPV
sensitivity	0.43	0.81	specificity	

Table 4: Statistical examination of the data of all three steps done by the single novice observer concerning the differentiation between benign and malignant lesions.

Statistical examination of the data obtained from the single novice observer concerning the problem of differentiating between malignant skin lesions and benign lesions showed the following results: The specificity of all three test methods was very high, varying from 79 % to 84 %. The sensitivity ranged from 43% for using clinical, dermoscopic and the RCM images up to 65 % if only the clinical and dermoscopic images were used. The positive predictive value (ppv = precision) was high from 0.80 to 0.83. The accuracy was only significant elevated up to 0.71 in the second step, while accuracy was poor up to 0.59 resp. 0.60 using RCM images. For more detail information see table no. 4.

Clinical and dermoscopic images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	12	7	0.63	PPV
Diagnosis: benign	15	55	0.79	NPV
sensitivity	0.44	0.89	specificity	
Only RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	14	5	0.74	PPV
Diagnosis: benign	13	57	0.81	NPV
sensitivity	0.52	0.92	specificity	
Clinical, dermoscopic and RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	16	8	0.67	PPV
Diagnosis: benign	11	54	0.83	NPV
sensitivity	0.59	0.87	specificity	

Table 5: View of the statistical data obtained from the single novice observer concerning the differentiation between non-MM skin cancer and benign lesions in all three steps.

Statistical examination of the data obtained from the novice observer concerning the differentiation between non- MM skin cancer and benign lesions showed a good specificity from 87 % to 92 % but a very low sensitivity from 44 % to 59 %. For more detail information see table no. 5.

Clinical and dermoscopic images (n = 160)				
MM	n = 71	n = 89	benign	
Diagnosis: MM	39	11	0.78	PPV
Diagnosis: non MM	32	78	0.71	NPV
sensitivity	0.55	0.88	specificity	
Only RCM images (n = 160)				
MM	n = 71	n = 89	benign	
Diagnosis: MM	29	10	0.74	PPV
Diagnosis: non MM	42	79	0.65	NPV
sensitivity	0.41	0.89	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
MM	n = 71	n = 89	benign	
Diagnosis: MM	25	2	0.93	PPV
Diagnosis: non MM	46	87	0.65	NPV
sensitivity	0.35	0.98	specificity	

Table 6: Statistical examination of the data obtained from the single novice observer concerning the diagnosis of melanoma in all three steps.

Statistical data obtained from the single novice observer for diagnosing MM showed a very high specificity from 88 % up to 98 % but sensitivity was very poor from 35 % to 55 %. The precision (positive predictive value = PPV) was elevated from 0.74 to considerable 0.93 using clinical, dermoscopic and RCM images. For more detail information see table no. 6.

Clinical and dermoscopic images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	5	9	0.36	PPV
Diagnosis: non BCC	10	136	0.93	NPV
sensitivity	0.33	0.94	specificity	
Only RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	9	3	0.75	PPV
Diagnosis: non BCC	6	142	0.96	NPV
sensitivity	0.60	0.98	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	11	6	0.65	PPV
Diagnosis: non BCC	4	139	0.97	NPV
sensitivity	0.73	0.96	specificity	

Table 7: Statistical view of the data obtained from the single novice observer concerning the diagnosis of BCC in all three steps.

Statistical examination of the data obtained from the single novice observer concerning the reliability for diagnosing basal cell carcinoma proofed a very high specificity ranging from 94 % to 98 % but sensitivity was only elevated up to 73 % in the third step where clinical dermoscopic and RCM images were used. The positive predictive value (ppv = precision) was widely varying from 0.36 to 0.73 but the negative predictive was very high from 0.94 to 0.98. For more detail information see table no. 7.

Clinical and dermoscopic images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	12	25	0.32	PPV
Diagnosis: non nevus	4	119	0.97	NPV
sensitivity	0.75	0.83	specificity	
Only RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	8	59	0.12	PPV
Diagnosis: non nevus	8	85	0.91	NPV
sensitivity	0.50	0.59	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	12	64	0.16	PPV
Diagnosis: non nevus	4	80	0.95	NPV
sensitivity	0.75	0.56	specificity	

Table 8: Statistical view of the data obtained from the single novice observer concerning the diagnosis of nevus in all three steps.

Statistical examination of the data obtained from the single novice observer concerning the reliability for diagnosing nevus showed a good sensitivity of 75 % and a specificity of 83 % only in the second step using clinical and dermoscopic images. In the two other steps the specificity was only elevated to 56 % and the precision was very poor at 0.12 resp. 0.16. However the negative predictive value was very high in all three steps varying from 0.90 to 0.97. For more information see table no. 8.

Single novice observer's second course of image evaluation:

The second evaluation of the clinical and dermoscopic images showed higher sensitivities for detection of melanoma, detection of malignant skin lesions and differentiation between non-melanoma skin cancer and benign lesions (63 %, 68% and 52 %, respectively) than for the first course of image evaluation (55 %, 65 %, and 44 %, respectively); specificities, however, were lower for the second image evaluation (79 %, 71 %, and 71 %, respectively) than for the first course of evaluation (88 %, 79 %, and 89 %, respectively). Evaluation of RCM images alone showed again higher sensitivities for the second (69 %, 71 %, and 52 %, respectively) than for the first (41 %, 48 %, and 52 %, respectively) evaluation; specificities were again lower for the second (83 %, 79 %, and 92 %, respectively) than for the first (89 %, 82 %, and 92 %, respectively) course. Second evaluation of all images together showed higher sensitivities (68 %, 71 %, and 59 %, respectively) than the first course of image evaluation (35 %, 43%, and 59%) and, except for detection of melanoma, higher specificities (88 %, 87%, and 97%) than for the first course of image evaluation (98%, 81%, and 87 %).

The sensitivity for melanoma detection was higher for RCM based image evaluations (68 % for evaluation of all images together and 69 % for RCM images alone) than for evaluation of clinical and dermoscopic images alone (63 %). The specificity was highest for evaluation of all images together (88 % vs. 79 % for evaluation of clinical and dermoscopic images and 63 % for RCM images alone). The positive predictive values (PPV) were higher for RCM based image evaluations (81 % for evaluation of all images together and 77 % for RCM images alone) than for evaluation of clinical and dermoscopic images alone (70 %); the negative predictive values were again higher for RCM based image evaluations (77 %) than for evaluation of clinical and dermoscopic images alone (73 %).

recommended procedure	Clinical and dermoscopic images		only RCM images		Clinical, dermoscopic, RCM images		Actually performed procedure	
excision	90	56.25%	88	55.00%	64	40.00%	128	80.00%
follow up	35	21.88%	33	20.63%	49	30.63%	18	11.25%
none	35	21.88%	39	24.38%	47	29.38%	14	8.75%

Table 9: Numbers and percentages of recommended procedures made in each step of the single novice observer´s second evaluation course.

Clinical and dermoscopic images				
	second course		first course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.68	0.71	0.65	0.79
non MM malignant vs. benign	0.52	0.71	0.44	0.89
MM vs. non MM	0.63	0.79	0.55	0.88
BCC vs. non BCC	0.47	0.97	0.33	0.94
nevus vs. non nevus	0.75	0.79	0.75	0.83
harmonic mean	0.59	0.78	0.50	0.86
geometric mean	0.60	0.79	0.52	0.86
arithmetic mean	0.61	0.79	0.54	0.87
Only RCM images				
	second course		first course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.71	0.79	0.48	0.82
non MM malignant vs. benign	0.52	0.92	0.52	0.92
MM vs. non MM	0.69	0.83	0.41	0.89
BCC vs. non BCC	0.60	0.98	0.60	0.98
nevus vs. non nevus	0.56	0.74	0.50	0.59
harmonic mean	0.61	0.84	0.49	0.81
geometric mean	0.61	0.85	0.50	0.83
arithmetic mean	0.62	0.85	0.50	0.84
Clinical, dermoscopic and RCM images				
	second course		first course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.71	0.87	0.43	0.81
non MM malignant vs. benign	0.59	0.97	0.59	0.87
MM vs. non MM	0.68	0.88	0.35	0.98
BCC vs. non BCC	0.67	0.98	0.73	0.96
nevus vs. non nevus	0.81	0.76	0.75	0.56
harmonic mean	0.69	0.88	0.52	0.80
geometric mean	0.69	0.89	0.55	0.82
arithmetic mean	0.69	0.89	0.57	0.84

Table 10: Harmonic, geometric and arithmetic mean of the sensitivity and specificity of different questions to the raw data obtained from the single novice observer and comparison with first course of image evaluation.

Clinical and dermoscopic images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	67	18	0.79	PPV
Diagnosis: benign	31	44	0.59	NPV
sensitivity	0.68	0.71	specificity	
Only RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	70	13	0.84	PPV
Diagnosis: benign	28	49	0.64	NPV
sensitivity	0.71	0.79	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	70	8	0.90	PPV
Diagnosis: benign	28	54	0.66	NPV
sensitivity	0.71	0.87	specificity	

Table 11: Statistical examination of the data of all three steps done by the single novice observer concerning the differentiation between benign and malignant lesions.

Statistical examination of the data obtained from the single novice observer concerning the problem of differentiating between malignant skin lesions and benign lesions showed the following results: The specificity of all three test methods was high, varying from 71 % to 80 %. The sensitivity ranged from 68 % to 71 %. The positive predictive value (ppv = precision) was high from 0.79 to 0.90. Note that nearly all parameters were significantly better in the second evaluation then compared with the first one. For more detail information see table no. 11.

Clinical and dermoscopic images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	14	18	0.44	PPV
Diagnosis: benign	13	44	0.77	NPV
sensitivity	0.52	0.71	specificity	
Only RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	14	5	0.74	PPV
Diagnosis: benign	13	57	0.81	NPV
sensitivity	0.52	0.92	specificity	
Clinical, dermoscopic and RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	16	2	0.89	PPV
Diagnosis: benign	11	60	0.85	NPV
sensitivity	0.59	0.97	specificity	

Table 12: View of the statistical data obtained from the single novice observer concerning the differentiation between non-MM skin cancer and benign lesions in all three steps.

Statistical examination of the data obtained from the novice observer concerning the differentiation between non- MM skin cancer and benign lesions showed a very high specificity from 71 % to 98 % but a moderate sensitivity from 52 % to 59%. For more detail information see table no. 12.

Clinical and dermoscopic images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	45	19	0.70	PPV
Diagnosis: non MM	26	70	0.73	NPV
sensitivity	0.63	0.79	specificity	
Only RCM images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	59	15	0.77	PPV
Diagnosis: non MM	22	74	0.77	NPV
sensitivity	0.69	0.63	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	48	11	0.81	PPV
Diagnosis: non MM	23	78	0.77	NPV
sensitivity	0.68	0.88	specificity	

Table 13: Statistical examination of the data obtained from the single novice observer concerning the diagnosis of melanoma in all three steps.

Statistical examination of the data obtained from the single novice observer concerning the differentiation between melanoma and other skin lesions showed a high specificity varying from 79 % to 88 % and a moderate sensitivity varying from 63 % to 69 %. The precision was elevated from 0.70 to 0.81 and the NPV was ranging from 0.73 to 0.77. Note that the sensitivity was far higher in the second examination than in the first one. For more detail information see table no. 13.

Clinical and dermoscopic images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	7	5	0.58	PPV
Diagnosis: non BCC	8	140	0.95	NPV
sensitivity	0.47	0.97	specificity	
Only RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	9	3	0.75	PPV
Diagnosis: non BCC	6	142	0.96	NPV
sensitivity	0.60	0.98	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	10	3	0.77	PPV
Diagnosis: non BCC	5	142	0.97	NPV
sensitivity	0.67	0.98	specificity	

Table 14: Statistical view of the data obtained from the single novice observer concerning the diagnosis of BCC in all three steps.

Statistical examination of the data obtained from the novice observer concerning the reliability for diagnosing BCC proofed a very high specificity ranging from 97 % to 98 % but sensitivity was moderate from 47 % to 67 %. The precision was varying from 0.58 to 0.77 and the NPV was very high from 0.95 to 0.97. The positive likelihood ratio was extremely high elevated from 13.53 to 32.22. For more detail information see table no. 14.

Clinical and dermoscopic images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	12	30	0.29	precision
Diagnosis: non nevus	4	114	0.97	NPV
sensitivity	0.75	0.79	specificity	
Only RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	9	38	0.19	precision
Diagnosis: non nevus	7	106	0.94	NPV
sensitivity	0.56	0.74	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	13	34	0.28	precision
Diagnosis: non nevus	3	110	0.97	NPV
sensitivity	0.81	0.76	specificity	

Table 15: Statistical view of the data obtained from the single novice observer concerning the diagnosis of nevus in all three steps.

Statistical examination of the data obtained from the single novice observer concerning the reliability for diagnosing nevus showed a widely varying sensitivity from 56 % to 81 % and a high specificity from 74 % to 79 %. The precision was poor at ranging from 0.19 to 0.29. However the negative predictive value was very high in all three steps varying from 0.94 to 0.97. Note that all parameters were better in the second examination done by the single novice observer than in the first one. For more information see table no. 15.

Expert observer's image evaluation:

In the image evaluation done by the two expert observers the harmonic, geometric and arithmetic mean of the sensitivity of all different questions showed a good sensitivity ranging from 57 % to 72 %. These parameters were slightly better than those of the single novice observer in his second course ranging from 59 % to 69 %. The harmonic, geometric and arithmetic mean of the specificity was closely ranging from 92 % to 93 %, significantly higher than those obtained by the single novice observer ranging from 78 % to 89 %. The highest sensitivity of 72 % and the highest specificity of 93 % were attained by the expert observers by using clinical, dermoscopic and RCM images together. Curiously, the arithmetic mean of the sensitivities for using only RCM images were minor to those obtained by the single novice observer (58 % vs. 62 %) but specificities were significantly higher (92 % vs. 85 %)

The highest sensitivity of 76 % and the highest specificity of 91 % for detecting melanoma were found by using clinical, dermoscopic and RCM images together. These were quite higher than those obtained by the single novice observer in his second course (58 % and 84 % respectively). The highest sensitivity of 80 % and a good specificity of 84 % were attained for differentiating malignant versus benign lesions by using clinical, dermoscopic and RCM images together, comparable to those obtained by the single novice observer (87 % and 71 % respectively). The highest specificity of 99 % and a good sensitivity of 67 % were obtained for diagnosing BCC by using clinical, dermoscopic and RCM images together, comparable to 98 % and 67 % respectively, obtained by the single novice observer.

For more information see tables no. 16-17.

recommended procedure	Clinical and dermoscopic images		only RCM images		Clinical, dermoscopic, RCM images		Actually performed procedure	
excision	131	81.88%	96	60.00%	116	72.50%	128	80.00%
follow up	9	5.63%	7	4.38%	16	10.00%	18	11.25%
none	20	12.50%	57	35.63%	28	17.50%	14	8.75%

Table 16: Numbers and percentages of recommended procedures made in each step of the expert observer´s evaluation course.

Clinical and dermoscopic images				
	Expert observers		novice observer 2nd course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.71	0.82	0.68	0.71
non MM malignant vs. benign	0.78	0.92	0.52	0.71
MM vs. non MM	0.65	0.90	0.63	0.79
BCC vs. non BCC	0.60	0.99	0.47	0.97
nevus vs. non nevus	0.63	0.98	0.75	0.79
harmonic mean	0.67	0.92	0.59	0.78
geometric mean	0.67	0.92	0.60	0.79
arithmetic mean	0.67	0.92	0.61	0.79
Only RCM images				
	Expert observers		novice observer 2nd course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.67	0.89	0.71	0.79
non MM malignant vs. benign	0.56	0.94	0.52	0.92
MM vs. non MM	0.56	0.90	0.69	0.83
BCC vs. non BCC	0.67	0.97	0.60	0.98
nevus vs. non nevus	0.44	0.94	0.56	0.74
harmonic mean	0.57	0.92	0.61	0.84
geometric mean	0.57	0.92	0.61	0.85
arithmetic mean	0.58	0.92	0.62	0.85
Clinical, dermoscopic and RCM images				
	Expert observers		novice observer 2nd course	
	sensitivity	specificity	sensitivity	specificity
malignant versus benign	0.80	0.84	0.71	0.87
non MM malignant vs. benign	0.74	0.95	0.59	0.97
MM vs. non MM	0.76	0.91	0.68	0.88
BCC vs. non BCC	0.67	0.99	0.67	0.98
nevus vs. non nevus	0.63	0.98	0.81	0.76
harmonic mean	0.71	0.93	0.69	0.88
geometric mean	0.71	0.93	0.69	0.89
arithmetic mean	0.72	0.93	0.69	0.89

Table 17: Harmonic, geometric and arithmetic mean of the sensitivity and specificity of different questions to the raw data obtained from the two expert observers and the novice observer.

Clinical and dermoscopic images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	70	11	0.86	precision
Diagnosis: benign	28	51	0.65	NPV
sensitivity	0.71	0.82	specificity	
Only RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	66	7	0.90	precision
Diagnosis: benign	32	55	0.63	NPV
sensitivity	0.67	0.89	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
malignant	n = 98	n = 62	benign	
Diagnosis: malignant	78	10	0.89	precision
Diagnosis: benign	20	52	0.72	NPV
sensitivity	0.80	0.84	specificity	

Table 18: Statistical examination of the data of all three steps done by the two expert observers concerning the differentiation between benign and malignant lesions.

Statistical examination of the data obtained from the two expert observers concerning the problem of differentiating between malignant skin lesions and benign lesions showed the following results:

The specificity of all three test methods was high, varying from 82 % to 89 %. The sensitivity ranged from 67 % to 80 %. The positive predictive value (ppv = precision) was high from 0.86 to 0.90. For more detail information see table no. 18.

Clinical and dermoscopic images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	21	5	0.81	precision
Diagnosis: benign	6	57	0.90	NPV
sensitivity	0.78	0.92	specificity	
Only RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	15	4	0.79	precision
Diagnosis: benign	12	58	0.83	NPV
sensitivity	0.56	0.94	specificity	
Clinical, dermoscopic and RCM images (n = 89)				
non MM malignant	n = 27	n = 62	benign	
Diagnosis: malignant	20	3	0.87	precision
Diagnosis: benign	7	59	0.89	NPV
sensitivity	0.74	0.95	specificity	

Table 19: View of the statistical data obtained from the two expert observers concerning the differentiation between non-MM skin cancer and benign lesions in all three steps.

Statistical data obtained from the two expert observers concerning the differentiation between non- melanoma skin cancer and benign lesions showed a very high specificity from 92 % to 95 % but a widely varying sensitivity from 56 % to 78 %. Accuracy was high from 0.82 to 0.89. The negative predictive value was elevated from 0.83 to 0.90. The positive likelihood ratio was elevated up to 15.31 if using clinical, dermoscopic and RCM images. For more detail information see table no. 19.

Clinical and dermoscopic images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	46	9	0.84	precision
Diagnosis: non MM	25	80	0.76	NPV
sensitivity	0.65	0.90	specificity	
Only RCM images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	40	9	0.82	precision
Diagnosis: non MM	31	80	0.72	NPV
sensitivity	0.56	0.90	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
MM	n = 71	n = 89	non MM	
Diagnosis: MM	54	8	0.87	precision
Diagnosis: non MM	17	81	0.83	NPV
sensitivity	0.76	0.91	specificity	

Table 20: Statistical examination of the data obtained from the two expert observers concerning the diagnosis of melanoma in all three steps.

Statistical examination of the data obtained from the two expert observers concerning the problem of differentiating between melanoma and other skin lesions showed a very high specificity varying from 90 % to 91 % and a moderate sensitivity varying from 56 % to 76 %. For more information see table no. 20.

Clinical and dermoscopic images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	9	2	0.82	precision
Diagnosis: non BCC	6	143	0.96	NPV
sensitivity	0.60	0.99	specificity	
Only RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	10	5	0.67	precision
Diagnosis: non BCC	5	140	0.97	NPV
sensitivity	0.67	0.97	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
BCC	n = 15	n = 145	non BCC	
Diagnosis: BCC	10	2	0.83	precision
Diagnosis: non BCC	5	143	0.97	NPV
sensitivity	0.67	0.99	specificity	

Table 21: Statistical view of the data obtained from the two expert observers concerning the diagnosis of BCC in all three steps.

Statistical examination of the data obtained from the two expert observers concerning the reliability for diagnosing basal cell carcinoma proofed a very high specificity ranging from 97 % to 99 % but sensitivity was moderate from 60 % to 67 %. The positive predictive value (ppv = precision) was varying from 0.67 to 0.83 and the negative predictive was very high from 0.96 to 0.97. The positive likelihood ratio was extremely high elevated from 19.33 to 48.33. For more detail information see table no. 21.

Clinical and dermoscopic images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	10	3	0.77	precision
Diagnosis: non nevus	6	141	0.96	NPV
sensitivity	0.63	0.98	specificity	
Only RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	7	9	0.44	precision
Diagnosis: non nevus	9	135	0.94	NPV
sensitivity	0.44	0.94	specificity	
Clinical, dermoscopic and RCM images (n = 160)				
nevus	n = 16	n = 144	non nevus	
Diagnosis: nevus	10	3	0.77	precision
Diagnosis: non nevus	6	141	0.96	NPV
sensitivity	0.63	0.98	specificity	

Table 22: Statistical view of the data obtained from the two expert observers concerning the diagnosis of nevus in all three steps.

Statistical examination of the data obtained from the two expert observers concerning the reliability for diagnosing nevus showed a widely varying sensitivity from 44 % to 63 % and a high specificity from 94 % to 98 %. The precision was widely ranging from 0.44 to 0.77. However the negative predictive value was very high in all three steps varying from 0.94 to 0.96. The positive likelihood ratio was extremely high from 7.00 to 30.00. For more information see table no. 22.

Comparison of detail RCM criteria

All databases collected from all three observers contained the lists of the ticked detail RCM criteria which were obtained from the first and the third step were compared together. Unfortunately data was well available in four databases from the two courses of image evaluation done by the novice observer but only one database from the second step done by the two expert observer was available in which only RCM images were evaluated.

The known and histopathological verified diagnoses from the patient's reports were used as a filter for the observed detail RCM criteria. Detail criteria were widely varying between each observer and each step too; thus meaning a high interobserver and intraobserver variability. Added percentages over hundred percent are caused by multiple observed detail RCM criteria in single lesions without any dominant pattern.

Melanoma:

Regular skin folds were observed in only the half cases compared to the novice observer (23 (32.39%) versus 42 (59.15%) to 49 (69.01%) cases) comparably to irregularly, bright reflecting, and structure-less skin islands with disruption of the stratum corneum found less cases by the expert observer (9 (12.68%) versus 19 (26.76% to 25 (35.21%) cases). In less than 14 (19.72%) cases nucleated cells (parakeratosis), epidermal projections or Keratin-filled invaginations of the lesion surface were observed in each step by each observer.

In the spinous granular layer a typical honeycomb pattern was found by the two expert observer in 38 (53.52%) cases similar to the numbers of the novice observer (38 (53.52%) to 46 (64.79%) cases). A disarranged pattern with focal or complete pattern loss was found in 29 (40.85%) to 31 (43.66%) by the single novice observer in each course but only in 2 (2.82%) cases by the expert observers. Furthermore atypical honeycomb pattern was mentioned in 11 (15.49%) cases by the two expert observers and in 7 (9.89 %) resp. 17 (23.94%) cases by the single novice observer. Sheets of bright reflecting dendrites were mentioned in 30 (42.25%) to 35 (49.30%) cases by the single novice observer but only in 8 (11.27%) cases by the two expert observers. Curiously melanocytic nests

were observed in 36 (50.70%) to 43 (60.56%) cases by the single novice observer but none could be detected by the two expert observers.

At the dermo-epidermal junction data obtained from all observers showed no domination for any detail criterion except very large interobserver and intraobserver variability. Absolute numbers and percentages between different observers were ranging from zero to 31 (43.66%) cases for a single criterion in the same step. Melanocytic signs were frequently mentioned however without any domination of any criterion.

In the upper dermis no domination of any possible criterion could be found in the data of all three observers except very large interobserver and intraobserver variability. Absolute numbers and percentages between different observers were ranging from 1 (1.41%) to 20 (28.17%) cases for a single criterion in the same step.

Nevus:

The statistical data obtained from all three observers concerning previously verified nevus mentioned a symmetrical lesion in 9 (56.25%) to 15 (93.75%) cases. In 1 (6.25%) to 5 (31.25%) cases the lesion has been reported as asymmetrical. Regular skin folds were dominating at the skin surface in 9 (56.25%) to 11 (68.75%) cases.

At the spinous-granular layer typical honeycomb pattern was dominating in 10 (62.50%) to 14 (87.50%) cases. Melanocytic nests were found in 9 (56.25%) to 12 (75.00%) cases by the single novice observer but curiously none could be found by the two expert observers. Interestingly all other criteria were mentioned in up to 7 (43.75%) cases in the first image evaluation done by the single novice observer, while in the second evaluation absolute number were only elevate 4 (25.00%) cases similar to those mentioned by the two expert observers. This fact demonstrates a high learning curve for recognizing detail criteria concerning the diagnosis of benign nevus.

At the dermo-epidermal junction all data from all observers showed no significant domination of any detail criterion except large interobserver and intraobserver

variability. Absolute numbers and percentages between different observers were ranging from zero to 6 (37.50%) cases for a single criterion in the same step.

In the upper dermis dilated round blood vessel were found in 6 (37.50%) to 12 (75.00%) cases by the single novice observer but only 3 (18.75%) were found by the expert observers. All other criteria were mentioned each in up to 6 (37.50%) cases in the first evaluation done by the single novice observer but in the second evaluation each criterion were mentioned only in 3 (18.75%) similar to those mentioned by the two expert observers. This fact demonstrates again a high learning curve for recognizing detail criteria concerning the diagnosis of benign nevus.

BCC:

The statistical data gained from all observers concerning previously verified BCC mentioned a symmetrical lesion in 5 (33.33%) to 8 (53.33%) cases. In 7 (46.67%) to 10 (66.67%) cases the lesion has been reported as asymmetrical. Regular skin folds were dominating at the skin surface in 5 (33.33%) to 7 (46.67%) cases. Irregularly shaped, bright reflecting, structureless skin islands were mentioned in 8 (53.33%) to 10 (66.67%) cases. Keratin filled invaginations of the lesion surface were observed in 3 (20.00%) to 7 (46.67%) cases.

Typical honeycomb pattern was found in 7 (46.67%) to 12 (80.00%) cases; while atypical honeycomb pattern has been mentioned in 4 (26.67%) cases to 9 (60.00%) cases. Furthermore disarranged pattern with focal or complete pattern loss has been observed in 3 (20.00%) to 8 (53.33%) cases. The characteristically “Streaming” could be found in 6 (40.00%) to 10 (66.67%) cases. All other detail criteria of the spinous-granular layer were mentioned each in less than 4 (26.67%) cases.

At the dermo-epidermal junction a disarranged pattern was found in 8 (53.33%) to 12 (80.00%) cases. Ill-defined non-edged papillae were observed in 2 (13.33%) to 6 (40.00%) cases. All other possible detail criteria were mentioned each in less than 3 (20.00%) cases.

At the upper dermis fibrotic tumor stroma were observed in 8 (53.33%) to 12 (80.00%) cases. Characteristically tumor islands were seen in 5 (33.33%) to 10

(66.67%) cases and larger dendritic structures within tumor islands in 6 (40.00%) to 8 (53.33%) cases.

Cleft-like spaces around tumor islands were mentioned in 6 (40.00%) to 10 (66.67%). The main vessel pattern were linear tortuous and branching vessels which were mentioned in 6 (40.00%) to 8 (53.33%) cases. However a mixed vascular pattern was observed in 5 (33.33%) to 8 (53.33%) cases. All other possible detail criteria were mentioned each in less than 3 (20.00%) cases.

Solar lentigo:

The statistical data obtained from all observers reported a symmetrical lesion in 10 (62.50%) to 15 (93.75%) cases. Regular skin folds were observed in 8 (50.00%) to 14 (87.50%) cases; while other criteria were mentioned each in less than 5 (31.25%) cases.

At the spinous-granular layer typical honeycomb pattern was observed in 8 (50.00%) to 15 (93.75%) cases. Surprisingly melanocytic nests were observed in 10 (62.50%) to 13 (81.25%) cases.

At the dermo-epidermal junction sharply demarcated (edged) papillae could be found in 7 (43.75 %) to 11 (68.75%) cases; while densely packed, round to polymorphous, edged papillae were observed in 5 (31.25%) to 9 (56.25%) cases. Furthermore round to oval, junctional dense melanocytic nests were found in 6 (37.50%) to 8 (50.00%) cases. All other criteria were mentioned in less than 4 (25.00%) cases.

In the upper dermis dilated round blood vessels within dermal papillae were observed in 4 (25.00%) to 12 (75.00%) cases; while linear, tortuous and branching vessels were mentioned in 1 (6.25%) to 5 (31.25%) cases. All other possible detail criteria were mentioned each in less than 4 (25.00%) cases.

Other (miscellaneous) lesions:

The statistical data obtained from all observers reported a symmetrical lesion in 16 (72.73%) to 19 (86.36%) cases; while an asymmetrical lesion was found in 4 (18.18%) to 5 (22.73%) cases. Regular skin folds were found in 11(50.00%) to 17 (77.27%) cases; while irregularly shaped, bright reflecting and structureless skin

islands were observed in 3 (13.64%) to 11 (77.27%) cases. All other criteria of the stratum corneum were mentioned in less than 4 (18.18%) cases.

At the spinous-granular layer typical honeycomb pattern was observed in 16 (72.73%) to 19 (86.36%) cases. All other detail criteria were observed each in less than 6 (27.27%) cases.

At the dermo-epidermal junction sharply demarcated edged papillae were seen in 4 (18.18%) to 13 (59.09%) cases. Densely packed, round to polymorphous, edged papillae and ill-defined papillae were mentioned each in 4 (18.18%) to 8 (36.36%) cases. All other criteria of the dermo-epidermal junction were observed in less than 5 (22.73%) cases.

At the upper dermis dilated round blood vessels within dermal papillae were observed in 5 (22.73%) to 16 (72.73%) cases. Fibrotic tumor stroma was observed in 3 (13.64%) to 7 (31.82%) cases. All others criteria of the dermo-epidermal junction were each seen in less than 4 (18.18%) cases.

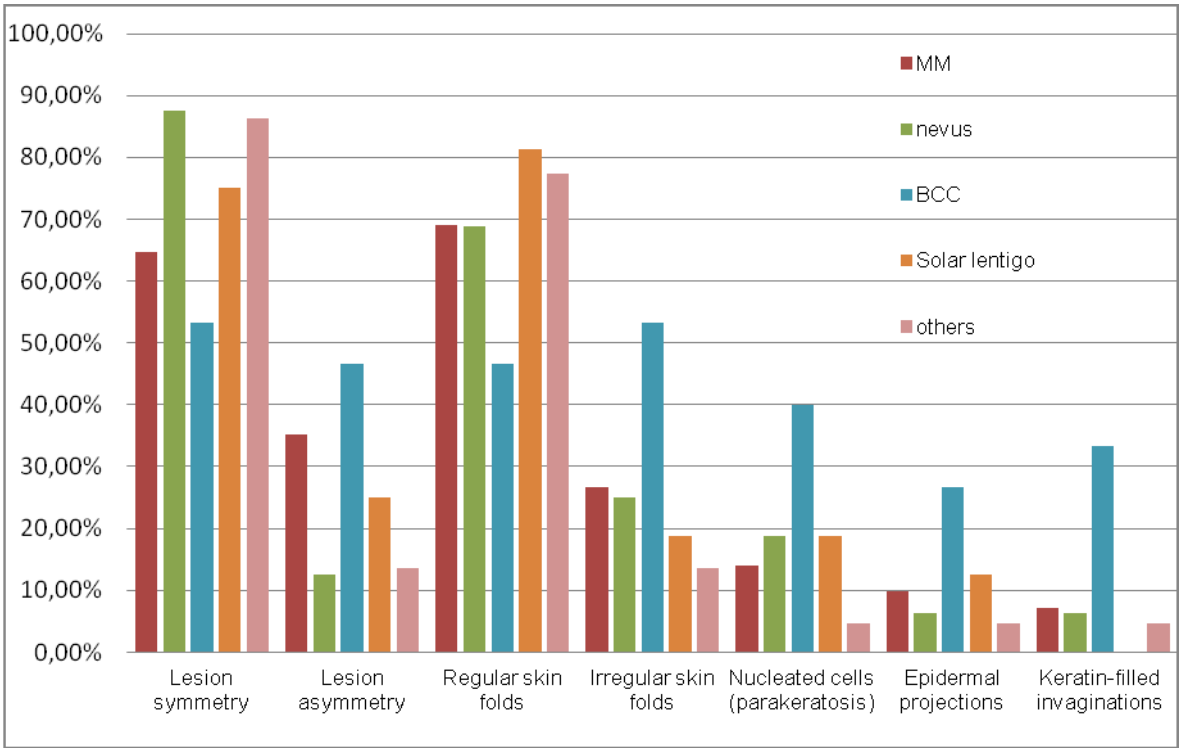

Figure 20: This diagram displays the statistical distribution of the observed criteria in the corneal layer for five different diagnoses using clinical, dermoscopic and RCM images. Data was obtained by the single novice observer's first course of image evaluation.

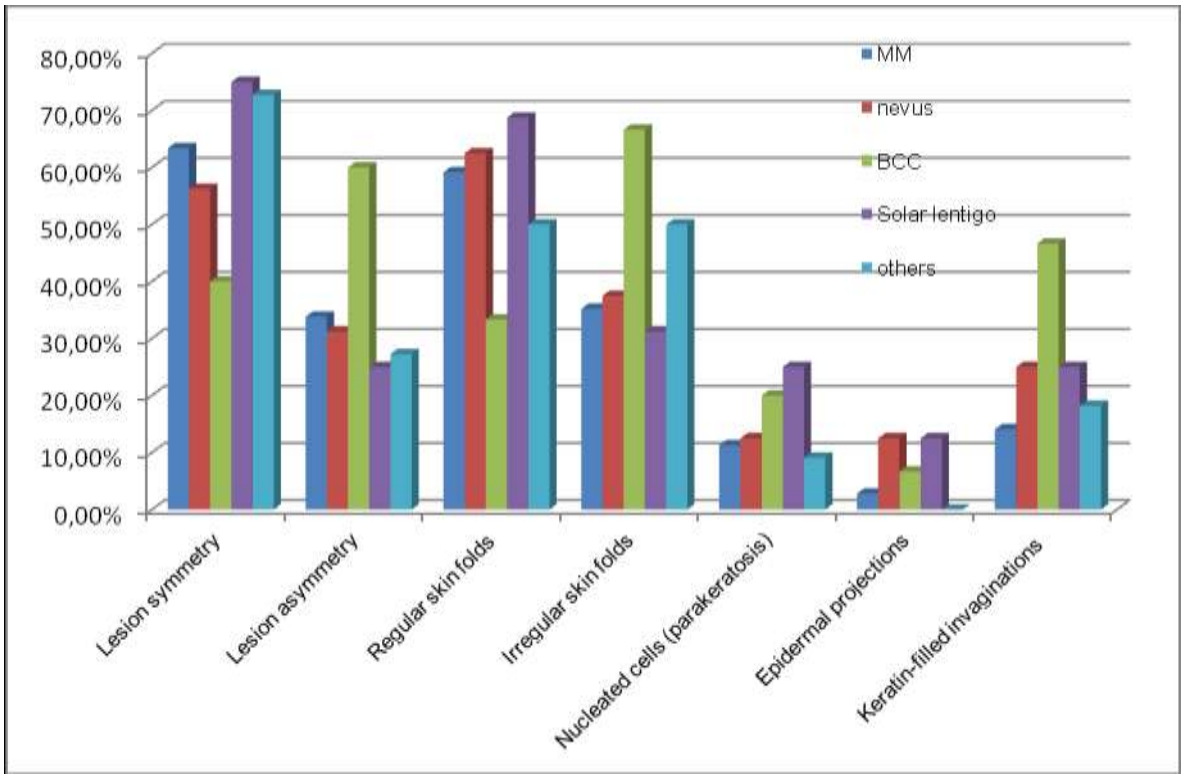

Figure 21: This diagram displays the statistical distribution of the observed criteria in the corneal layer for five different diagnoses using only the RCM images. Data was obtained by the single novice observer's first course of image evaluation.

Discussion:

The aim of our study was to evaluate the impact of RCM as adjunct to clinical and dermoscopic image examination. Smaller facial skin lesions can be challenging even for experienced observers. No larger study concerning the sensitivity and specificity for diagnosing benign and malign facial skin lesions has been published yet.

160 facial skin lesions were imaged in totally 148 patients, 98 (61.25 %) lesions were malignant and 62 (38.75 %) were benign. 87 (54.38 %) skin lesions were melanocytic and 73 (45.63 %) were non-melanocytic. Previously verified histopathological diagnoses was available in 128 cases (80 %)

Three data sets were acquired in three consecutive steps; in the first step only RCM images were evaluated by the observers (two experts and one novice in RCM) who were all blinded to the diagnoses of the skin lesions. In the second step the clinical and dermoscopic images were assessed together by the same blinded observers. In the third step all images were evaluated together. The single novice observer performed two courses of image evaluation (each with three steps) to evaluate the learning curve for RCM image analysis. The time gap between the two courses was six months.

For RCM image analysis, mosaics were first assessed concerning the overall lesion appearance (symmetry / asymmetry) and then for specific RCM criteria at various imaging depths ranging from stratum corneum to the upper dermis.

Our study clearly demonstrated three findings:

1.) Improvement of diagnostic accuracy:

Our data showed an improvement both in sensitivity and in specificity if the RCM images were evaluated adjunct to dermoscopic and clinical images. This fact is best shown by the two expert observers: Using only the RCM images proofed a low sensitivity of 57 % but a very high specificity of 92 %. When dermoscopic and clinical images were added the sensitivity climbed to 71 % with a persisting specificity of 93 %.

This finding has to be compared with the first step, where only clinical and dermoscopic images were evaluated the sensitivity dropped to 67 % while the specificity was high at 92 %.

2.) Change of Management Procedure:

RCM has been described has a helpful tool facilitating optimal patient treatment and care for both melanocytic (28) and non-melanocytic skin cancers. In our study recommended procedures included excision, follow- up and none; meaning that no further measures were needed. In the novice observer's first course far less excisions were recommended than actually performed (clinical and dermoscopic images: 47% / only RCM: 46 % / clinical, RCM and dermoscopic images: 40 % versus 80 % actually performed). An explanation for this astonishing finding could be the low arithmetic sensitivity of all diagnoses of 54 % for clinical and dermoscopic images, 50 % for only RCM, and 57 % for clinical, RCM and dermoscopic images respectively. In the novice observer's second course the overall arithmetic sensitivity was quite higher (clinical and dermoscopic images: 61 %, only RCM: 62 % and clinical, RCM and dermoscopic images: 69 %) but recommended excision was still lower (clinical and dermoscopic images: 56 %, only RCM: 55 %, clinical, RCM and dermoscopic images: 40 %) than actually performed (80 %). The expert observers demonstrated a higher overall arithmetic sensitivity (67 %, 58 %, and 72 %) as well as higher rates for recommended excision (82 % / 60 % / 76 % versus 80 % actually performed).

3.) Learning curve of novice observer:

An interesting finding was depicted in the single novice observer's second course of image evaluation: The overall arithmetic sensitivities (61 % / 62 % / 69 %) improved in all three steps compared with the first course (sensitivity: 54 % / 50 % / 57 %) while the specificities remained high. (First course: 87 % / 84 % / 84 % second course: 79 % / 85 % / 89 %).

However; the expert observers still proofed a slightly higher sensitivity (67 % / 58 % / 72 %) and specifity (92% / 92 % / 93 %). As seen in the novice observer's course the highest sensitivity was obtained by the two expert observers as well as by using clinical and dermoscopic images adjunct to the RCM images.

4.) Melanoma detection:

A two-step scoring algorithm for detecting malignant melanoma was proposed by Segura et al. (18) in 2009 and reevaluated by Carrera et al. in 2012 (29). However the examination of our detail RCM criteria showed no prevalence for any detail RCM criteria at any skin layer depth. A large range for any detail criterion could be found between the single novice observer's first and the second course of image evaluation. The expert observers' course showed the same bandwidth of range, thus representing a high intra- as well as a high interobserver variability. Absolute numbers and percentages between different courses from the novice observer were ranging from 1 (1.41%) to 20 (28.17%) cases for a single criterion in the same step.

A study published by Gerger et al. (30) described a high sensitivity and specificity (97 % and 83 % respectively) comparable to a study published by Longo et al. (31) demonstrating a sensitivity of 96 % and a specificity of 94 %. Another study published in 2006 by Gerger et al. (32) showing a sensitivity of 96 %. However a study involving 216 melanomas by Guitera et al. (33) proved a lower sensitivity and specificity (87 % and 70 % respectively). All these studies included melanomas on all body sites. In our study only the expert observers' course showed comparable results (sensitivity: 76 % and specificity: 91%) when using clinical and dermoscopic images adjunct to RCM images. Results from the single novice observer were far minor demonstrating a poor sensitivity (35 % / 41 % / 55%) in the first course but a high specificity (88 % / 89 % / 89 %). In the novice observer's second course sensitivities (63 % / 69 % / 68 %) and specificities (79 % / 83 % / 84 %) were higher than in the first course but still inferior to the expert observers. These findings underline the importance of training in the evaluation of RCM and dermoscopic images.

5.) Basalioma detection:

In a large study published by Guitera et al. (33) containing 119 BCCs; BCCs were detected by RCM with an astonishing sensitivity of 100 % and a very high specificity of 88%. In our study a high diagnostic accuracy for BCC with a good sensitivity (33 % / 60 % / 73 %) and specificity (94 %/ 98 % / 96 %) was seen even in the novice observer's first course.

In the novice observer's second course the sensitivity (47 % / 60 % / 67 %) improved while specificity (97 % / 98 % / 98 %) remained high. These results are similar to those seen in the expert observers' course depicting a good sensitivity (60% / 67% / 67%) and specificity (99 % / 97 % / 99 %).

Several detail RCM criteria have been mentioned as pathognomonic for BCC in several different studies (33) (34) (35) however the typical "Streaming" could be found by all observers in all steps in only 6 (40.00%) to 10 (66.67%) cases. The characteristically tumor islands have been described has critical tool to define operative margins (36) however tumor islands were only seen in 5 (33.33%) to 10 (66.67%) cases and larger dendritic structures within tumor islands were mentioned in 6 (40.00%) to 8 (53.33%) cases.

6.) Nevus detection

Nevus has been reported by Pellacani et al. (37) as marked by a ringed pattern and melanocytic nests. In our study nevus has been seen as a more symmetrical lesion with regular skin folds, typical honeycomb pattern and melanocytic nests at the spinous granular layer. This fact probably is due the special histological anatomy of facial skin. Interestingly all other detail criteria were mentioned in up to 7 (43.75%) cases in the single novice observer's first course, while in the second course absolute number were only mentioned in 4 (25.00%) cases, similar to those mentioned by the two expert observers. This fact demonstrates a high learning curve for recognizing detail criteria concerning the diagnosis of benign nevus.

Data derived from the single novice observer's second course concerning detail criteria for nevus showed a curious preference for typical honeycomb pattern and melanocytic nests. The number of nevi in this study is limited (n = 16) so to analyze this fact a new study with a higher number of nevi and newly defined detail criteria is required to prove the exact percentages of detail RCM criteria for diagnosing nevi.

Conclusion:

Our study has proven RCM as a helpful method for increasing both the specificity and the sensitivity for diagnosing both malignant and benign facial lesions. RCM alone is inferior to clinical and dermoscopic images however using clinical, dermoscopic and RCM images together clearly improves the sensitivity and the specificity. The single novice observer's second course of image evaluation showed a far better sensitivity and only a slightly minor specificity thus representing a short learning curve.

References:

1 Rajadhyaksha M, Grossman M, Esterowitz D, Webb RH, Anderson RR. In vivo confocal scanning laser microscopy of human skin: melanin provides strong contrast. J Invest Dermatol. 1995 Jun;104(6):946-52.

2 Rajadhyaksha M, González S, Zavislan JM, Anderson RR, Webb RR. In vivo confocal scanning laser microscopy of human skin II: Advances in instrumentation and comparison with histology. J Invest Dermatol. 1999; 113: 293– 303.

3 Huzaira M, Rius F, Rajadhyaksha M, Anderson RR, González S. Topographic variations in normal skin, as viewed by in vivo reflectance confocal microscopy. J Invest Dermatol. 2001; 116: 846– 852.

4 Yamashita T, Kuwahara T, González S, Takahashi M. Non-invasive visualization of melanin and melanocytes by reflectance-mode confocal microscopy . J Invest Dermatol. 2005; 124: 235– 24.

5 Braga JC, Scope A, Klaz I, Mecca P, González S, Rabinovitz H, et al. The significance of reflectance confocal microscopy in the assessment of solitary pink skin lesions. J Am Acad Dermatol. 2009; 61 (2): 230– 241.

6 Roewert-Huber J, Stockfleth E, Kerl H. Pathology and pathobiology of actinic (solar) keratosis - an update Br J Dermatol. 2007 Dec;157 Suppl 2:18-20.

7 Ulrich M, Maltusch A, Rius-Diaz F, Röwert-Huber J, González S, Sterry W, et al. Clinical applicability of in vivo reflectance confocal microscopy for the diagnosis of actinic keratoses. Dermatol Surg. 2008; 34: 610– 619.

8 Aghassi D, Anderson RR, González S. Confocal laser microscopic imaging of actinic keratoses in vivo: a preliminary report. J Am Acad Dermatol. 2000; 43: 42– 48.

9 González S, Tannous Z. Real-time, in vivo confocal reflectance microscopy of basal cell carcinoma. J Am Acad Dermatol. 2002; 47: 869– 874.

10 Ahlgrimm-Siess V, Cao T, Oliviero M, Hofmann-Wellenhof R, Rabinovitz H, Scope A. The Vasculature of Nonmelanocytic Skin Tumors in Reflectance Confocal Microscopy: Vascular Features of Basal Cell Carcinoma. Arch Dermatol. 2010 Mar; 146(3): 353-4.

11 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. J Am Acad Dermatol. 2004; 51: 923–930.

12 Marra DE, Torres A, Schanbacher CF, Gonzalez S. Detection of residual basal cell carcinoma by in vivo confocal microscopy. Dermatol Surg. 2005 May; 31(5):538-41.

13 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. J Am Acad Dermatol. 2001 Sep;45(3):365-76.

14 Ahlgrimm-Siess V, Massone C, Koller S, Fink-Puches R, Richtig E, Wolf I. In vivo confocal scanning laser microscopy of common naevi with globular, homogeneous and reticular pattern in dermoscopy. Br J Dermatol. 2008 May;158(5):1000-7. Epub 2008 Feb 18.

15 Scope A, Benvenuto-Andrade C, Agero AL, Malvehy J, Puig S, Rajadhyaksha M, et al. In vivo reflectance confocal microscopy imaging of melanocytic skin lesions: consensus terminology glossary and illustrative images. J Am Acad Dermatol. 2007; 57 (4): 644– 658.

16 Pellacani G, Cesinaro AM, Seidenari S. In vivo assessment of melanocytic nests in nevi and melanomas by reflectance confocal microscopy. Mod Pathol. 2005 Apr;18(4):469-74.

17 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. J Invest Dermatol. 2007; 127: 2759– 2765.

18 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. J Am Acad Dermatol. 2009 Aug;61(2):216-29. Epub 2009 Apr 29.

19 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. Arch Dermatol. 2008 Dec;144(12):1597-608.

20 Lorber A, Wiltgen M, Hofmann-Wellenhof R, Koller S, Weger W, Ahlgrimm-Siess V. Correlation of image analysis features and visual morphology in melanocytic skin tumours using in vivo confocal laser scanning microscopy. Skin Res Technol. 2009 May;15(2):237-41.

21 Pellacani G, Cesinaro AM, Seidenari S. Reflectance-mode confocal microscopy for the in vivo characterization of pagetoid melanocytosis in melanomas and nevi. J Invest Dermatol. 2005; 125: 532– 537.

22 Gerami P, Barnhill RL, Beilfuss BA, LeBoit P, Schneider P, Guitart J. Superficial melanocytic neoplasms with pagetoid melanocytosis: a study of interobserver concordance and correlation with FISH. Am J Surg Pathol. 2010 Jun;34(6):816-21.

23 Tannous ZS, Mihm MC, Flotte TJ, González S. In vivo examination of lentigo maligna and malignant melanoma in situ, lentigo maligna type by near-infrared reflectance confocal microscopy: comparison of in vivo confocal images with histologic sections. J Am Acad Dermatol. 2002 Feb;46(2):260-3.

24 Ahlgrimm-Siess V, Massone C, Scope A, Fink-Puches R, Richtig E, Wolf IH Reflectance confocal microscopy of facial lentigo maligna and lentigo maligna melanoma: a preliminary study. Br J Dermatol. 2009 Dec;161(6):1307-16. Epub 2009 May 5.

25 Guitera P, Pellacani G, Crotty KA, Scolyer RA, Li LX, 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. J Invest Dermatol. 2010 Aug;130(8):2080-91. Epub 2010 Apr 15.

26 Ahlgrimm-Siess V, Cao T, Oliviero M, Laimer M, Hofmann-Wellenhof R, Rabinovitz HS, Scope A. Seborrheic keratosis: Reflectance confocal microscopy features and correlation with dermoscopy. J Am Acad Dermatol. 2013 Feb (epub ahead of print)

27 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. J Am Acad Dermatol. 2006 Jul;55(1):88-97.

28 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 Dermatol. 2013 Apr 3:1-7. [Epub ahead of print]

29 Carrera C, Puig S, Malvehy J. In vivo confocal reflectance microscopy in melanoma. Dermatol Therapy. 2012 Vol. 25 410-422

30 Gerger A, Hofmann-Wellenhof R, Samonigg H, Smolle J. In vivo confocal laser scanning microscopy in the diagnosis of melanocytic skin tumours. Br J Dermatol. 2009 Mar;160(3):475-81Epub 2009 Jan 10. Review.

31 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 on 140 cases. Br J Dermatol. 2013 Feb 3. [Epub ahead of print]

32 Gerger A, Koller S, Weger W, Richtig E, Kerl H, Samonigg H, et al. Sensitivity and specificity of confocal laser-scanning microscopy for in vivo diagnosis of malignant skin tumors. Cancer. 2006 Jul 1;107(1):193-200.

33 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. J Invest Dermatol. 2012 Oct;132(10):2386-94. Epub 2012 Jun 21.

34 Seguira S. Puig S. Carrera C, Lecha M, Borges V, Malvehy J. Non-invasive management of non-melanoma skin cancer in patients with cancer predisposition

genodermatosis: a role for confocal microscopy and photodynamic therapy. J Eur Acad Dermatol Venereol. 2011 Jul;25(7):819-27. Epub 2010 Oct 15.

35 Manfredini M, Arginelli F, Dunsby C, French P, Talbot C, König K et al. High-resolution imaging of basal cell carcinoma: a comparison between multiphoton microscopy with fluorescence lifetime imaging and reflectance confocal microscopy. Skin Res Technol. 2013 Feb;19(1):e433-43.

36 Pan ZY, Lin JR, Cheng TT, Wu JQ, Wu WY. In vivo reflectance confocal microscopy of Basal cell carcinoma: feasibility of preoperative mapping of cancer margins. Dermatol Surg. 2012 Dec;38(12):1945-50.

37 Pellacani G, Scope A, Farnetani F, Casaretta G, Zalaudek I, Moscarella E et al. Towards an in vivo morphologic classification of melanocytic nevi. J Eur Acad Dermatol Venereol. 2013 May 10. [Epub ahead of print]