In Vivo Reflectance Confocal Microscopy as a Response Monitoring Tool for Actinic Keratoses Undergoing Cryotherapy and Photodynamic Therapy

Clara Curiel-Lewandrowski 1,2,*, Caitlyn N. Myrdal 1, Kathylynn Saboda 2, Chengcheng Hu 3, Edith Arzberger 4, Giovanni Pellacani 5, Franz Josef Legat 4, Martina Ulrich 6, Petra Hochfellner 4, Margaret C. Oliviero 7, Paola Pasquali 8,9,*, Melissa Gill 9,10 and Rainer Hofmann-Wellenhof 4

Abstract: Reflectance confocal microscopy (RCM) presents a non-invasive method to image actinic keratoses (AK) at a cellular level. However, RCM criteria for AK response monitoring vary across studies and a universal, standardized approach is lacking. We aimed to identify reliable AK response criteria and to compare the clinical and RCM evaluation of responses across AK severity grades. Twenty patients were included and randomized to receive either cryotherapy (n = 10) or PDT (n = 10). Clinical assessment and RCM evaluation of 12 criteria were performed in AK lesions and photodamaged skin at baseline, 3 and 6 months. We identified the RCM criteria that reliably characterize AK at baseline and display significant reduction following treatment. Those with the highest baseline odds ratio (OR), good interobserver agreement, and most significant change over time were atypical honeycomb pattern (OR: 12.7, CI: 5.7–28.1), hyperkeratosis (OR: 13.6, CI: 5.3–34.9), stratum corneum disruption (OR: 7.8, CI: 3.5–17.3), and disarranged epidermal pattern (OR: 6.5, CI: 2.9–14.8). Clinical evaluation demonstrated a significant treatment response without relapse. However, in grade 2 AK, 10/12 RCM parameters increased from 3 to 6 months, which suggested early subclinical recurrence detection by RCM. Incorporating standardized RCM protocols for the assessment of AK may enable a more meaningful comparison across clinical trials, while allowing for the early detection of relapses and evaluation of biological responses to therapy over time.
Keywords: reflectance confocal microscopy; actinic keratosis; response monitoring; cryotherapy; photodynamic therapy; RCM

1. Introduction

Actinic keratoses are common skin lesions with the risk to evolve into squamous cell carcinoma (SCC). It is currently not possible to predict which AK lesions will progress and various sources differ on the potential rate of progression. A landmark paper from Callen et al. approximated a 0.25–20% annual transformation rate for a single AK [1]. Cockerell estimated that the 10-year neoplastic transformation rate for the average person with 7 AKs is between 10–20% [2]. The diagnosis of AK is typically made via clinical evaluation and those that appear suspicious for SCC are biopsied for histopathologic confirmation.

Treatment options for AK are abundant and include photodynamic therapy (PDT), cryotherapy, 5-fluorouracil, diclofenac gel, imiquimod, and chemical peels, among others. Cryotherapy represents a lesion-directed therapy while PDT is a field-directed therapy that requires a photosensitizer, light, and oxygen. Photosensitizing agents include topical methyl aminolevulinate (MAL), 5-aminolevulinic acid (ALA), or BF-200 ALA, a nanoemulsion-based gel that optimizes epidermal penetration [3]. While light sources typically include red or blue light, daylight PDT has also been demonstrated to be effective [4,5].

To date, preventive and therapeutic approaches targeting AK have primarily relied on clinical and histological assessment to measure the effectiveness of the intervention used. Both approaches have significant limitations, including the low sensitivity and subjectivity of clinical assessment and the invasiveness associated with skin biopsy procedures, which hampers the possibility for the longitudinal monitoring of individual lesion response.

Reflectance confocal microscopy is a non-invasive method to image the skin at the cellular level using an in vivo approach and allows for a resolution similar to that of histology. RCM can be used to diagnose AK with 80% sensitivity and 98% specificity [6]. Furthermore, it has been suggested that RCM can detect early morphological changes and cellular atypia before an AK is clinically apparent [7–9]. Hence, due to its non-invasive nature, RCM is advantageous as subclinical AK can be detected and monitored over a period of time with relative ease, and little discomfort for the patient. Pellacani et al. demonstrated good interobserver correlation and high concordance between RCM and histological grading of keratinocyte atypia, highlighting that RCM can be as effective as a histopathologic examination [10].

Studies have also examined the use of RCM technology to follow the outcomes of therapeutic interventions in AK. In 2018, Ishioka et al. described a reduction in the RCM parameters of nucleated polygonal cells and isolated keratinocytes with 5% fluorouracil treatment [11]. Pasquali et al. monitored AK treatment with sequential cryotherapy and ingenol mebutate, identifying an atypical honeycomb pattern, round papillary vessels, parakeratosis, and stratum corneum detachment as reliable RCM parameters for response characterization [12]. In 2019, Benati et al. identified a significant improvement of keratinocyte disarray, parakeratosis, and crust following topical imiquimod treatment [13]. The above studies all concluded that RCM is effective for AK response monitoring; however, methods of evaluating responses varied considerably. Inconsistency of reliable RCM parameter identification is also seen in non-therapeutic AK characterization studies at baseline [14–18]. The establishment of a standard, universally accepted protocol for the evaluation of AK treatment responses by RCM could greatly equalize results between trials and add clinical benefit to surveillance.

The primary goal of our study was to contribute to this need by examining RCM criteria for AK and identifying those that display a significant response to treatment, good interobserver agreement, and high prevalence at baseline when compared to nearby photodamaged (PD) skin. Our secondary aims were to compare RCM and the clinical
evaluation of responses, evaluate differences across AK grades, and assess cryotherapy and PDT treatments of AK as measured by RCM. Unique contributions to the current state of the field include the incorporation of nearby photodamaged (PD) skin as a reference for baseline assessment, comparison of RCM parameters across AK grades, and utilization of three experienced confocal evaluators to assess interobserver agreement.

2. Materials and Methods

2.1. Study Population

This single-center prospective study included patients recruited from the Dermatology Department at the Medical University of Graz, Austria. A total of 20 healthy patients aged 45–85 years with Fitzpatrick skin types I-II and a clinical diagnosis of two AKs (3–6 mm in diameter, grades 1–2) on the forehead, scalp, hands, or forearms were included. All participants signed an informed consent form, and the study was approved by the institutional review board and conducted in accordance with the Declaration of Helsinki principles.

2.2. Study Design

Patients were randomly assigned to receive either treatment with cryotherapy (n = 10) or photodynamic therapy (PDT, n = 10). Three study skin sites were selected for evaluation: 1. Grade 1 AK (G1); 2. Grade 2 AK (G2); 3. Photodamaged skin (PD) at a separate location and in close proximity to the AKs. Clinical grading was completed through assessment of hyperkeratosis from 0 (absent) to 3 (3+), based on the Olsen et al. grading criteria [19]. Lesions were graded based on palpability: 0 = not palpable when compared to surrounding skin, 1 = slight palpability (the AK is felt better than seen), 2 = moderate palpability (the AK is seen and felt), and 3 = severe (the AK easily is visualized and felt). For more granular clinical scoring, measures of erythema and hyperkeratosis were scored with half-points ranging from 0 to 3 to characterize lesions more accurately. For instance, 0.5 = trace palpability (lesion cannot be seen visually and is minimally palpable) and 2.5 = moderate palpability but not considered severe. Measures of erythema ranged from 0 = none, 0.5 = minimal, 1 = mild, 2 = moderate, 3 = severe. Clinical scoring was performed by one investigator, CCL. Skin sites were photographed at each visit by standard protocol, which included regional photography followed by close-up and dermoscopy images. In addition, detailed measurements (3 per lesion) indicating the location of the study lesions were obtained. Individuals agreed to limit significant sun exposure to study areas at least 15 days prior to a scheduled imaging session and agreed to wear protective clothing when outdoors.

Study skin sites were evaluated clinically, by RCM, and by the standard imaging protocol at baseline, 3 and 6 months post-treatment. For both groups, the first visit consisted of baseline photographic documentation and RCM followed by treatment administration to AKs.

RCM imaging was performed using a confocal microscope (Vivascope 1500–Mavig) for which details of this technique have been described previously [20,21]. Four horizontally mapped mosaic images (VivaBlocks) were obtained at representative levels of the stratum corneum, mid epidermis, dermoepidermal junction and papillary dermis. Only areas without artifacts within the mosaics were scored. Two vertically mapped images (VivaStacks) were obtained per lesion, one each from the left upper and right lower quadrants of the field of view captured with RCM, with regular step intervals from the surface of the stratum corneum to the dermis. Stratum corneum thickness was calculated by counting the number of step intervals from the top of the stratum corneum to the top of the stratum granulosum. Stratum corneum thickness >20 µm on the face and >40 µm on other sites was considered hyperkeratosis [22].
2.3. Treatment Protocol

2.3.1. Cryotherapy

AKs were treated using a Brymill’s Cry-Ac® 500 cc hand-held device (Brymill Cryogenic Systems, Ellington, CT, USA). Liquid nitrogen was delivered through a C aperture (Model #102-C, 0.022in.). The surface temperature was measured using an IR (Infrared) thermometer and each lesion was frozen until $-22$ to $-25$ $^\circ$C was obtained. Spraying was perpendicular to the skin and 5 cm from the skin surface. The depth of freezing front was measured using 22 MHz high-frequency ultrasound immediately after freezing to conform that the ice block had reached the dermoepidermal junction.

2.3.2. PDT

Topical methyl aminolevulinate (MAL) pre-light photosensitizer was administered to AK lesions after removal of the superficial crust. Three hours later, the area was irradiated with a red-light source (light dose 37 J/cm$^2$ from a medical lamp Aktilite® CL128 (Galderma Laboritories, Lausanne, Switzerland), emitting heat-free, visible red light at a peak wavelength of 630 nm; irradiation time was 8 min) to activate the photosensitizer. The PD skin sites were not directly treated in either group.

2.4. Reflectance Confocal Microscopy Evaluation

RCM analysis included the scoring of 12 diagnostic criteria for AK based on previous work which examined features through RCM [8,16]. All parameters were scored as either “absent” or “present” except for presence of round blood vessels, polymorphous blood vessels, inflammatory infiltrate dermis (0 = absent, 1 = 1+, 2 = 2+, 3 = 3+), and atypical honeycomb pattern (0 = absent, 1 = mild, 2 = moderate, 3 = severe). To assess baseline interobserver variability, RCM readers (GP, CCL, RHW) were asked to score images from a subset of 10 study subjects at baseline through which a preliminary proportion of agreement was calculated for each criterion. After the consensus meetings, the drafting of a reference document incorporating consensus definitions, full subject enrollment, and a washout period of approximately 4 months, the RCM analysis was completed on all cases using the reference document. Interobserver proportion of agreement was calculated again. Study RCM image analysis was completed by three experienced RCM evaluators (cryotherapy group: GP, CCL/EA, MU; PDT group: GP, CCL/EA, MO). At the Graz study site, images were read together by CCL and EA.

2.5. Statistical Methods

Dichotomous RCM parameters were compared using logistic regression with GEE to adjust for intrasubject correlations. Ordinal RCM parameters were compared using ordinal logistic regression, with clustering within subjects to adjust for intrasubject correlations. Clinical features were scored on a semi-quantitative scale from 0–3. For comparisons of clinical features, Kruskal-Wallis tests were used. The percentage of agreement is presented with binomial 95% confidence intervals. Analyses were performed using Stata 16 (StataCorp, College Station, TX, USA).

3. Results

3.1. Patient Population

Twenty patients met the criteria and were randomized into the study. The mean age was 74 (SD 6.7) and 18 (90%) of the patients were male. Patients were recruited into the Dermatology Department at the Medical University of Graz, Austria; 19 had Fitzpatrick skin type II (95%), and one had Fitzpatrick skin type I (5%). A total of 40 AKs were evaluated and treated, one G1 AK and one G2 AK per individual. Fourteen AKs were located on the forehead (35%), thirteen elsewhere on the face (32.5%), seven on the scalp (17.5%), and six on the dorsal hand and forearm (15%).
3.2. Clinical Assessment of AKs

Clinically, AKs in both treatment groups demonstrated a significant decrease in the parameters of hyperkeratosis and erythema at 3 and 6 months (Table 1). In the cryotherapy group, treatment outcome at 6 months was evaluated clinically as a complete response in 8/10 G1 AKs (80%) and 8/10 G2 AKs (80%). A partial response at 6 months was seen in 2/10 G1 AKs (20%) and 2/10 G2 AKs (20%). In the PDT group, clinical evaluation at 6 months was scored as a complete response in 6/9 G1 AKs (67%), and 5/9 G2 AKs (56%). A partial response was seen in 3/9 G1 AKs (33%) and 4/9 G2 AKs (44%). Clinical grading scores were not available for one patient in the PDT group.

Table 1. Clinical scoring based on measures of erythema and hyperkeratosis, with green indicating complete response at 6 months, orange indicating partial response at 6 months, and pink indicating clinical progression.

| Subject ID | Baseline Erythema | 3 Mo Erythema | 6 Mo Erythema | Baseline Hyperkeratosis | 3 Mo Hyperkeratosis | 6 Mo Hyperkeratosis |
|------------|-------------------|---------------|---------------|-------------------------|-------------------|-------------------|
| C1         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C2         | 1                 | 0             | 1             | 1                       | 0.5               | 0                 |
| C3         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C4         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C5         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C6         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C7         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| C8         | 1                 | 1             | 0             | 1                       | 1                 | 0                 |
| C9         | 1                 | 1             | 0             | 1                       | 0                 | 0                 |
| C10        | 1                 | 0             | 0             | 0.5                     | 1                 | 0                 |

| Subject ID | Baseline Erythema | 3 Mo Erythema | 6 Mo Erythema | Baseline Hyperkeratosis | 3 Mo Hyperkeratosis | 6 Mo Hyperkeratosis |
|------------|-------------------|---------------|---------------|-------------------------|-------------------|-------------------|
| P1         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| P2         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| P3         | 1                 | 0             | 1             | 0                       | 0                 | 0                 |
| P4         | 1                 | 0             | 0             | 1                       | 0.5               | 1                 |
| P5         | 1                 | 0.5           | 0.5           | 1                       | 0                 | 0                 |
| P6         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| P7         | 1                 | -             | -             | 1                       | -                 | -                 |
| P8         | 1                 | 0             | 0             | 1                       | 0.5               | 2                 |
| P9         | 1                 | 0             | 0             | 1                       | 0                 | 0                 |
| P10        | 1                 | 0             | 0             | 1                       | 0.5               | 0                 |

1 Per protocol, G1 AK had baseline clinical score of “1” for erythema and hyperkeratosis, and G2 AK had a baseline clinical score of “2” for erythema and hyperkeratosis. Clinical scoring was performed on a scale of 0 (absent) to 3 (3+). Subject P7 did not receive clinical scoring.

3.3. RCM Interobserver Agreement

Table 2 displays the RCM interobserver agreement. Prior to the start of this study and before the criteria standardization, the proportion of agreement between three experienced RCM readers ranged from 37 (CI: 24–52) to 87 (CI: 80–92). Standardization efforts were achieved through consensus meetings, in which criteria definitions and examples were discussed to implement consistency among readers. Post-standardization effort, the proportion of agreement between three experienced RCM readers, ranged from 61 (CI: 57–65) to 86 (CI: 83–88). Interobserver agreement improvements for stratum corneum disruption, round nucleated cells, disarranged epidermal pattern, and polymorphous blood vessels reached statistical significance following the standardization efforts. Parakeratosis and atypical honeycomb patterns approached significance. Interobserver agreement was excellent before and after standardization for atypical honeycomb pattern, intraepidermal dendritic cells, and dermal inflammatory cell infiltrate.
Table 2. Interobserver agreement of RCM parameters for AK.

| Skin Layer                | RCM Parameter                      | Pre-Standardization Proportion of Agreement | Post-Standardization Proportion of Agreement |
|---------------------------|------------------------------------|---------------------------------------------|---------------------------------------------|
|                           |                                    | (95% CI)                                    | (95% CI)                                    |
| Stratum Corneum           | Parakeratosis                      | 54 (45–63)                                 | 65 (60–69)                                 |
|                           | Hyperkeratosis                     | 63 (53–71)                                 | 63 (58–68)                                 |
|                           | Stratum Corneum Disruption         | 37 (24–52)                                 | 61 (57–65) **                             |
|                           | Atypical Honeycomb Pattern         | 80 (68–88)                                 | 86 (83–88)                                 |
|                           | Round Nucleated Cells              | 52 (42–61)                                 | 79 (76–82) **                             |
|                           | Disarranged Epidermal Pattern      | 37 (29–47)                                 | 63 (59–66) **                             |
|                           | Presence of Inflammatory Cells     | 61 (55–67)                                 | 62 (58–66)                                 |
|                           | Presence of Dendritic Cells        | 87 (80–92)                                 | 80 (77–83)                                 |
| Epidermis                 | Presence of Inflammatory Cells     | 85 (82–89)                                 | 83 (80–86)                                 |
|                           | Solar Elastosis                     | N/A                                         | 73 (69–76)                                 |
|                           | Round Blood Vessels                | 69 (57–78)                                 | 64 (60–67)                                 |
|                           | Polymorphous Blood Vessels         | 53 (74–65)                                 | 81 (78–84) **                             |

1 Pre-standardization time-point: prior to any RCM consensus meetings or standardization (pre-study). Pre-standardization readers included CCL, RHW, and GP. Readers for post-standardization proportion of agreement, following subject RCM analysis, included GP, CCL/EA*, MU (cryotherapy) and GP, CCL/EA*, MO (PDT). *CCL/EA read images together. **Non-dichotomous variables, agreement within 1 point. ** Indicates a statistically significant improvement in interobserver agreement.

3.4. Baseline RCM Scoring

Twelve RCM criteria were evaluated per lesion. A comparison between AKs and adjacent PD skin yielded statistical significances in every parameter except for the presence of dendritic cells and round nucleated cells (Table 3). RCM criteria demonstrating the highest, most significant OR included hyperkeratosis (OR: 13.6, CI: 5.3–34.9), atypical honeycomb pattern (OR: 12.7, CI: 5.7–28.1), stratum corneum disruption (OR: 7.8, CI: 3.5–17.3), and disarranged epidermal pattern (OR: 6.5, CI: 2.9–14.8).

Table 3. Odds ratios comparing RCM parameters in AK to PD skin at baseline.

| Skin Layer                | RCM Parameter                              | Grade 1 AK (95% CI) | Grade 2 AK (95% CI) | All AK (95% CI) |
|---------------------------|--------------------------------------------|---------------------|---------------------|-----------------|
| Stratum Corneum           | Parakeratosis                              | 5.5 ** (2.6–11.7)  | 5.3 ** (2.3–11.9)  | 5.3 ** (2.6–10.6) |
|                           | Hyperkeratosis                             | 10.6 ** (3.8–29.0) | 18.3 ** (6.1–54.5) | 13.6 ** (5.3–34.9) |
|                           | Stratum Corneum Disruption                 | 7.1 ** (2.9–17.4)  | 8.9 ** (3.5–22.8)  | 7.8 ** (3.5–17.3) |
|                           | Atypical Honeycomb Pattern                 | 9.7 ** (3.8–24.5)  | 18.0 ** (7.7–41.8) | 12.7 ** (5.7–28.1) |
| Epidermis                 | Round Nucleated Cells                      | 3.0 * (1.01–8.9)   | 1.6 (0.5–5.1)      | 2.3 (0.8–6.5)   |
|                           | Disarranged Epidermal Pattern              | 5.4 ** (2.2–12.3)  | 8.3 ** (3.2–21.5)  | 6.5 ** (2.9–14.8) |
|                           | Presence of Inflammatory Epidermal Cells   | 1.9 (0.9–4.1)      | 2.4 * (1.1–5.2)    | 2.1 * (1.1–4.0) |
|                           | Presence of Dendritic Cells                | 0.8 (0.2–2.6)      | 1.1 (0.3–3.5)      | 0.9 (0.4–2.6)   |
### Table 3. Cont.

| Skin Layer | RCM Parameter | Grade 1 AK (95% CI) | Grade 2 AK (95% CI) | All AK (95% CI) |
|------------|---------------|---------------------|---------------------|----------------|
| Dermis     | Inflammatory Infiltrate | 3.6 ** (1.5–8.6) | 2.2 (0.8–5.8) | 2.8 * (1.2–6.5) |
|           | Solar Elastosis | 3.5 (0.9–12.5) | 1.6 (0.6–4.4) | 2.2 * (0.9–5.5) |
|           | Round Blood Vessels | 1.5 (0.8–2.7) | 1.8 * (1.02–3.2) | 1.6 * (1.0–2.6) |
|           | Polymorphous Blood Vessels | 3.8 ** (1.6–8.7) | 3.0 * (1.1–7.7) | 3.3 ** (1.5–7.4) |

Scores from RCM evaluators are averaged to obtain odds ratios. Odds ratios are presented for AK lesions compared to PD skin, *p* < 0.05, **p < 0.01. 1 For non-dichotomous variables, odds ratios were calculated using ordinal logistical regression.

3.5. RCM Evaluation over Time

Table 4 shows RCM scores at 3 and 6 months following cryotherapy or PDT treatment.

### Table 4. Mean (standard deviation) clinical scores over time and RCM parameter prevalence over time in cryotherapy and PDT groups.

| Treatment | Clinical Parameter | Grade 1 AK | Grade 2 AK | PD Skin |
|-----------|--------------------|------------|------------|---------|
|           | Erythema           |            |            |         |
| Cryo      |                    | 1.00 1     | 0.20 (0.42) | 0.15 (0.34) | 2.00 1 |
| PDT       |                    | 1.00 1     | 0.55 (0.17) | 0.17 (0.35) | 2.00 1 |
|           | Hyperkeratosis     |            |            |         |
| Cryo      |                    | 1.00 1     | 0.35 (0.47) | 0.05 (0.16) | 2.00 1 |
| PDT       |                    | 1.00 1     | 0.33 (0.43) | 0.05 (0.17) | 2.00 1 |
|           | Parakeratosis      |            |            |         |
| Cryo      |                    | 85%        | 43% **     | 57% *     | 58% |
| PDT       |                    | 52%        | 29%        | 29%       | 79% |
|           | Hyperkeratosis     |            |            |         |
| Cryo      |                    | 71%        | 45% *      | 23% **    | 80% |
| PDT       |                    | 45%        | 22%        | 20%       | 63% |
|           | Stratum Corneum Disruption | | | | |
| Cryo      |                    | 88%        | 47% **     | 40% **    | 83% |
| PDT       |                    | 56%        | 39%        | 12% *     | 69% |
|           | Atypical Honeycomb Pattern | | | | |
| Cryo      |                    | 100%       | 67% *      | 63% **    | 96% |
| PDT       |                    | 97%        | 63% *      | 53% **    | 100% |
|           | Round Nucleated Cells | | | | |
| Cryo      |                    | 23%        | 23%        | 7%        | 13% |
| PDT       |                    | 23%        | 4%         | 6%        | 15% |
|           | Disarranged Pattern | | | | |
| Cryo      |                    | 48%        | 23% *      | 13% **    | 59% |
| PDT       |                    | 53%        | 16% *      | 25%       | 62% |
|           | Inflammatory Cells Present | | | | |
| Cryo      |                    | 61%        | 40%        | 28% *     | 43% |
| PDT       |                    | 50%        | 22%        | 22%       | 74% |
|           | Dendritic Cells Present | | | | |
| Cryo      |                    | 18%        | 10%        | 17%       | 13% |
| PDT       |                    | 4%         | 8%         | 0%        | 15% |
|           | Inflammatory Cells Present | | | | |
| Cryo      |                    | 61%        | 54%        | 27% *     | 39% |
| PDT       |                    | 65%        | 48% **     | 29% *     | 54% |
|           | Solar Elastosis     | | | | |
| Cryo      |                    | 96%        | 79%        | 90%       | 81% |
| PDT       |                    | 89%        | 96%        | 100%      | 88% |
|           | Round Blood Vessels | | | | |
| Cryo      |                    | 67%        | 79%        | 77%       | 64% |
| PDT       |                    | 72%        | 85% *      | 94% *     | 86% |
|           | Polymorphous Blood Vessels | | | | |
| Cryo      |                    | 44%        | 31%        | 13%       | 26% |
| PDT       |                    | 38%        | 23%        | 23%       | 43% |

Significance compared to baseline (BL), *p < 0.05, **p < 0.01. 1 Per protocol, G1 AK had baseline clinical score of “1” for erythema and hyperkeratosis, and G2 AK had a baseline clinical score of “2” for erythema and hyperkeratosis. Clinical scoring was performed on a scale of 0 (absent) to 3 (3+). 2 p-value from 3 to 6 months = 0.6815. 3 p-value from 3 to 6 months = 0.8702. 4 For non-dichotomous variables, percent prevalence calculated using ordinal logistical regression.
3.5.1. Stratum Corneum

Overall, parameters in the stratum corneum and epidermis exhibited the greatest response to treatment (Figure 1). In the stratum corneum of G1 AKs, 3/3 criteria in the cryotherapy group and 1/3 criteria in the PDT group displayed a significant reduction at 6 months compared to baseline. Hyperkeratosis was the only stratum corneum parameter that exhibited a statistically significant reduction at 6 months in G2 AKs, and this was seen in the cryotherapy treatment group. However, all (3/3) stratum corneum parameters in G2 AKs displayed a significant decrease at 3 months.

![Figure 1. Clinical and reflectance confocal microscopy images over time in an AK lesion. Baseline (a) and 6 months (d) dermoscopic assessment of a G1 AK in the cryotherapy group; improvement in stratum corneum disruption from baseline (b) to 6 months (e) post-cryotherapy; improvement in atypical honeycomb pattern from baseline (c) to 6 months (f) post-cryotherapy.](image)

3.5.2. Epidermis

In G1 AKs, 3/5 epidermal parameters in the cryotherapy group and 1/5 in the PDT group showed a significant reduction at 6 months compared to baseline. In G2 AKs, 2/5 epidermal parameters in both the cryotherapy and PDT groups displayed a significant reduction at 6 months.

3.5.3. Dermis

In G1 AKs, both cryotherapy and PDT significantly reduced the intradermal inflammatory infiltrate at 6 months. In the PDT-treated G1 lesions, round blood vessels significantly increased at both 3 and 6 months. Measures of polymorphous vessels and solar elastosis did not significantly change over time.

RCM parameters with the greatest significance across groups included atypical honeycomb pattern, hyperkeratosis, disarranged epidermal pattern, and stratum corneum disruption (Supplementary Figure S1). Of note, these four criteria also displayed the highest OR at baseline. Interestingly, various parameters in G2 AKs increased in prevalence from 3 to 6 months, likely representing early AK recurrence. Though these increases did not reach significance, this trend can be seen in 10/12 criteria in the PDT-treated G2 AKs, and 6/12 criteria in the cryotherapy-treated G2 AKs, with the most prominent increases...
observed in parakeratosis, stratum corneum disruption, and intraepidermal inflammatory cells.

3.6. Comparison between Treatment Groups

Overall, both treatments decreased AK severity as measured by RCM. Direct comparisons of the cryotherapy and PDT treatment group did not reach statistical significance in AK. Both the PDT and cryotherapy treatments were tolerated by all patients.

4. Discussion

RCM has the potential to standardize the therapeutic monitoring of actinic keratosis over time. However, the identification of relevant and reliable confocal parameters varies greatly across clinical trials, and a standardized protocol for RCM measurement of AK treatment response is needed. Although we observed significant reductions in many of the RCM criteria, our results suggest that atypical honeycomb patterns, hyperkeratosis, disarranged epidermal patterns, and stratum corneum disruptions are the most reliable criteria for assessing AK treatment response. These criteria displayed the highest baseline OR and exhibited significant responses to treatment across AK grades and treatment modalities. These results, taken together with recent publications by Benati et al., Ishioka et al., and Sousa et al., suggest that these criteria could be effectively implemented in future AK clinical studies using in vivo RCM as the primary monitoring tool for endpoint assessment [11,13,23]. It would be of significant benefit to simplify the RCM criteria used to date in order to maximize their adoption, and to incorporate evaluation of PD skin as an internal control.

Our identification of atypical honeycomb patterns as a reliable RCM parameter for AK is consistent with many reports highlighting the utility of keratinocyte disarray/atypical honeycomb patterns in AK response monitoring [11–13,23–28]. Hyperkeratosis [8,13,15,17,23,25,29] and stratum corneum disruption [8,11,12] have also been identified as significant RCM parameters, although many studies have not included these criteria in the RCM analysis of AK treatment responses. Interestingly, following its use in early clinical trials [6,16], future studies largely omitted disarranged epidermal pattern/epidermal disarray as an RCM criterion for AK, likely due to difficulty in standardization related to its varying definitions (ranging from cytological atypia, disordered keratinocyte maturation and/or architectural patterns) and challenges in its differentiation from similar epidermal parameters. It is possible that disarranged epidermal patterns displayed success in this study due to extensive standardization efforts. However, due to the difficulty in normalization, there are considerable limitations to its reliable use in clinical practice. Replacing this term with previously vetted and defined dermatopathology terminology may be useful. For instance, epidermal hyperplasia, epidermal atrophy, acanthosis, digitated, papillated, reticulated, basilar proliferation, or budding are histological terms used to describe changes to the epidermal architecture. Applying these terms to RCM may enable more specific descriptions that are readily understood and disseminated, as they have already been widely adopted and defined in histopathology. Although some studies have reported significant reductions in nucleated cells [11,25,29] and dendritic cells [8,12] our results did not indicate a reliable, significant change in those parameters.

Actinic keratoses arise in areas of photodamaged skin and thus are evaluated in that setting. This has important implications in selecting which RCM parameters should be most heavily weighted for diagnosis and evaluation of AK. For instance, we observe that the baseline comparison of AK and PD skin (Table 3) yielded different key criteria than solely considering baseline percent prevalence alone (Supplementary Table S1). To efficiently distinguish AK from surrounding sun damaged areas, characteristics of the surrounding skin must be considered. In addition, PD skin can also serve as a reference point for the “normalization” of the RCM parameters. It is important to note that in field-directed interventions, the therapeutic effect on the PD skin can also be observed.
One challenge to RCM is the ability to image thick, hyperkeratotic lesions. As the thickness increases, the edge sensitivity and resolution decrease [30]. Some reports of RCM in AK exclude thicker lesions and most that examine a variety of AKs do not stratify their results based on grade, thus limiting their conclusions across lesion types. By studying and analyzing G1 and G2 AK separately, we found that the baseline odds ratios were higher in thicker (G2) AKs than thinner (G1) AKs, suggesting that pathologic RCM criteria are more easily detected in higher-grade AKs. Additionally, more parameters displayed a statistically significant decrease at 6 months in G1 AKs as compared to G2 AKs, demonstrating that obtaining a durable treatment response is more feasible in thinner AKs.

Previous reports have demonstrated that RCM allows for the increased detection of subclinical AK [8,11,24,25,28]. Ulrich et al. concluded that architectural disruption and cellular pleomorphism best characterize subclinical AK [5]. Ishioka et al. reinforced these findings, reporting stratum corneum disruption and atypical honeycomb patterns as the most prevalent subclinical AK characteristics post-treatment [11]. We observed the rebound of multiple criteria from 3 to 6 months, which likely indicated subclinical, early signs of relapse, particularly in G2 AKs. The criteria with the largest increases from 3 to 6 months included stratum corneum disruption, parakeratosis, and intraepidermal inflammatory cells. These findings suggest that, along with atypical honeycomb patterns, stratum corneum disruption, parakeratosis, and intraepidermal inflammatory cells may also serve as a clue for early AK recurrence. It is important to note that this evidence of early relapse was observed via RCM even though clinical evaluation demonstrated a sustained response. It has been suggested that all AK lesions, regardless of thickness, have a significant risk for invasive progression [31]. This emphasizes the importance of surveillance and supports the utility of RCM for the non-invasive monitoring of subclinical disease, and identification of early post-treatment AK recurrence.

Subject to operator-dependence and differences in individual experience, interobserver variability is an inherent limitation to RCM [32]. Similar studies assessing AK treatment response through RCM have frequently incorporated two experienced evaluators, and some have included a single experienced reader [11–13,24–26,28]. In our study, interobserver agreement was measured between three experienced RCM evaluators. We have shown that through focused consensus meetings and standardization efforts, interobserver agreement can be improved. This can be best exemplified through the parameters of stratum corneum disruption and disarranged epidermal pattern. Disruption of the stratum corneum may be difficult to assess in images obtained with a tangential angle, due to the anatomical location of the skin. Through consensus meetings, it was agreed that scattered evidence of black lagoons and a presence of isolated, hyperkeratotic scales favors the presence of SC disruption, while abrupt geometrical loss of the field favors artifacts, due to the tangential angle at the time of image acquisition. Disruption of the epidermal pattern was defined as disruption of the architectural arrangement of the honeycomb pattern. Readers agreed that it was important to evaluate the surrounding epidermis pattern before deciding on the presence or absence of this feature, as even normal skin may have a certain degree of disarrangement due to multiple factors, including tangential imaging or artifacts. Consensus significantly increased interrater agreement. A similar standardization approach may be useful in future trials with multiple evaluators. Additionally, we demonstrated that even if parameters only display moderate agreement, significant changes can still be observed. For instance, though stratum corneum disruption was among the lowest proportion of agreement (61%, CI: 57–65), there was still a significant decrease in the RCM criteria at both time points, in both treatment groups.

Clinical trials have demonstrated efficacy of both PDT and cryotherapy in the treatment of AK; however, PDT has demonstrated increased efficacy in head-to-head comparisons [33]. Although we did not identify a significant difference between treatment groups, both individually demonstrated success in the treatment of AK, as measured by both RCM and clinical evaluation.
Limitations of this study include the length of follow-up and the number of patients. A larger sample size would aid in comparing PDT and cryotherapy treatment. Additionally, the sample size of 20 participants affected the width of confidence intervals for the odds ratios that are listed in Table 3. Given that the clinical scores in Table 4 are allotted a half-value score from zero to three, the standard deviation of these scores had expanded due to the variability of these measurements. Furthermore, this study was also limited by the length of follow-up. Expanding the study length past 6 months would have better characterized the recurrence of AK. Trials published after our data were collected suggest that the classification of AK by the degree of basal proliferation and growth patterns is a more reliable indicator of progression to invasive carcinoma than grading based on the degree of cytological atypia [31,34–36]. Our definition of disarranged epidermal pattern encompasses the basal proliferation patterns described in these recent reports. Future studies should evaluate whether the incorporation of a more “granular PRO-like” AK scoring system, or a scoring system based on the growth pattern of basal keratinocytes, could also be applied to RCM. Additional consensus studies to further refine RCM terminology and descriptors, followed by larger studies’ assessments of a wider range of RCM parameters, are necessary to progress this initial work in creating a standardized, universal protocol for RCM monitoring of AK treatment responses. This step is critical if we are to establish a more objective, sensitive, and reliable methodology to comparatively assess the multiple AK treatment options and lead the field of therapeutic prevention into a cost-effective practice.

5. Conclusions

We identified RCM parameters that both effectively differentiated AK from nearby, photodamaged skin and reliably assessed AK responses to treatment. There is great value in the implementation of RCM in AK therapeutic response trials, as non-invasive, objective monitoring of lesions over time will lead to the more reliable characterization of longitudinal responses. A standardized, universally accepted protocol will decrease subjectivity, simplify clinical practice, and allow for the more meaningful comparison of results amongst clinical trials. Furthermore, the cellular level resolution obtained by RCM, unlike clinical scoring, provides insight into the biological mechanisms of treatment responses and thereby could identify actionable data towards optimizing therapy.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cancers13215488/s1, Table S1: Percent prevalence among reads at baseline, Figure S1: Graphical representation of key RCM criteria over time.

Author Contributions: Conceptualization, C.C.-L., E.A. and R.H.-W.; Formal analysis, C.C.-L., C.N.M., K.S. and C.H.; Funding acquisition, C.C.-L.; Investigation, C.C.-L., E.A., G.P., F.J.L., M.U., P.H., M.C.O. and P.P.; Methodology, C.C.-L., K.S., C.H., E.A. and R.H.-W.; Project administration, C.C.-L. and E.A.; Writing—original draft, C.C.-L., C.N.M., K.S. and M.G.; Writing—review & editing, C.C.-L., C.N.M., K.S., C.H., E.A., G.P., F.J.L., M.U., P.H., M.C.O., P.P., M.G. and R.H.-W. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Targeted Prevention for Non-Melanoma Skin Cancer PO1 Grant (CA229112, NCI/NIH) and The Alan and Janice Levin Family Endowed Chair for Excellence in Cancer Research at the Arizona Cancer Center.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (EK-number: 25-095 ex 12/13).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.
Acknowledgments: We thank Ingrid Wolf, Martin Inzinger, Angelika Hofer, Alexandra Gruber-Wackernagel, and Regina Fink-Puches for their contributions to this study.

Conflicts of Interest: F.J.L. has received financial support from Galderma to attend scientific EURO-PDT-Meetings. M.C.O. is a consultant for Caliber ID and DermaSensor. M.G. is a consulting investigator on an investigator-initiated study for DBV technologies on a topic unrelated to those described in this study and serves as a research consultant to the Dermatology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. The rest of the authors state no conflict of interest in relation to this manuscript.

References

1. Callen, J.P.; Bickers, D.R.; Moy, R.L. Actinic Keratoses. J. Am. Acad. Dermatol. 1997, 36, 650–653. [CrossRef]
2. Cockerell, C.J. Histopathology of Incipient Intraepidermal Squamous Cell Carcinoma ("actinic Keratosis"). J. Am. Acad. Dermatol. 2000, 42, 11–17. [CrossRef]
3. Novak, B.; DuBois, J.; Chahrou, O.; Papusha, T.; Hirt, S.; Philippi, T.; Zogel, C.; Osenberg, K.; Schmitt, B.; Lübbert, H. Clinical Pharmacokinetics and Safety of a 10% Aminolevulinic Acid Hydrochloride Nanoemulsion Gel (BF-200 ALA) in Photodynamic Therapy of Patients Extensively Affected with Actinic Keratosis: Results of 2 Maximal Usage Pharmacokinetic Trials. Clin. Pharmacol. Drug Deo. 2021. [CrossRef]
4. Crow, L.D.; Saylor, D.; Griffith-Bauer, K.; Tello, T.; Aroyan, C.; Lazar, A.A.; Arron, S.T. Comparative Tolerability and Efficacy of Daylight, Conventional, and Combination Aminolevulinic Acid Photodynamic Therapy for Treatment of Actinic Keratosis. J. Am. Acad. Dermatol. 2021, 85, 967–969. [CrossRef]
5. Rubel, D.M.; Spelman, L.; Murrell, D.F.; See, J.-A.; Hewitt, D.; Foley, P.; Bosc, C.; Kerob, D.; Kerrouche, N.; Wulf, H.C.; et al. Daylight Photodynamic Therapy with Methyl Aminolevulinate Cream as a Convenient, Similarly Effective, Nearly Painless Alternative to Conventional Photodynamic Therapy in Actinic Keratosis Treatment: A Randomized Controlled Trial. Br. J. Dermatol. 2014, 171, 1164–1171. [CrossRef] [PubMed]
6. Aghassi, D.; Anderson, R.R.; González, S. Confocal Laser Microscopic Imaging of Actinic Keratoses in Vivo: A Preliminary Report. J. Am. Acad. Dermatol. 2000, 43, 42–48. [CrossRef] [PubMed]
7. Richtig, E.; Ahlgrimm-Siess, V.; Koller, S.; Gerger, A.; Horn, M.; Smolle, J.; Hofmann-Wellenhof, R. Follow-up of Actinic Keratoses after Shave Biopsy by in-Vivo Reflectance Confocal Microscopy—A Pilot Study. J. Eur. Acad. Dermatol. Venereol. 2010, 24, 293–298. [CrossRef] [PubMed]
8. Ulrich, M.; Krueger-Corcoran, D.; Roewert-Huber, J.; Sterry, W.; Stockfleth, E.; Astner, S. Reflectance Confocal Microscopy for Noninvasive Monitoring of Therapy and Detection of Subclinical Actinic Keratoses. Dermatology 2010, 220, 15–24. [CrossRef]
9. Ulrich, M.; Lange-Asschenfeldt, S.; González, S. In Vivo Reflectance Microscopy for Early Diagnosis of Nonmelanoma Skin Cancer. Actas Dermosifiliogr. 2012, 103, 784–789. [CrossRef]
10. Pellacani, G.; Ulrich, M.; Casari, A.; Prow, T.W.; Cannillo, F.; Benati, E.; Losi, A.; Cesinero, A.M.; Longo, C.; Argenziano, G.; et al. Grading Keratinocyte Atypia in Actinic Keratosis: A Correlation of Reflectance Confocal Microscopy and Histopathology. J. Eur. Acad. Dermatol. Venereol. 2015, 29, 2216–2221. [CrossRef]
11. Ishioka, P.; Maia, M.; Rodrigues, S.B.; Lellis, R.F.; Hirata, S.H. In Vivo Confocal Laser Microscopy for Monitoring of Actinic Keratosis Treatment: A Comparison with Histopathologic Assessment after Treatment with Topical 5% 5-Fluorouracil. J. Eur. Acad. Dermatol. Venereol. 2018, 32, 1155–1163. [CrossRef]
12. Pasquali, P.; Segurado-Miravalles, G.; González, S. Sequential Treatment of Actinic Keratosis with Cryotherapy and Ingenol Mebutate: Reflectance Confocal Microscopy Monitoring of Efficacy and Local Skin Reaction. Int. J. Dermatol. 2018, 57, 1178–1181. [CrossRef]
13. Benati, E.; Longhitanato, S.; Pampena, R.; Mirra, M.; Raucci, M.; Pellacani, G.; Longo, C. Digital Follow-up by Means of Dermatoscopy and Reflectance Confocal Microscopy of Actinic Keratosis Treated with Imiquimod 3.75% Cream. J. Eur. Acad. Dermatol. Venereol. 2020, 34, 1471–1477. [CrossRef]
14. Horn, M.; Gerger, A.; Ahlgriim-Siess, V.; Weger, W.; Koller, S.; Kerl, H.; Samonig, H.; Smolle, J.; Hofmann-Wellenhof, R. Discrimination of Actinic Keratoses from Normal Skin with Reflectance Mode Confocal Microscopy. Dermatol. Surg. 2008, 34, 620–625. [CrossRef]
15. Mota, A.; Piñeiro-Maceira, J.M.; Barcaui, C. Evaluation of Diagnostic Criteria of Actinic Keratosis through Reflectance Confocal Microscopy. Skin Res. Technol. 2020, 26, 883–890. [CrossRef]
16. Rishpon, A.; Kim, N.; Scope, A.; Porges, L.; Oliviero, M.C.; Braun, R.P.; Marghoob, A.A.; Fox, C.A.; Rabinovitz, H.S. Reflectance Confocal Microscopy Criteria for Squamous Cell Carcinomas and Actinic Keratoses. Arch. Dermatol. 2009, 145, 766–772. [CrossRef]
17. Tan, J.M.; Lambie, D.; Sinnya, S.; Sahebian, A.; Soyer, H.P.; Prow, T.W.; Ardigion, M. Histopathology and Reflectance Confocal Microscopy Features of Photodamaged Skin and Actinic Keratoses. J. Eur. Acad. Dermatol. Venereol. 2016, 30, 1901–1911. [CrossRef] [PubMed]
18. Ulrich, M.; Maltusch, A.; Rius-Diaz, F.; Röwert-Huber, J.; González, S.; Sterry, W.; Stockfleth, E.; Astner, S. Clinical Applicability of in Vivo Reflectance Confocal Microscopy for the Diagnosis of Actinic Keratoses. Dermatol. Surg. 2008, 34, 610–619. [CrossRef] [PubMed]
