Reliability and validity of conjunctival ultraviolet auto-fluorescence measurement

Justin C Sherwin,1,2 Charlotte M McKnight,3 Alex W Hewitt,1 Lyn R Griffiths,4 Minas T Coroneo,5 David A Mackey1,3,6

ABSTRACT

Background Conjunctival ultraviolet auto-fluorescence (UVAF) photography was developed to detect and characterise pre-clinical sunlight-induced UV damage. The reliability of this measurement and its relationship to outdoor activity are currently unknown.

Methods 599 people aged 16–85 years in the cross-sectional Norfolk Island Eye Study were included in the validation study. 196 UVAF individual photographs (49 people) and 60 UVAF photographs (15 people) of Norfolk Island Eye Study participants were used for intra- and inter-observer reliability assessment, respectively. Conjunctival UVAF was measured using UV photography. UVAF area was calculated using computerised methods by one grader on two occasions (intra-observer analysis) or two graders (inter-observer analysis). Outdoor activity category, during summer and winter separately, was determined with a UV questionnaire. Total UVAF equalled the area measured in four conjunctival areas (nasal/temporal conjunctiva of right and left eyes).

Results Intra-observer (p_c=0.988, 95% CI 0.967 to 0.996, p<0.001), and inter-observer concordance correlation coefficients (p_c=0.924, 95% CI 0.870 to 0.956, p<0.001) of total UVAF exceeded 0.900. When grouped according to 10 mm² total UVAF increments, intra- and inter-observer reliability was very good (k=0.81) and good (k=0.71), respectively. Increasing time outdoors was strongly with increasing total UVAF in summer and winter (P_{intra}<0.001). Conclusion Intra- and inter-observer reliability of conjunctival UVAF is high. In this population, UVAF correlates strongly with the authors’ survey-based assessment of time spent outdoors.

INTRODUCTION

Ophthalmomelioses describe a group of ophthalmic conditions in which sunlight is implicated in the pathogenesis. To date, sunlight has been associated with several different ophthalmic conditions, affecting every major anatomical component of the eye, as well as systemic conditions with the potential for ocular involvement.2 Excessive ultraviolet radiation (UVR) exposure plays a pathogenic role in a number of eye diseases: pterygium, cortical cataract, squamous cell carcinoma of the cornea and conjunctiva, acute photokeratitis and conjunctivitis and acute solar retinopathy.3 Developing a valid and reliable system to measure ocular sunlight exposure is pivotal for future epidemiological research into ophthalmomelioses and preventive strategies targeting diseases associated with insufficient or excessive UVR. In order to detect and characterise pre-clinical ocular surface sunlight-induced UV damage, conjunctival ultraviolet auto-fluorescence (UVAF) was developed.4 5 Underlying this paradigm is the concept of Wood’s light used in dermatology.4 Initial results suggested areas of fluorescence represented precursors to ophthalmomelioses.4 In people with established ophthalmomelioses, in particular pterygium, it was hypothesised that UVAF represented areas of cellular activity within pterygia.7 These initial studies were limited by small sample sizes and lack of quantitative assessment of the area of UVAF.

There were two aims for this study. First, we sought to determine the intra- and inter-observer reliability of conjunctival UVAF. Second, we wished to determine the validity of conjunctival UVAF in relation to outdoor activity level.

METHODS

Study population and recruitment

Norfolk Island is an Australian territory that features a sub-tropical climate most of the year. It is located approximately 1600 km NE of Sydney, Australia, in the Pacific Ocean at latitude 29° 02’ South—167° 56’ East. The Norfolk Island Eye Study (NIES), which commenced in late 2007, is a cross-sectional study performed on Norfolk Island. In the 2011 census, there were 2302 people (all ages) on Norfolk Island, of whom approximately 88% were permanent residents (sum of residents and general entry permit holders), and a similar proportion were aged ≥15 years.9 Comprehensive details about NIES have been described previously.9 In brief, we recruited permanent residents aged ≥15 years. More than 70% of permanent residents aged ≥50 years participated.

Ethics

Ethics approval was obtained from human research and ethics committees of Griffith University and the Royal Victorian Eye and Ear Hospital in Melbourne. Research was conducted in accordance with the Declaration of Helsinki and its subsequent tenets. Written informed consent was obtained prior to conducting the ophthalmic examination. In addition, there was local community consultation with the hospital administration, local doctors and optometrists and visiting ophthalmologists to check that all concerns were met regarding the possible long-term impact of the study.

Outdoor activity assessment questionnaire

The UV-exposure questionnaire contained questions relating to hair and skin colour, skin phenotype (history of sunburn; response to bright...
Conjunctival UV autofluorescence assessment and measurement

Digital photographs were obtained using both reflected visible light (control) and UV-induced fluorescence with the aid of two portable photographic systems. Each consisted of a height adjustable table equipped with subject headrest, camera positioning assembly, digital single-lens reflex camera, macro lens and filtered electronic flash. Each eye was photographed at 0.94 magnification with separate views of the nasal and temporal regions of both eyes. Coloured low-voltage light emitting diodes were positioned on stands in the subject’s visual field, at approximately 35° to the camera—subject axis to aid fixation.

The UVAF photography used a specially adapted electronic flash system fitted with UV-transmission filters (transmittance range 500 to 400 nm, peak 365 nm) as the excitation source. Subject fluorescence was recorded with a Nikon D100 (Nikon, Melville, New York, USA) digital camera and 105 mm f/2.8 Micro Nikkor (Nikon, Melville, New York, USA) lens fitted with infrared and UV barrier filters. Thus, only fluorescence was recorded by the camera. Images were saved in red-green-blue format. Each eye was photographed at 0.94 magnification immediately after fixation. The degree of UVAF against our questionnaire-based assessment of outdoor activity. Trends across categorical variables of were assessed using Cuzick’s nonparametric test for trend or Kendall’s rank correlation. The OR of having a UVAF area in the first quartile compared with the other three quartiles, according to outdoor activity level, was performed using logistic regression. Statistical significance was set at p<0.05. All analyses were undertaken in Intercooled Stata V.10.1 for Macintosh (Stata-Corp, College Station, 2009).

RESULTS

Inter-observer reliability assessment

We evaluated 196 photographs of 49 people. Mean age was 53.3±14.2 years (range 16–85 years) and 28 (57.1%) were women. A histogram of inter-observer differences revealed that differences in measurements between the two assessors approximated a normal distribution (not shown). CCCs were performed for all UVAF photograph combinations (table 1). The highest CCC corresponded to the total UVAF category. Difference in mean area between graders, and 95% limits of agreement were calculated as described above. Five categories for the area of UVAF were defined, based on previous observations of UVAF levels with and without sun protection. The area of UVAF was also calculated using one photograph per participant.

Association between UVAF and outdoor activity

There were 599 people with both outdoor activity and UVAF measurements. UVAF photographs were then used to validate degree of UVAF against our questionnaire-based assessment of outdoor activity. Trends across categorical variables of were assessed using Cuzick’s nonparametric test for trend or Kendall’s rank correlation. The OR of having a UVAF area in the first quartile compared with the other three quartiles, according to outdoor activity level, was performed using logistic regression. Statistical significance was set at p<0.05. All analyses were undertaken in Intercooled Stata V.10.1 for Macintosh (Stata-Corp, College Station, 2009).

Table 1 Concordance correlation coefficients (CCCs), inter-grader difference and 95% limits of agreement (LOA) for inter-observer reliability testing of area of UVAF in 196 photographs (49 people), and were masked to each other’s results. For the intra-observer component, one assessor JCS determined the area of UVAF twice with the same photographs (60 photographs; 15 people).

Using UVAF as a continuous variable, we calculated concordance correlation coefficients (CCC), termed p_c, according to the methods initially proposed, and later refined, by Lin. The CCC calculates measures of both precision and accuracy to determine whether the observed data significantly deviate from the line of perfect concordance plotted at 45°. Normal plots (Q-Q) are also presented graphically. Mean difference (and SD) between the two graders (inter-observer) or measurements (intra-observer) was reported. Ninety-five per cent limits of agreement (LOA) between the two different graders or measurements were calculated by the method outlined by Bland and Altman. We also categorised measurements of total UVAF. First, we grouped total UVAF into increasing increments of 5 mm² or 10 mm². Next we grouped total UVAF into either ‘none’ (0 mm²) or ‘any’ UVAF. For this approach, proportion of agreement and the Kappa statistic were used to measure agreement between graders. Interpretation of κ was performed according to the guidelines provided by Altman. Association between UVAF and outdoor activity.

RESULTS

Inter-observer reliability assessment

We evaluated 196 photographs of 49 people. Mean age was 53.3±14.2 years (range 16–85 years) and 28 (57.1%) were women. A histogram of inter-observer differences revealed that differences in measurements between the two assessors approximated a normal distribution (not shown). CCCs were performed for all UVAF photograph combinations (table 1). The highest CCC corresponded to the total UVAF category. Difference in mean area between graders, and 95%

Table 1 Concordance correlation coefficients (CCCs), inter-grader difference and 95% limits of agreement (LOA) for inter-observer reliability testing of area of UVAF in 196 photographs (49 people), and were masked to each other’s results. For the intra-observer component, one assessor JCS determined the area of UVAF twice with the same photographs (60 photographs; 15 people).

Using UVAF as a continuous variable, we calculated concordance correlation coefficients (CCC), termed p_c, according to the methods initially proposed, and later refined, by Lin. The CCC calculates measures of both precision and accuracy to determine whether the observed data significantly deviate from the line of perfect concordance plotted at 45°. Normal plots (Q-Q) are also presented graphically. Mean difference (and SD) between the two graders (inter-observer) or measurements (intra-observer) was reported. Ninety-five per cent limits of agreement (LOA) between the two different graders or measurements were calculated by the method outlined by Bland and Altman. We also categorised measurements of total UVAF. First, we grouped total UVAF into increasing increments of 5 mm² or 10 mm². Next we grouped total UVAF into either ‘none’ (0 mm²) or ‘any’ UVAF. For this approach, proportion of agreement and the Kappa statistic were used to measure agreement between graders. Interpretation of κ was performed according to the guidelines provided by Altman. Association between UVAF and outdoor activity.

RESULTS

Inter-observer reliability assessment

We evaluated 196 photographs of 49 people. Mean age was 53.3±14.2 years (range 16–85 years) and 28 (57.1%) were women. A histogram of inter-observer differences revealed that differences in measurements between the two assessors approximated a normal distribution (not shown). CCCs were performed for all UVAF photograph combinations (table 1). The highest CCC corresponded to the total UVAF category. Difference in mean area between graders, and 95%

Table 1 Concordance correlation coefficients (CCCs), inter-grader difference and 95% limits of agreement (LOA) for inter-observer reliability testing of area of UVAF in 196 photographs (49 people), and were masked to each other’s results. For the intra-observer component, one assessor JCS determined the area of UVAF twice with the same photographs (60 photographs; 15 people).

Using UVAF as a continuous variable, we calculated concordance correlation coefficients (CCC), termed p_c, according to the methods initially proposed, and later refined, by Lin. The CCC calculates measures of both precision and accuracy to determine whether the observed data significantly deviate from the line of perfect concordance plotted at 45°. Normal plots (Q-Q) are also presented graphically. Mean difference (and SD) between the two graders (inter-observer) or measurements (intra-observer) was reported. Ninety-five per cent limits of agreement (LOA) between the two different graders or measurements were calculated by the method outlined by Bland and Altman. We also categorised measurements of total UVAF. First, we grouped total UVAF into increasing increments of 5 mm² or 10 mm². Next we grouped total UVAF into either ‘none’ (0 mm²) or ‘any’ UVAF. For this approach, proportion of agreement and the Kappa statistic were used to measure agreement between graders. Interpretation of κ was performed according to the guidelines provided by Altman. Association between UVAF and outdoor activity.

RESULTS

Inter-observer reliability assessment

We evaluated 196 photographs of 49 people. Mean age was 53.3±14.2 years (range 16–85 years) and 28 (57.1%) were women. A histogram of inter-observer differences revealed that differences in measurements between the two assessors approximated a normal distribution (not shown). CCCs were performed for all UVAF photograph combinations (table 1). The highest CCC corresponded to the total UVAF category. Difference in mean area between graders, and 95%
LOA increased with increasing numbers of photos contributing to the UVAF measurement. Inter-observer reliability was also assessed using normal (Q-Q) and repeatability plots of total UVAF (figure 1). The linearity of the Q-Q plot demonstrated a normal distribution. Variation between assessors was mildly dependent on the magnitude of measurements. When the area of UVAF was much greater, the variation between the two observers increased slightly. There was no evidence of a major systematic bias. The null hypothesis of equal variance between the two measurements was not rejected for total UVAF (Fisher’s test of difference in variance, r=0.099, p=0.500).

For total UVAF categories delineated by 10 mm², agreement was good: agreement=79.6%, κ statistic (κ) =0.71, p<0.001. Reliability was reduced, but still good, with 5 mm² categories: agreement =73.5%, κ =0.61, p<0.001. Reliability of presence or absence of UVAF revealed: agreement=98.0%, κ=0.79, p<0.001.

Intra-observer reliability assessment

The total UVAF category was associated with a substantial strength of agreement, greater than that of the inter-observer analysis (table 2). Mean differences between the two measurements, and in the 95% LOA, were considerably less compared with the results of the inter-observer analysis. A histogram of differences between the two measurements approximated a normal distribution, as supported by the normal (Q-Q) plot (figure 2). There was no evidence of systematic bias. Pitman’s test for total UVAF revealed r=0.053, p=0.801.

For both 5 mm² and 10 mm² categories, reliability was very high: agreement =86.7%, κ=0.81, p<0.001 for both categories. There were no total UVAF measurements given as 0 mm² in any measurements in the intra-observer component.

Association between UVAF and outdoor activity

Time spent outdoors was higher in men than women (Kendall’s τ, p<0.001), but there was no statistically significant change in time spent outdoors with increasing age category (p<0.001). Total UVAF increased with increasing time outdoors in both sexes. For men, median UVAF increased across the three groups of outdoor activity (p<0.001): <1/4 day outside (median UVAF 28.3 mm² (IQR 11.4–57.0)); 1/2 day outside (32.8 mm² (17.6–55.7)); >3/4 day outside (45.1 mm² (26.6–66.1)). Similarly, for women, median total UVAF increased from 20.0 mm² (9.1–34.1) to 25.5 mm² (15.2–45.5) to 35.0 mm² (17.8–65.9) across the three categories (p<0.001). Median UVAF increased across the three categories when both sexes were analysed together (figure 3). The correlation between total UVAF and outdoor activity was also assessed (Spearman’s ρ=0.29, p<0.001). There was also a significant trend of increasing median UVAF with increasing time spent outdoors in winter (p<0.001): mostly indoors 23.3 (IQR 9.7–55.2) mm²; half indoors/outdoors 28.7 (IQR 13.8–49) mm²; mostly outdoors 33.9 (IQR 21.6–58.1) mm².

We divided total UVAF into two categories: lowest quartile and highest three quartiles. The bottom quartile equalled total UVAF <14.5 mm². With increasing outdoor activity in summer, there were reduced odds of having a UVAF area in the first quartile compared with the other three quartiles (p<0.001). Compared with the lowest category of outdoor activity, the ORs of having a UVAF in the lowest quartile were

Table 2 Concordance correlation coefficients (CCCs), intra-grader difference and 95% limits of agreement for intra-observer reliability UV autofluorescence measurement

<table>
<thead>
<tr>
<th>Measurement</th>
<th>CCC_p_c (95% CI)</th>
<th>p Value</th>
<th>Difference between measurements (mm²) mean (SD)</th>
<th>95% LOA (mm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right nasal (1 photo per participant)</td>
<td>0.915 (0.721 to 0.964)</td>
<td>&lt;0.001</td>
<td>−0.81 (3.59)</td>
<td>−7.63 to 6.01</td>
</tr>
<tr>
<td>Right temporal (1 photo per participant)</td>
<td>0.824 (0.558 to 0.937)</td>
<td>&lt;0.001</td>
<td>0.60 (3.61)</td>
<td>−6.46 to 7.62</td>
</tr>
<tr>
<td>Left nasal (1 photo per participant)</td>
<td>0.991 (0.947 to 0.996)</td>
<td>&lt;0.001</td>
<td>0.14 (0.84)</td>
<td>−1.48 to 1.70</td>
</tr>
<tr>
<td>Left temporal (1 photo per participant)</td>
<td>0.897 (0.741 to 0.944)</td>
<td>&lt;0.001</td>
<td>1.54 (2.90)</td>
<td>−4.22 to 7.10</td>
</tr>
<tr>
<td>Total nasal (sum of 2 photos)</td>
<td>0.919 (0.782 to 0.972)</td>
<td>&lt;0.001</td>
<td>−0.79 (3.83)</td>
<td>−8.14 to 6.64</td>
</tr>
<tr>
<td>Total temporal (sum of 2 photos)</td>
<td>0.884 (0.678 to 0.940)</td>
<td>&lt;0.001</td>
<td>−0.75 (4.43)</td>
<td>−9.30 to 8.05</td>
</tr>
<tr>
<td>Total left (sum of 2 photos)</td>
<td>0.895 (0.729 to 0.962)</td>
<td>&lt;0.001</td>
<td>−1.63 (3.02)</td>
<td>−7.43 to 4.22</td>
</tr>
<tr>
<td>Total right (sum of 2 photos)</td>
<td>0.954 (0.870 to 0.984)</td>
<td>&lt;0.001</td>
<td>0.20 (2.72)</td>
<td>−5.21 to 5.53</td>
</tr>
<tr>
<td>Total UVAF (sum of 4 photos)</td>
<td>0.988 (0.967 to 0.996)</td>
<td>&lt;0.001</td>
<td>−1.41 (1.90)</td>
<td>−5.26 to 2.39</td>
</tr>
</tbody>
</table>
Figure 2 (A) Normal (Q-Q) Plot of intra-reliability measurement of total UV autofluorescence. (B) Repeatability plot of intra-reliability measurement of total UV autofluorescence. The dashed line at 45 degrees equals the line of perfect concordance. \( r = 0.998 \) (95% CI 0.967 to 0.996); Pearson's \( r \) = 0.993; bias correction factor \( (C_b) = \frac{r}{\sqrt{1-r^2}} = 0.996 \); reduced major axis: slope = 0.994; intercept = -1.245.

DISCUSSION

It is necessary to develop ocular-specific UVR exposure methods as the eye has a unique morphology such that the degree of limbal focusing of peripheral light is related to the corneal shape curvature.14 15 Methods used previously were generally impractical and unfeasible for large epidemiological studies, health promotion or clinical practice.16 17 We have found that conjunctival UVAF is a highly reliable measurement, and differences between intra- and inter-observer reliability measurements were minor. Conjunctival UVAF correlated very well with our questionnaire-based outdoor activity level.

The highest CCCs were for total UVAF measurements, representing sum of area in four measurements in both eyes; \( \kappa \) statistics for total UVAF in 5 mm\(^2\) or 10 mm\(^2\) increments indicated good and very good agreement, respectively.13 Possibly of more importance is the ability to discriminate between individuals with and without any discernible UVAF, and our inter-observer agreement was nearly perfect at 98%. The inter-observer CCC figure revealed that the variation between the measurements increased as the degree of UVAF increased, as displayed by increased numbers of outlier values. This is to be expected, as the recognition of area that autofluoresces is operator dependent, and the delineation between an area of UVAF and no UVAF is often not unequivocal, especially when the quality of the image is reduced.

The mean difference between graders (inter-observer) and measurements (intra-observer) was approximately 5 and 1.5 mm\(^2\), respectively for total UVAF. The corresponding 95% LOA was between -15.7 and 19.7 mm\(^2\) and between -5.3 and 2.4 mm\(^2\), for inter- and intra-reliability measures, respectively. To aid clinical interpretability we have used a measurement of 10 mm\(^2\) total UVAF to understand the clinical relevance of UVAF. In individuals aged >15 years in NIES, total UVAF was highest in men (median UVAF approximately 10 mm\(^2\) larger than women) and younger individuals \( (P_{\text{trend}} < 0.001) \).5 For a 10 mm\(^2\) increase in total UVAF area, the OR for a pterygium (any eye) was 1.16 (95% CI 1.05 to 1.28 \( p = 0.002 \)) in a multivariate model that adjusted for age, sex, skin type and time spent outdoors.16 Hence, the 95% LOA in the inter-observer results indicate a possible clinically meaningful difference.

Median UVAF correlates well with outdoor activity and can be considered a valid biomarker for usual outdoor exposure in both summer and winter. With the lowest category of outdoor activity, there was an approximate threefold increase in odds of having a UVAF reading in the first quartile compared with those with the highest outdoor activity \( (>3/4 \text{ day}) \): OR 2.95, 95% CI 1.69 to 5.16, \( p < 0.001 \). This association was slightly attenuated following adjustment for age and sex (OR 2.45, 95% CI 1.34 to 4.47, \( p = 0.005 \)).

Figure 3 Degree of total UV autofluorescence (UVAF) according to category of outdoor exposure in summer. There was an increase in median total UVAF (mm\(^2\)) with increasing time. Outdoors \( (P_{\text{trend}} < 0.001) \). Error bars correspond to the interquartile range (IQR) of total UVAF.
difficulty exists in finding a balance between insufficient and excessive UVR exposure. At one end of the spectrum, insufficient UVR puts one at increased risk of vitamin D deficiency and skeletal disease, whereas on the other end, increased risks of ophthalmohelioses and skin cancer are manifest. Such findings are difficult to convey to the public and to translate into effective public health messages regarding what constitutes a safe level of UV exposure.

There were several additional limitations inherent in our study. Our findings were obtained in a small, geographically isolated population. Even so, one benefit of using the Norfolk Island population is that it has a homogenous environment (including latitude, weather patterns and UVR exposure). Our study is also limited by the use of our questionnaire-based assessment of outdoor exposure, which was only categorised into three groups. Using questionnaire methods, a more complex outdoor exposure could have been obtained, for example, by asking about outdoor activity at different times of the week (eg, weekday vs weekend), measuring outdoor activity as a continuous variable (eg, hours per day/week) or with use of an activity diary. Finally, the numbers of measurements used in the reliability analyses, especially intra-observer reliability, were relatively small.

In conclusion, we have shown that conjunctival UVAF is a highly reliable measurement that correlates strongly with level of outdoor activity. We propose that conjunctival UVAF can be used as a biomarker of sub-acute sunlight exposure. Further study is required to illustrate the precise nature of the tissue that autofluoresces at a histological level and to study the relationship between UVAF and ophthalmohelioses in other populations.

Acknowledgements We wish to thank the community of Norfolk Island for participation in this research and for welcoming us into their community.

Contributors JCS conceived the idea of the study, performed the ultraviolet autofluorescence analysis, completed the statistical analysis and also wrote the first draft of the manuscript. CMM also performed the ultraviolet autofluorescence analysis. AWH and DAM took photographs of the participants. AWH, LRG, MTC and DAM were supervisors of the project. All authors were responsible for revising and approving subsequent drafts of the article prior to submission.

Funding Financial support was provided by the Royal Victorian Eye and Ear Hospital research committee and the Peggy and Leslie Cranbourne Foundation. Centre for Eye Research Australia receives operational infrastructure support from the Victorian government. We wish to thank the community of Norfolk Island for participation in this research and for welcoming us into their community.

Competing interests The authors indicate no financial conflict of interest. DAM is a recipient of the Pfizer Australia Senior Research Fellowship. MTC discloses that he is the inventor of US Patent 7 217 289: Ocular scaffold for stem cell cultivation and methods of use, US patent application 20060204474: Treatment of epithelial layer lesions, US patent application 20050267115: Treatment of ocular lesions; is a consultant to Allergan, Inc. in the area of medical treatment of pterygium and dry eye and has received research funds and travel support; receives royalties from Eagle Vision Inc. in relation to a dry eye product; and has been a consultant to Johnson and Johnson Vision Care Inc and has received research funds and travel support.

Ethics approval The ethics approval was provided by the human research and ethics committees of Griffith University, Royal Victorian Eye and Ear Hospital in Melbourne.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES