Quantification of Sun-related Changes in the Eye in Conjunctival Ultraviolet Autofluorescence Images

Emily Huynh\textsuperscript{1,*}, Danuta M. Bukowska\textsuperscript{1}, Seyhan Yazar\textsuperscript{1}, Charlotte M. McKnight\textsuperscript{1}, Ajmal Mian\textsuperscript{2}, David A. Mackey\textsuperscript{1}

\textsuperscript{1}Centre for Ophthalmology and Visual Science, University of Western Australia / Lions Eye Institute, 2 Verdun Street, Nedlands WA 6009, Perth, Australia
\textsuperscript{2}School of Computer Science and Software Engineering, University of Western Australia, 35 Stirling Highway
Crawley WA 6009 Perth, Australia

*Corresponding author
Email: EmilyHuynh@lei.org.au (EH)

Keywords: image processing; photography; sunlight; ophthalmoelioses
Abstract

Quantification of sun-related changes in Conjunctival Ultraviolet Autofluorescence (CUVAF) images is a subjective and tedious task in which reproducibility of results is difficult. Thus, we have developed a semi-automatic method in Matlab® to analyze CUVAF images. The algorithm was validated on 200 images from 50 randomly selected participants from the Western Australian Pregnancy Cohort (Raine) Study 20 year follow-up assessment where CUVAF area measurements were available from previous manual analysis. Algorithm performance was compared to manual measurements and yielded better than 95% correspondence in both intra- and inter-observer agreement. Furthermore, the semi-automatic method significantly reduced analysis time by 50%.

Introduction

Conjunctival Ultraviolet Autofluorescence (CUVAF) photography was developed based on the principle of Wood’s light used in dermatology to detect and characterize pre-clinical ocular sunlight-induced changes on the conjunctiva [1, 2]. A pilot study on school-aged children demonstrated that CUVAF was a more sensitive method of detecting precursors of ocular sun damage than traditional reflected visible light (control) photography. As precursors of sun damage occur years before clinical manifestation of sunlight-associated ophthalmic disease, early detection could be a key step in prevention of ophthalmohelioses [1]. Furthermore, the area of CUVAF has been found to be highly correlated with time spent outdoors, suggesting that CUVAF could be used as an objective marker for sun exposure [3]. CUVAF photography has since been used in studies to explore the epidemiology and pathogenesis of ophthalmohelioses as well as the association between sun exposure and other eye disorders.
In a study led by Coroneo and colleagues, [4] ultraviolet (UV) fluorescence patterns of pterygia were described using CUVAF photography, furthering our understanding of pterygium pathophysiology. In 2011, the first quantifiable estimates of CUVAF in the general population were reported as part of the Norfolk Island Eye Study. The aim of the study was to determine the relationship between age, gender and CUVAF in an adult population-study, using CUVAF as a clinical marker of facial ultraviolet exposure [5]. Other studies have investigated the association between area of CUVAF and established ophthalmic disease such as pterygium and refractive error. Increasing CUVAF is associated with prevalent pterygia in both the adult population and the young adult population [6, 7]. In contrast, myopia in young adults is inversely correlated to the amount of CUVAF [3, 8]. These studies were conducted in Australia and Norfolk Island, where the level of exposure to ultraviolet radiation is high due to an outdoor lifestyle, relatively low latitudes and sub-tropical climate [9]. A similar study using CUVAF photography was conducted in the United Kingdom, to investigate the prevalence of CUVAF in a diverse population of European and Middle Eastern eye care practitioners.

The CUVAF area was quantified using Adobe Photoshop CS4 Extend (Adobe Systems Inc., San Jose, CA, USA) in the Australian studies [5-8, 10] and ImageJ (US National Institute of Health, Bethesda, MA, USA) in the United Kingdom study [11]. For each photograph, the region of CUVAF was manually delineated by an assessor, requiring great attention to detail and a steady hand. Once the area was defined, the area was converted to mm$^2$ using a calibration factor determined by photographing a ruler with the same camera system. Although the reliability of this method has been validated [2], it is inherently subjective, tedious, and prone to human error. Furthermore, the marked images were not saved, making retrospective review of the CUVAF region impossible.
The aim of this study was to develop and validate the reliability of a semi-automatic method for CUVAF analysis, and to determine whether it could replace the current manual method. Ideally, if both the area measurements and saved images showing manual delineation of the CUVAF region were available, it might be possible to utilize a machine learning technique to develop a fully automatic algorithm to eliminate subjectivity. However, as only area measurements were available, an algorithm for semi-automatic CUVAF quantification is proposed.

**Methods**

**Retrospective data**

Data for this validation were derived from the Raine Study 20 year follow-up assessment where CUVAF images were previously measured using the manual method. Full details of the study can be found in a previous publication [10]. In brief, 1344 individuals were enrolled in the 20 year follow-up study and completed a comprehensive eye assessment, which included the acquisition and analysis of CUVAF photographs as described in [7]. Written informed consent was obtained from all participants prior to the examinations. The study obtained ethics approval from the Human Research Ethics Committee at the University of Western Australia and adhered to the Declaration of Helsinki.

CUVAF images from 50 participants (total of 200 images) were randomly selected from the Raine study to validate the algorithm. Two assessors (EH and DB) analyzed these images using the proposed algorithm (see next section). Both assessors were given an inclusion criterion to define what constitutes a CUVAF region. The results from each assessor were compared with the existing manual measurements to determine, whether the manual and semi-automatic methods for CUVAF analysis could be used interchangeably. Furthermore, intra- and inter-
observer reliability was also evaluated for the semi-automatic method. The Bland-Altman test was used to assess the level of agreement between each dataset where the limits of agreement (LOA) was defined as $1.96 \times \text{Standard Deviation (SD)}$ [12, 13].

**Algorithm**

The algorithm for semi-automatic quantification of sun-related changes in CUVAF images was implemented in Matlab® 2011b. A Graphic User Interface (GUI) was designed to allow users to interact with the algorithm to assist defining the CUVAF regions. The measured areas, delineated images and parameters required to reproduce the results were automatically saved in the process.

**Pre-processing**

CUVAF digital images often need to be adjusted to enable assessors to see signs of sun-related changes. To ensure no bias was introduced in the process, all images were pre-processed as follows. Images were resized by a factor of 0.5 to reduce processing time and the red channel was removed to reduce the effects of red/white light artifacts arising from the reflection of the camera flash off the surface of the eye. To enhance the brightness and contrast of the images, the tolerance values to saturate 2% of the darkest pixels and 1% of the brightest pixels in the blue channel was determined using ‘stretchlim’. This tolerance value was used as an argument for the contrast stretching function ‘imadjust’ which was applied to the green and blue channels. The pre-processed RGB image was reconstructed by concatenating a zeroed red channel with the modified green and blue channels, which was converted to grayscale for analysis. An example of a pre-processed image for the selection of the Region of Interest (ROI) is shown in Fig. 1b.
If CUVAF is present, the region of interest (ROI) is defined by the assessor by roughly marking the region of UV-induced autofluorescence using ‘roipoly’. As the algorithm can be confounded by imaging artifacts, these should be excluded from the ROI. A binary mask was created from the ROI and applied to the grayscale image. The coordinates of the ROI are automatically saved in the Excel Spreadsheet.

Local thresholding

Due to the presence of uneven illumination in CUVAF photographs, a local thresholding method is used to segment the CUVAF regions from the ROI in the masked grayscale image. As the amount of CUVAF in each photograph is highly variable in a population, having a fixed window size to process all images would not be effective in cases where the CUVAF region is much larger or smaller than the window. Therefore, we chose to define a square window for local thresholding that is adapted to the size of the CUVAF region, by making it approximately equal to 15% of the ROI area ($A_{ROI}$), where the width is determined by Eq. (1):

$$w = round\left(\sqrt{0.15 \cdot A_{ROI}}\right)$$  \hspace{1cm} (1)
A sliding window approach is used to segment the CUVAF region in the ROI, where the center is shifted at steps equal to 1/3 the width across the grayscale image to allow overlap. To ensure the edges of the potential CUVAF regions are not underrepresented from overlapping the windows, a dilated ROI mask is created to discriminate between the ROI and its surrounding edge. The ROI mask is dilated by a factor equal to the step size, which is then added to the original ROI mask, such that the pixels representing the original mask have a value of 2, and the dilated edges, a value of 1 (Fig. 2). The original window size is used for thresholding regions on the mask with a value of 2 whereas a window a quarter of the size of the original was used for edges regions with value 1. Reducing the window size at the edges helps to preserve local details and as well as to better distinguish CUVAF from the background as the local region under scrutiny is highly likely to contain an even proportion of CUVAF and background.

Fig 2. Example of the dilated mask used to distinguish the edges of the potential CUVAF regions. White regions (value 2) represent the ROI undergo local thresholding with the original window size. Grey regions (value 1) represent the dilated ROI are thresholded with a window, a quarter of the size of the original.

A modified version of Niblack’s local thresholding method [14] is used to binarize each window. The local threshold \( T \) is calculated using Eq. (2) where \( m \) and \( \sigma \) are the mean and standard deviation of the intensity in the local window respectively, and \( k \) is a constant set as 0.5 [10].

\[
T = m + k \cdot \sigma
\]  

(2)
Pixels with intensities greater or equal to the local threshold are considered to represent sun-related change are given a value of 1, and 0 otherwise. The output of the overall local thresholding method is the sum of all the binary images from each window, creating an - image we call the overlap map. Shown in Fig. 3a below, is an example of an overlap map, where colors are used to represent the number of overlaps. The overlap map is smoothed with a Gaussian filter of dimensions equal to the sliding window step (Fig. 3b), to remove horizontal and vertical lines introduced from summing overlapping windows. Furthermore, smoothing the overlap map also ‘joins’ any gaps in the affected regions such that the final result would be more similar to the results expected from the manual measurements. The overlap map image was multiplied with the dilated mask to remove any noise picked up from the local thresholding method. Small regions were also removed as they were deemed insignificant.

Fig. 3. (a) Color map of the summed thresholded windows. (b) Smoothed overlap map; the color bar shows the number of overlaps for each pixel.

Refining the CUVAF Region

Pixels on the overlap map with a value greater or equal to the default overlap threshold of 1, are considered to represent CUVAF regions. The GUI allows the assessor to adjust the overlap threshold to refine the CUVAF region. If the assessor is not satisfied with any of the threshold
values for the overlap map, the ROI can be redrawn to obtain a different overlap map. Increasing/decreasing the size of the ROI can change the distribution of values in the overlap map. The CUVAF region can also be refined by removing isolate regions. An example of the refined CUVAF region is shown in Fig. 4.

Fig. 4. Output image of semi-automatic CUVAF analysis algorithm, where the CUVAF region is delineated in red.

**CUVAF Area Calculation**

To determine the CUVAF area in mm$^2$, the number of pixels constituting the CUVAF region is counted. The number of pixels is then converted to mm$^2$ using Eq. (3), where $A_m$ and $A_p$ represents the CUVAF area in mm$^2$ and number of pixels respectively, and $f_c$ is the calibration factor that has been adjusted according to the image resize factor in the pre-processing step.

$$A_m = A_p \cdot f_c^2$$

(3)

**Saving Data**

A Microsoft Excel spreadsheet containing information from CUVAF analysis is automatically created for each subject. The data is saved in a specific format, allowing it to be easily collated, using a macro written in Visual Basics for Applications in Microsoft Excel. In this spreadsheet, the measured area, as well as the parameters: resize factor, overlap threshold,
window width ratio, area threshold filter, ROI coordinates and coordinates of removed regions are saved, enabling us to reproduce the marked image for retrospective review.

**Graphic User Interface**

The GUI (Fig. 5) was developed to provide a convenient platform in which the assessor can access all CUVAF images for analysis, and have the results automatically saved. The GUI has functions that allow users to: (1) preview the pre-processed images; (2) change image display during CUVAF delineation (original, enhanced, grayscale or color image); (3) identify assessor with initials; (4) flag images as subjective or weak CUVAF; (5) write comments on individual images; (6) use a previously saved ROI or reproduce results for post-analysis modifications.

![Screenshot of the GUI for CUVAF analysis.](image)

**Results**
The algorithm and GUI provided an efficient method to access, measure and save all data, and was found to reduce analysis time by approximately 50% when compared to the manual method. The results of the Bland-Altman tests comparing the semi-automatic and the manual method for CUVAF measurement, as well as the reliability and validity of the semi-automatic method are shown in Table 1 with the plots in Fig. 6.

Table 1: Results of the Bland-Altman test comparing the semi-automatic and manual method, and the inter- and intra-observer reliability of the semi-automatic method

<table>
<thead>
<tr>
<th>Type of Comparison</th>
<th>Assessors</th>
<th>Mean (SD) difference (mm²)</th>
<th>95% LOA (mm²)</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-automatic vs manual</td>
<td>1A vs Manual</td>
<td>-8.10 (9.93)</td>
<td>-27.55 to 11.36</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>2 vs Manual</td>
<td>-11.95 (11.83)</td>
<td>-35.14 to 11.25</td>
<td>96</td>
</tr>
<tr>
<td>Inter-observer</td>
<td>1A vs 2</td>
<td>-3.85 (6.70)</td>
<td>-16.99 to 9.29</td>
<td>96</td>
</tr>
<tr>
<td>Intra-observer</td>
<td>1A vs 1B</td>
<td>0.72 (2.96)</td>
<td>-5.09 to 6.53</td>
<td>98</td>
</tr>
</tbody>
</table>
Fig. 6. Bland-Altman plot of difference against mean CUVAF comparing: (a) Assessor 1A and manual measurements. (b) Assessor 2 and manual measurements. (c) Assessor 1A and Assessor 2 (inter-observer reliability). (d) Assessor 1A and Assessor 1B (intra-observer reliability). The red lines represent the mean bias and the LOAs are shown by the green lines.

Manual vs Semi-automatic Method for CUVAF Analysis

Comparing the semi-automatic and manual method for CUVAF analysis, the LOA ranged from -27.55 to 11.36 for Assessor 1 and -35.14 to 11.25 for Assessor 2, with a mean bias of -8.10 and -11.95 respectively (Table 1). The negative mean bias observed in the results of both assessors suggests that measurements made using the semi-automatic method yields smaller CUVAF area than with the manual method. Both assessors using the semi-automatic method
achieved 96% agreement with the manual method, where % agreement is the percentage of the points lying within their respective LOA. As over 95% agreement was attained, the semi-automatic and manual methods of quantifying CUVAF can be used interchangeably, provided that the observed difference would not impact clinical management [12]. Looking at the Bland-Altman plots in Figs 6(a) and 6(b) respectively, we note that there is a slight negative relationship between the measurement variability and range of CUVAF measurements. As the CUVAF area increases, there appears to be a larger negative difference between manual and semi-automatic method.

**Inter-observer reliability assessment**

In an assessment of inter-observer reliability between Assessor 1A and Assessor 2 using the semi-automatic method, the LOA was found to range from -16.99 to 9.29 with a mean bias of -3.85 (6.70) (Table 1). The negative mean bias indicates that Assessor 2 has a tendency to mark a slightly smaller CUVAF region than Assessor 1. 96% agreement was achieved between the two assessors using the semi-automatic method. The Bland-Altman plot between Assessor 1A and 2 (Fig. 6c) showed no significant change in inter-observer variability for the range of measured CUVAF areas.

**Intra-observer reliability assessment**

In an assessment of intra-observer reliability, an agreement of 98% was achieved with a mean bias of 0.72 (2.96) and LOA ranging from -5.09 to 6.53 from the measurements of Assessor 1 at two different time points. As the Bland-Altman plot shows no significant change in intra-observer variability for the range of CUVAF measurements, and the mean difference is close to zero and LOA range is small, we conclude that measurements made using the semi-automatic method are repeatable.
A previous study by Sherwin et al. [2] looked at the inter- and intra-observer reliability and validity of quantifying CUVAF using the manual method. In their study, 196 photographs from 49 participants from the Norfolk Island Eye study were analyzed for inter-observer reliability assessment, and 60 photographs (15 participants) for the intra-observer assessment. It was noted the LOA and mean bias increases with the number of images analyzed [2]. Direct comparison of the results of the our study with the Sherwin study cannot be made due to the large differences in age of the participants (inverse relationship between age and CUVAF [5]) as well as the difference in the number of images analyzed for intra-observer reliability assessment. However they can be used as a guide for tolerable ranges of variation.

The inter-observer reliability assessment results from the Sherwin study reported the LOA between its two assessors using the manual method to be -13.67 to 19.71 with a mean difference of 3.02 (8.52) compared to ours using the semi-automatic method -16.99 to 9.29 with a mean bias of -3.85 (6.70). A similar number of images were analyzed in both our studies and as the range of the LOA, mean bias and standard deviation is closer to zero using the semi-automatic method, suggests that it may be a more reliable method of quantifying CUVAF.

For intra-observer reliability assessment, the Sherwin study reported the LOA of -5.26 to 2.39 with a mean bias of -1.41 (1.90) compared to our findings of LOA ranging from -5.09 to 6.53 and mean bias of 0.72 (2.96). The Sherwin study analyzed only 15 participants for intra-observer reliability compared to the 50 participants we analyzed which could account for the wider LOA observed in our analysis. Regardless, the variations noted in our study are small enough suggesting that the semi-automatic method also has high reliability.
One limitation of our study was that the manual assessor of the retrospective data was not recruited to analyze the subset of images using the semi-automatic method. As a result, larger measurement error is expected when comparing the manual and semi-automatic method for quantifying CUVAF as it would include variability due to both inter-observer and measurement method. Hence, two assessors were recruited to measure the retrospective data using the semi-automatic method and a mean bias of -8.10 (9.93) and -11.95 (11.83) was reported between the manual and semi-automatic method. As the mean bias for both assessors were within the LOA defined for inter-observer reliability assessment of both the Sherwin study (-13.67 to 19.71) and our study (-16.99 to 9.29), we find that this is an acceptable variation between the two methods considering the circumstances and the lack of a gold standard.

CUVAF quantification is difficult due to variability in the shape, size and intensity of the detected changes. Furthermore, reflections off the surface of the eye, eye conditions (e.g. hemorrhages, nevi) or debris on camera lenses create artifacts in the image that can obscure or be mistaken for an area of autofluorescence. These aberrations, in addition to poor image acquisition techniques result in unfocussed, decentered or low contrast images which contribute to the difficulty of accurate CUVAF quantification. Ideally, measurements obtained using the semi-automatic method would match the manual method as this data has already been published and used in a number of studies. However, there is no clear definition of what constitutes a CUVAF region and no gold standard for CUVAF measurements which would be necessary for future comparative studies.

The semi-automatic method allows users to flag images as ‘Subjective’ or ‘Weak CUVAF’ to help identify images susceptible to highly variable measurements. Subjective images are defined as those where CUVAF is present, but the border of the region is difficult to define due to factors
such as flash artifacts obscuring the region, or regions with diffuse edges. A ‘Weak CUVAF’
region was defined as one where CUVAF appears to be present, but the signal is so weak such
that it could be mistaken for uneven illumination or a camera artifact, especially when the image
is not in focus. Having these options allow these images to be included/excluded in data analysis,
or to be re-analyzed by one or more assessors to reach a consensus on the true area of CUVAF.

Furthermore, users can write comments on the image which can be used to characterize the
appearance of the CUVAF region. Both qualitative and quantitative data obtained from this
computerized analysis can help to establish a universally accepted criterion for characterizing
CUVAF and work towards developing a standard method for CUVAF analysis across the globe.

**Conclusions**

The aim of this study was to validate whether the proposed semi-automatic method could
replace the manual method for CUVAF analysis. Although it has been previously established
that the measurement of CUVAF using the manual method is reliable [2], this method is
subjective and tedious and results cannot be retrospectively reviewed. Having a semi-automatic
method can provide a more efficient means of quantifying CUVAF that is less prone to human
error, with reduced subjectivity and enhanced reproducibility. The length of time taken to
perform manual measurements is onerous, requiring many hours or even days of researcher time
that could be spent on other research activity. The time-consuming nature of manual
measurement also effectively limits the number of study participants that can be analyzed,
making the technique impractical for very large population-based cohorts. A semi-automatic
method of measurement reduces the time and cost burden of the technique, and potentially
allows analysis of very large cohorts As both assessors achieved over 95% agreement with the
manual measurements, we conclude that the two methods agree sufficiently well for them to be
used interchangeably, and that the semi-automatic method is a valid tool for assessment of ocular
sun exposure in the research setting [12]. Furthermore, the semi-automatic method for CUVAF
analysis also shows high reliability, with both intra- and inter-observer agreement over 95%. An
additional benefit of the computer program is that all steps of the analysis process, including the
outline of the area of fluorescence, are saved and can be retrospectively reviewed at any time.
This will be of high importance for characterizing the CUVAF region for defining a standard
inclusion criterion as well as for future studies that may want to investigate the implications of
CUVAF intensity.

Further work on methods for CUVAF analysis would be focused on defining a universally
accepted inclusion criteria and develop a fully automated algorithm to remove the subjectivity
involved in evaluating CUVAF images. Data obtained from the manual CUVAF analysis of the
Raine Study only consisted of the measured CUVAF region for each image without a
corresponding marked image to show the CUVAF border. With both area measurements and
marked images obtained from this study, a machine learning technique can be utilized to try to
fully automate the analysis of CUVAF images.

Determining accurately and objectively lifetime sun exposure is fundamental to research into
the detrimental and beneficial effects of ultraviolet radiation on human health. This semi-
automatic method for analyzing pre-clinical ocular sun damage is efficient and enables
retrospective review of images, giving it great potential for large cohort and interventional
studies.

Acknowledgements

We are grateful to the Raine Study participants and thank the Raine Study and Lions Eye
Institute research staff for cohort coordination and data collection. We would like to
acknowledge ongoing support from Professor Minas Coroneo and his group from the University of New South Wales. The Raine Study 20-year follow-up was supported by the National Health and Medical Research Council (1021105) and The Ophthalmic Research Institute of Australia grants awarded to David Mackey. Ajmal Mian was supported the Australian Research Council (DP110102399).

References


Emily Huynh completed a Master of Engineering (Biomedical) at the University of Melbourne. Currently, she works at the Lions Eye Institute as a researcher assistant where she is involved with developing software for image and data analysis. Her research interests are image processing, computer vision and machine learning for medical applications.

Danuta Bukowska received her PhD at the Nicolaus Copernicus University (NCU) in Torun, Poland under the supervision of Professor Maciej Wojtkowski, a leading researcher in optical coherence tomography. In October 2014, she has been employed full time as a Research Associate in the Imaging & Informatics Group and the Lions Eye Institute (LEI) in Perth, Western Australia. The main objective of her present work is design and development of methodology for data visualization and analysis.

Seyhan Yazar completed her PhD at the University of Western Australia in Perth, Australia under the supervision of Professor David A Mackey and Associate Professor Alex W Hewitt. Her main research focus has been the genetic and environmental factors in common, complex eye diseases such as myopia, corneal astigmatism, and keratoconus. In 2016, she received the Australian National Health and Medical Research Council’s CJ Martin Overseas Biomedical Fellowship to incorporate new bioinformatics methods into eye research.
Charlotte McKnight is a trainee ophthalmologist in the Western Australian training network of the Royal Australian and New Zealand College of Ophthalmologists. She has completed a Masters of Medical Science at the Centre for Ophthalmology and Visual Science at the University of Western Australia.

Ajmal Mian is a Research Fellow in the School of Computer Science and Software Engineering at the University of Western Australia. His research interests include computer vision, face and 3D object recognition, and multispectral image analysis.

David Mackey is Managing Director of the Lions Eye Institute and Professor of Ophthalmology and Director of the Centre for Ophthalmology and Vision Science at the University of Western Australia. Professor Mackey has achieved international recognition as a genetic ophthalmologist. His original research into the genetics of glaucoma and in the fields of optic atrophy and congenital cataract has received continued research funding support for the past two decades.