

## Validation of a clinical assessment tool for cicatrising conjunctivitis

Ong, Hon Shing; Minassian, Darwin; Rauz, Saaeha; Mehta, Jodhbir S.; Dart, John K.

DOI:

[10.1016/j.jtos.2019.10.010](https://doi.org/10.1016/j.jtos.2019.10.010)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Ong, HS, Minassian, D, Rauz, S, Mehta, JS & Dart, JK 2020, 'Validation of a clinical assessment tool for cicatrising conjunctivitis', *Ocular Surface*, vol. 18, no. 1, pp. 121-129. <https://doi.org/10.1016/j.jtos.2019.10.010>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



## Original Research

## Validation of a clinical assessment tool for cicatrising conjunctivitis

Hon Shing Ong<sup>a,b,c,d,e,\*</sup>, Darwin Minassian<sup>f</sup>, Saaeha Rauz<sup>g,h</sup>, Jodhbir S. Mehta<sup>a,b,e,i</sup>,  
John K. Dart<sup>c,d</sup>



<sup>a</sup> Corneal and External Diseases Department, Singapore National Eye Centre, Singapore

<sup>b</sup> Tissue Engineering and Stem Cell Group, Singapore Eye Research Institute, Singapore

<sup>c</sup> Corneal and External Diseases Department, Moorfields Eye Hospital NHS Foundation Trust, United Kingdom

<sup>d</sup> National Institute for Health Research (NIHR) Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, United Kingdom

<sup>e</sup> Duke-NUS Graduate Medical School, Singapore

<sup>f</sup> EpiVision Ophthalmic Epidemiology Consultants, United Kingdom

<sup>g</sup> Academic Unit of Ophthalmology, Institute of Inflammation and Ageing, University of Birmingham, United Kingdom

<sup>h</sup> Birmingham and Midland Eye Centre, Sandwell and West Birmingham Hospitals NHS Trust, United Kingdom

<sup>i</sup> School of Material Science and Engineering, Nanyang Technological University, Singapore

## ARTICLE INFO

## Keywords:

Cicatrizing conjunctivitis

Conjunctiva

Cornea pathology

Fornix depth measurer (FDM)

Grading tool

Measurements

Mucous membrane pemphigoid

Stevens-Johnson syndrome

## ABSTRACT

**Purpose:** This study was designed to validate a semi-quantitative clinical assessment tool for cicatrising conjunctivitis (CC).

**Methods:** Fifty-five patients (109 eyes) with mucous membrane pemphigoid (MMP) and 31 patients (61 eyes) with Stevens-Johnson syndrome (SJS) were included. Three methods were used for validation: (1) comparison of inter- and intra-observer reproducibility for the components selected for the initial version of the tool, (2) quantitative measurement of the scarring component with a fornix depth measurer, compared with qualitative Tauber grading methodology, (3) the final version of the tool was compared with the published Sotozono SJS grading system. Main outcome measures included: inter- and intra-observer reproducibility, calculation of composite measures of scarring and morbidity, component redundancy, and correlation with other grading systems.

**Results:** Inter- and intra-observer agreement was moderate-to-excellent for graded components of conjunctival hyperaemia, upper and lower symblepharon, upper and lower fornix depth, corneal vascularisation, and corneal opacity. There was poor-to-good agreement for limitation of motility which was rejected from inclusion in the final tool. Composite scores for scarring components and morbidity components showed good-to-excellent agreement and distribution of ocular disease severity. Analysis of the composite components showed no redundancy - all components contributed independently. Comparison with both Tauber and Sotozono grading methodologies showed good concordance.

**Conclusions:** This study has developed the first validated assessment tool applicable to causes of CC. The tool is concise and discriminates patients with varying disease severity. It measures both disease activity and severity and is suitable for clinical and research applications.

## Introduction

‘Cicatrising conjunctivitis’ (CC) is a term used to describe a heterogeneous group of disorders that cause inflammation and scarring within the conjunctiva [1]. In countries with a predominantly Caucasian population, such as the United Kingdom, ocular mucous membrane pemphigoid (MMP) is the commonest cause of CC [2]. Worldwide, other important causes of CC include Stevens-Johnson syndrome (SJS),

trachoma, atopic keratoconjunctivitis, mechanical or chemical injuries, ocular rosacea, and Sjogren's syndrome [1,2].

Much of the research into the therapy of CC is focused on the development of both anti-inflammatory therapies [3,4] and, more recently, anti-scarring therapies [1]. Anti-inflammatory therapy is currently the mainstay of treatment, for which adjunctive conjunctival/mucosal anti-scarring therapies are at present unavailable but are being researched [5,6]. Development of new therapies, as well as normal

\* Corresponding author. Singapore National Eye Centre, 11 Third Hospital Avenue, 168751, Singapore.

E-mail address: [honshing@gmail.com](mailto:honshing@gmail.com) (H.S. Ong).

<https://doi.org/10.1016/j.jtos.2019.10.010>

Received 9 June 2019; Received in revised form 19 October 2019; Accepted 31 October 2019

1542-0124/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

clinical practice, ideally requires reproducible inter- and intra-observer agreement measurement of scarring, inflammation, and morbidity (the sight threatening corneal complications of the disease), using a concise clinical assessment tool. These measurements are essential to identify disease progression and the effects of treatments, both for routine clinical practice and for clinical trials [7]. Currently, no validated tool exists for these purposes although the requirement for one has been recognised by the development of 11 methodologies over the last four decades. Most of these have focused on the evaluation of conjunctival scarring in MMP [8–14]. These ocular scoring systems have been recently reviewed; the lack of a validated system and their failure to provide assessments for disease activity as well as damage were highlighted [7,15]. Three systems include simple qualitative grading of conjunctival inflammation,[16,17] although one of these does not include any damage indices [18]

A method for the quantification of scarring associated with glaucoma medications [19] has been subsequently modified and validated for use in CC in general [12]. A further system, developed for the assessment of SJS with ocular involvement, incorporates 13 graded components to measure corneal disease, conjunctival inflammation, and eyelid damage [20]. This system of evaluating the chronic ocular complications of SJS has been further expanded recently [21]. Only two of these 11 systems, measuring conjunctival scarring using different quantitative methods, have had both inter- and intra-observer variability evaluated and have been compared with another methodology. One has not been used in any studies [13] whereas the other [12], a fornix depth measurer (FDM), has been used to evaluate the normal upper and lower fornix depths in both healthy South-Asian [22] and Caucasian populations [23]. The latter has also been used to assess the progression of conjunctival scarring in ocular MMP [24].

This study was designed to develop a concise, validated, semi-quantitative clinical severity assessment tool for CC. This tool incorporated the measures of inflammation (disease activity), scarring (disease progression), and morbidity (the sight-threatening effects of the disease) that are required for diagnosis, measurement of disease progression, and monitoring responses to treatment, and to have the sensitivity to discriminate between small differences in these. Our hypothesis was that a much simpler scoring system, closer in concept to that of the five-component trachoma scoring system [25] but with more subtle grading potential, could be achieved with fewer components than those demanded by existing grading systems. This hypothesis was tested through an initial study aimed at determining the inter- and intra-observer agreement of the components of a new clinical severity assessment tool. This was followed by an analysis that was used to eliminate components that could not be reproducibly measured. These data were used to develop a concise CC clinical assessment tool. This was further validated by comparison to previously described grading systems and then prepared in a form ready for use in clinical practice and clinical trials.

## Methods

This was a prospective cross-sectional study on a cohort of patients diagnosed with ocular MMP and SJS. The study protocol was approved by the UK National Health Service Research Ethics Committee (REC Reference 16/LO/0674) and the Singhealth Centralized Institutional Review Board (CIRB Reference 2017/2270). Informed consent was obtained from the participants after explanation of the nature of the study. The study adhered to the tenets of the Declaration of Helsinki. Patients diagnosed with MMP and SJS with ocular involvement were identified from databases of existing patients under the clinical care of the Corneal and External Diseases department at Moorfields Eye Hospital NHS Foundation Trust and the Singapore National Eye Centre respectively.

Patients recruited into the study received two appointments to evaluate the inter- and intra-observer agreements of two clinical

observers (JKD and HSO) using the clinical assessment tool. JKD is an experienced Ophthalmologist with over 30 years of clinical practice in the field of corneal and external diseases; at the time of this study, HSO was a junior Ophthalmologist in a training fellowship in corneal and external diseases. Patients were examined by both observers at both visits. The inter-observer interval at each visit was less than 10 min. To avoid changes in clinical status or treatment regimen that could affect the scores between the two visits, intervals between the two visits for all patients were within two weeks. Each clinical examiner performed the assessments independently and was masked to the severity scores graded by the other clinical examiner. To reduce bias, the severity scores graded by the examiners in the first clinical visit were not available to the examiners at the second visit.

### *Designing the clinical assessment tool for cicatrising conjunctivitis*

The design of the clinical assessment tool was based upon both the evaluation of the clinical outcome, reporting from the studies described above, together with previously described disease activity and damage indices for ocular surface diseases [7] and the findings of our cross-sectional and longitudinal MMP studies [24,26]. It comprised 12 components in three functional categories: a) *grading of inflammation* (bulbar conjunctival hyperaemia, limbitis), b) *grading of scarring* (subconjunctival fibrosis, limitation in ocular motility, upper and lower fornix symblephara, upper and lower fornix depth measure), and c) *grading of ocular morbidity*, which are the effects of inflammation and scarring (distichiasis, conjunctival and corneal keratinisation, corneal vascularisation, corneal opacity).

These 12 components were chosen from the 23 components, in 6 anatomical categories, that were used in our cross-sectional [26] and longitudinal studies [24], which in common with the Francis [17] and Sotozono [20,21] systems are the only studies that have graded inflammation, scarring, and ocular surface damage. To facilitate the sequence and the flow of the examination in the conduct of this validation study, the tool was converted into a case report form with guidelines for use (Supplementary Appendix 1).

Following data analyses, a finalized concise clinical assessment tool for CC was established (Supplementary Appendix 2). In this finalized tool, components in the original tool with inadequate levels of agreement were excluded; although components that were considered important in the assessment and prognosis in some CC cases, were retained.

Although this study used ocular MMP as the index disease, the tool has been designed to work with other causes of CC. The final version of the tool was evaluated in SJS, the commonest cause of CC in countries with predominantly Asian populations and compared with the Sotozono chronic SJS grading tool [20]. For this study 31 patients (61 eyes) having SJS were graded using both the final version of our CC assessment tool and the Sotozono grading tool.

### *The assessment of upper and lower central fornix depths*

The upper and lower central fornix depths were measured using our custom-made FDM validated in a previous study [12] and previously used to establish normal values for the upper and lower fornices in healthy Caucasians [23] and South-East Asians [22] (Supplementary Fig. 1; supplementary video). The guidance for the use of the FDM is found in our finalized concise clinical assessment tool for CC (Supplementary Appendix 2). Fornix depth measurements were recorded to the nearest 1 mm.

Supplementary video related to this article can be found at <https://doi.org/10.1016/j.jtos.2019.10.010>.

To assess the extent of forniceal foreshortening as a marker of severity of conjunctival scarring, the central fornix depths measured were converted into percentage foreshortening. Published normal age-specific values of mean depths of normal fornices were used to calculate

**Table 1**  
Characteristics of the study population.

	MMP (n = 55)	SJS (n = 31)
<b>Age of diagnosis in years</b> (range [R], median [M], interquartile range [IQR])	[R] 38–86, [M] 70.0 [IQR] 60 - 75	[R] 19–84, [M] 53.0 [IQR] 34 - 74
<b>Females</b>	29/55 (52.7%)	21/31 (67.7%)
<b>Race</b>		
White (British, Irish, other white)	46/55 (83.6%)	0/31 (0%)
Chinese	0/55 (0%)	23/31 (74.2%)
Asian or Asian British (Indian, Pakistani, Bangladeshi, any other South Asian background)	6/55 (10.9%)	2/31 (6.5%)
Malay	0/55 (0%)	4/31 (12.9%)
Black or Black British (Caribbean, African, other black)	1/55 (1.8%)	0/31 (0%)
Other ethnic group	2/55 (3.6%)	2/31 (6.4%)
<b>Previous fornix reconstruction surgery</b>		
Yes	10/109 (9.2%)	2/61 (3.3%)
No	99/109 (90.8%)	59/61 (96.7%)
<b>Previous lid surgery</b>		
None	69/109 (63.3%)	56/61 (91.8%)
Upper lid entropion repair	11/109 (10.1%)	0/61 (0%)
Lower lid entropion repair	15/109 (13.8%)	0/61 (0%)
Upper lid and lower lid entropion repair	9/109 (8.3%)	0/61 (0%)
Trichiasis surgery	9/109 (8.3%)	4/61 (8.3%)
Ptosis surgery	2/109 (1.8%)	0/61 (0%)
Tarsorrhaphy	1/109 (0.9%)	1/61 (0%)
<b>Vision score</b>		
6/7.5 or better	51/109 (46.8%)	26/57 (45.6%)
6/9–6/12	37/109 (33.9%)	14/57 (24.6%)
6/18–6/36	9/109 (8.3%)	9/57 (15.8%)
6/60–2/60	1/109 (0.9%)	1/57 (1.8%)
3/60 – counting fingers only	5/109 (4.6%)	2/57 (3.5%)
Hand movements only	3/109 (2.8%)	5/57 (8.8%)
Perception of light only	1/109 (0.9%)	0/57 (0%)
No perception of light	2/109 (1.8%)	0/57 (0%)
<b>Ocular co-morbidities</b>		
None	60/109 (55.0%)	49/61 (80.3%)
Glaucoma/ocular hypertension	10/109 (1.7%)	2/61 (3.3%)
Pseudophakia	30/109 (27.5%)	5/61 (8.2%)
Corneal graft	5/109 (4.6%)	5/61 (8.2%)
Other corneal pathologies	15/109 (13.8%)	1/61 (1.6%)
Retinal diseases	9/109 (8.3%)	2/61 (3.3%)
Uveitis	2/109 (1.8%)	0/61 (0%)
Previous glaucoma surgery	1/109 (0.9%)	0/61 (0%)
Previous vitrectomy	2/109 (1.8%)	0/61 (0%)

MMP Mucous membrane pemphigoid; SJS Stevens-Johnson Syndrome.

the percentage foreshortening of the central fornix depths [23]. To obtain age-adjusted normal values of fornix depths, regression lines were fitted onto published age-stratified values; regression equations were then used to predict the normal values for each age. Thus, fornix depth percentage foreshortening = [(fornix depth measured – age-adjusted normal fornix depth)/age-adjusted normal fornix depth] x 100%. To facilitate reckoning of fornix depth percentage foreshortening during clinical assessment, the percentage fornix depth foreshortening based on age-adjusted normative values were computed into reckoner tables (Supplementary Appendix 2).

#### Statistical methods

The statistical software used was Stata v.14.2 (StataCorp, Texas, USA).

#### Inter-observer and intra-observer agreement levels

The kappa statistic was estimated for binary clinical data, and the weighted kappa for ordinal data with more than two categories. For weighted kappa the linear weights were defined as  $1 - |k-j|/|q_{\max} - q_{\min}|$ , where k and j refer to the actual ratings and  $q_{\max}$  and  $q_{\min}$  are the maximum and minimum of all ratings. The intra-class correlation coefficient (ICC) was used for continuous data, and also for ordinal data with > 5 categories. A two-way mixed-effects model was used for estimating ICC, treating eyes (targets) as random, and observers (raters) as fixed. Other specifications for the model included the following: a)

*unit of analysis* used was individual ratings for the eye, and inter-observer and intra-observer agreements were between individual ratings rather than averages; b) *absolute agreements* were calculated where ratings match exactly, rather than *consistency of agreements* where ratings differed by constant amounts (i.e. perfect correlation, but agreement may be poor). 95% confidence intervals (CI) were calculated for both kappa statistics and ICC.

#### Interpretation of agreement levels

The levels of agreement were interpreted from point estimates of agreement coefficients (kappa statistic and ICC) and their respective 95% confidence bounds. Values < 0.55 were considered 'poor agreement'; 0.55 to 0.75 considered 'moderate agreement'; 0.76 to 0.90 considered 'good agreement'; > 0.90 considered 'excellent agreement'.

#### Composite scores

Composite scores were split into two main categories: a) grading of scarring and b) grading of morbidity (effects of inflammation and scarring). Composite scores were calculated as the sum of scores of different individual components divided by the number of components within each category. The decision to include an element in the composite score was made based on the interpretation of the level of inter-observer and intra-observer agreement observed. To obtain the composite scores, the scores of all components included were transposed on a percentage scale of 0–100. An *overall composite score* was calculated by adding the scores of all components in both categories and dividing

the sum by the total number of components.

To evaluate the internal consistency (reliability) of the composite score, Cronbach's alpha was estimated together with correlations between the composite score and each of its constituent components, and between each constituent component and the remaining components. Very high values of Cronbach's alpha could indicate some redundancy in the constituent components when one member of a pair of components is highly predictive of the other member. This was also explored.

*Comparison between two clinical severity-grading tools*

Correlation between the composite scores from our CC clinical assessment tool and the total score from the Sotozono grading tool was evaluated by calculating the Pearson correlation coefficient. Bland-Altman plots were used to assess the limits of agreement between the two grading tools.

**Results**

Patient characteristics are summarized in Table 1.

**MMP patients.** A total of 55 patients (109 eyes) with a diagnosis of ocular MMP were recruited. Two patients were unable to attend the second assessment visit. Of these, one patient had a corneal epithelial defect following the first clinical visit, which may have resulted from topical anesthetic toxicity and/or trauma when measuring the fornix depth. The median age of patients was 70.0 years (interquartile range (IQR) 60–75); 29/55 (52.7%) of patients were female; 46/55 (83.6%) were of Caucasian origin. In this series of MMP patients, 34/55 (61.8%) were direct immunofluorescence (DIF) positive and 20/55 (36.4%) were DIF negative. DIF investigation was not performed in one patient. The diagnosis of MMP in those without a positive DIF status was based on our previously described diagnostic criteria of 'DIF negative MMP' [26]

**SJS patients.** Thirty-one patients (61 eyes) with a diagnosis of SJS with ocular involvement were recruited having a median age of 53.0 years (IQR 34–74); 21/31 (67.7%) of patients were female; 23/31 (74.2%) were of Chinese origin.

*Inter-observer and intra-observer agreement*

Table 2 lists the inter-observer and intra-observer point estimates of agreement coefficients for each individual component of the severity-grading tool and their corresponding levels of agreement. Of the 12 components in the three categories of the assessment tool, moderate to excellent levels of agreements were found for the following 7 components:

- a) *Inflammation category:* bulbar conjunctival hyperaemia
- b) *Scarring category:* lower fornix depth percentage foreshortening, upper fornix depth percentage foreshortening, lower fornix symblepharon score, and upper fornix symblepharon score
- c) *Morbidity category:* corneal vascularisation and corneal opacity grading

These parameters were included in the final version of the CC assessment tool (Supplementary Appendix 2).

Due to the wide confidence intervals of the agreement levels for these components, inadequate levels of agreement were observed for the following 2 components: limitation in motility (scarring category) and ocular surface keratinisation (morbidity category). Limitation in motility (scarring category) was left out of the final tool. Ocular surface keratinisation (morbidity category) however, was retained (but not scored). The tool includes the identification of conjunctival keratinisation, lid margin/tarsal conjunctival keratinisation, and corneal keratinisation. The presence of ocular surface keratin has been shown in previous studies to be an important marker of severe dry eyes and overall disease severity in ocular surface disorders [17,20,27,28].

**Table 2**  
Inter-observer and intra-observer agreements of the individual components of the clinical severity-grading tool.

Component	Statistic <sup>a</sup>	Point estimate	95% CI		Agreement level
<b>INFLAMMATION</b>					
<b>Bulbar conjunctival hyperaemia</b>					
Between Observer	ICC	0.88	0.84	0.90	Good
Within Observer	ICC	0.76	0.69	0.81	Moderate to Good
<b>Limbitis</b>					
SCARRING					
<b>Limitation in motility</b>					
Between Observer	kappa	0.49	0.23	0.74	Poor
Within Observer	kappa	0.72	0.51	0.93	Poor to Good
<b>Lower fornix symblepharon</b>					
Between Observer	kappa	0.71	0.63	0.79	Moderate to Good
Within Observer	kappa	0.74	0.60	0.82	Moderate to Good
<b>Upper fornix symblepharon</b>					
Between Observer	kappa	0.68	0.59	0.78	Moderate to Good
Within Observer	kappa	0.66	0.55	0.77	Moderate to Good
<b>Lower fornix depth percentage foreshortening (%)<sup>b</sup></b>					
Between Observer	ICC	0.93	0.91	0.95	Excellent
Within Observer	ICC	0.94	0.92	0.95	Excellent
<b>Upper fornix depth percentage foreshortening (%)<sup>b</sup></b>					
Between Observer	ICC	0.89	0.81	0.93	Good to Excellent
Within Observer	ICC	0.92	0.89	0.94	Good to Excellent
<b>Subconjunctival fibrosis</b>					
MORBIDITY (EFFECTS OF INFLAMMATION AND SCARRING)					
<b>Distichiasis</b>					
Not analysed (only two eyes affected in this study)					
<b>Ocular surface keratinisation</b>					
Between Observer	kappa	0.59	0.43	0.75	Poor to Good
Within Observer	kappa	0.64	0.49	0.79	Poor to Good
<b>Corneal vascularisation</b>					
Between Observer	kappa	0.76	0.71	0.82	Moderate to Good
Within Observer	kappa	0.80	0.75	0.85	Moderate to Good
<b>Corneal opacity</b>					
Between Observer	ICC	0.84	0.79	0.87	Good
Within Observer	ICC	0.89	0.86	0.92	Good to Excellent

CI = confidence interval, ICC = Intraclass correlation coefficient.

<sup>a</sup> Kappa statistic was estimated for ordinal data with ≤ 5 categories and ICC for continuous data or ordinal data with > 5 categories.

<sup>b</sup> Fornix depth percentage foreshortening = (Normal age-adjusted fornix depth – measured fornix depth)/(Normal age-adjusted fornix depth) x 100.

Furthermore, in SJS the presence of lid margin keratinisation is thought to be a causative factor for poor visual prognosis [27,29]. Three further components, limbitis (inflammation category), subconjunctival fibrosis (scarring category), and distichiasis (morbidity category), were also not scored but have been retained in the final assessment tool because of their clinical importance in the assessment of individual cases. Limbitis is an important prognostic sign when present in severe conjunctival inflammation but was absent in our subjects. Subconjunctival fibrosis was identified in 98% of cases and, as a result, not analysed but has been retained as being important in the diagnosis of early CC. Distichiasis is relatively rare in MMP, used as the index disease for inter- and intra-observer evaluation of the tool in this study, but is an important sign of ocular surface scarring, particular in SJS.

This gives a total of 11 components used in the final assessment tool (Supplementary Appendix 2).

Composite scores

Visible inflammation in CC is a marker of ‘activity’ and is thus variable, and reversible. This has not been included in a composite scoring for conjunctival scarring, but has been included in the final tool as an index of reversible disease activity for monitoring response to immunosuppressive treatment. The aim of cicatrising conjunctivitis composite scores are to fully describe the scarring and morbidity categories of the assessment tool, both of which were represented by more than one component. Of the seven components with moderate-to-excellent agreement retained in the final version of the assessment tool, six components were used to calculate composite scores.

Composite score – grading of scarring

This comprised four components: upper fornix symblepharon score, lower fornix symblepharon score, upper fornix depth percentage foreshortening, and lower fornix depth percentage foreshortening. The composite score for the grading of scarring was calculated as a sum of all the following scores divided by 4. Both upper and lower symblepharon scores of 0–2 were transposed into a 0–100 scale by [symblepharon score] x 100/2. The upper and lower fornix depth percentage foreshortening were already in a 0–100 scale so that no transposition was required. There was excellent inter-observer agreement (ICC 0.93, 95% CI 0.91–0.94) and intra-observer agreement (ICC 0.93, 95% CI 0.91–0.94) for the composite scarring grading. This is summarized in Table 3.

Composite score – grading of morbidity

This comprised of two components: the corneal vascularisation score and corneal opacity scores. The composite score was calculated as a sum of the following scores divided by 2. The corneal vascularisation grading scores of 0–5 were transposed into a 0–100 scale by [corneal vascularisation grading score] x 100/5. The corneal opacity grading scores of 0–9 were transposed into a 0–100 scale by [corneal opacity grading score] x 100/9. There was good to excellent inter-observer agreement (ICC 0.92, 95% CI 0.89–0.94) and excellent intra-observer agreement (ICC 0.95, 95% CI 0.93–0.976) for the composite score on grading of morbidity. This is summarized in Table 3.

Composite score – combined scarring and morbidity categories

This combined score has the potential to provide one number that can quantify the largely irreversible effects of the disease both for routine clinical follow-up and in prospective studies. This combined score was evaluated to establish whether a score for any one of the six constituent components might predict the score of another indicating

redundancy of one or more components. The results of this analysis are described in Table 4, which summarizes the results of inter-component correlations, correlations between components of the two main categories (scarring and morbidity) and the overall composite score, and the internal consistency of the overall composite score. The component-rest of components correlation is an assessment of redundancy. The correlations between each component and the remaining components were poor to moderate, between  $r = 0.59$  (lower fornix symblepharon) and  $r = 0.73$  (corneal vascularisation score). This assessment of redundancy indicates that, when pairs of components were compared, one component does not have the potential to adequately predict the other such that each of the components contributes important additional information to the overall severity-grading tool. This was confirmed by the good internal consistency of the overall composite score (Cronbach's alpha = 0.845), which could not be improved by dropping any one of the 6 components; as a result, all these components were retained in the final CC assessment tool.

Frequency distributions of inflammation and composite grading for MMP ocular disease

Fig. 1 describes the distribution of the 109 eyes in the validation study by their inflammation and composite scores.

Practical examples and applications

Table 5 shows examples of the practical use of the CC assessment tool for patients from this series [26]. Supplementary Appendix 2 provides the finalized CC assessment tool and guidelines for its use in clinical and research settings. It also provides ready reckoners to identify age adjusted percentage reduction for both upper and lower fornix depth measurements. Supplementary Appendix 3 is a spreadsheet calculator, which provides individual and composite scores for the three functional categories of inflammation, scarring and morbidity; it is envisaged that this calculator will be used for calculation in clinical studies rather than for clinical use. It should be noted that the components of limbitis, subconjunctival fibrosis, distichiasis, and ocular surface keratinisation are to be recorded because of their diagnostic or prognostic importance in the finalized CC assessment tool but not scored. This has been done because there was poor inter- and intra-observer variability for keratinisation whereas for the other signs, the inter- and intra-observer agreements could not be assessed. Users of the tool may decide to score these components in clinical practice or research. However, in the absence of new data, the user must understand that the scoring of these components may not be reproducible by another observer.

Table 3

Inter-observer and intra-observer agreements of the composite scores in each category: scarring, morbidity (effects of inflammation and scarring), and overall composite score.

Composite Scores	Statistic	Point estimate	95% CI		Agreement level
i) Scarring score (%) <sup>a</sup>					
Between Observer Agreement	ICC	0.93	0.91	0.94	Excellent
Within Observer Agreement	ICC	0.93	0.91	0.94	Excellent
ii) Morbidity score (%) <sup>b</sup>					
Between Observer Agreement	ICC	0.92	0.89	0.94	Good to Excellent
Within Observer Agreement	ICC	0.95	0.93	0.96	Excellent
Overall Composite (Scarring & Morbidity) score (%) <sup>c</sup>					
Between Observer Agreement	ICC	0.95	0.93	0.96	Excellent
Within Observer Agreement	ICC	0.96	0.94	0.97	Excellent

CI = confidence interval, ICC = intraclass correlation coefficient.

<sup>a</sup> Composite score for scarring comprises of a) lower fornix depth percentage foreshortening, b) upper fornix depth percentage foreshortening, c) percentage lower fornix symblepharon, and d) percentage upper fornix symblepharon.

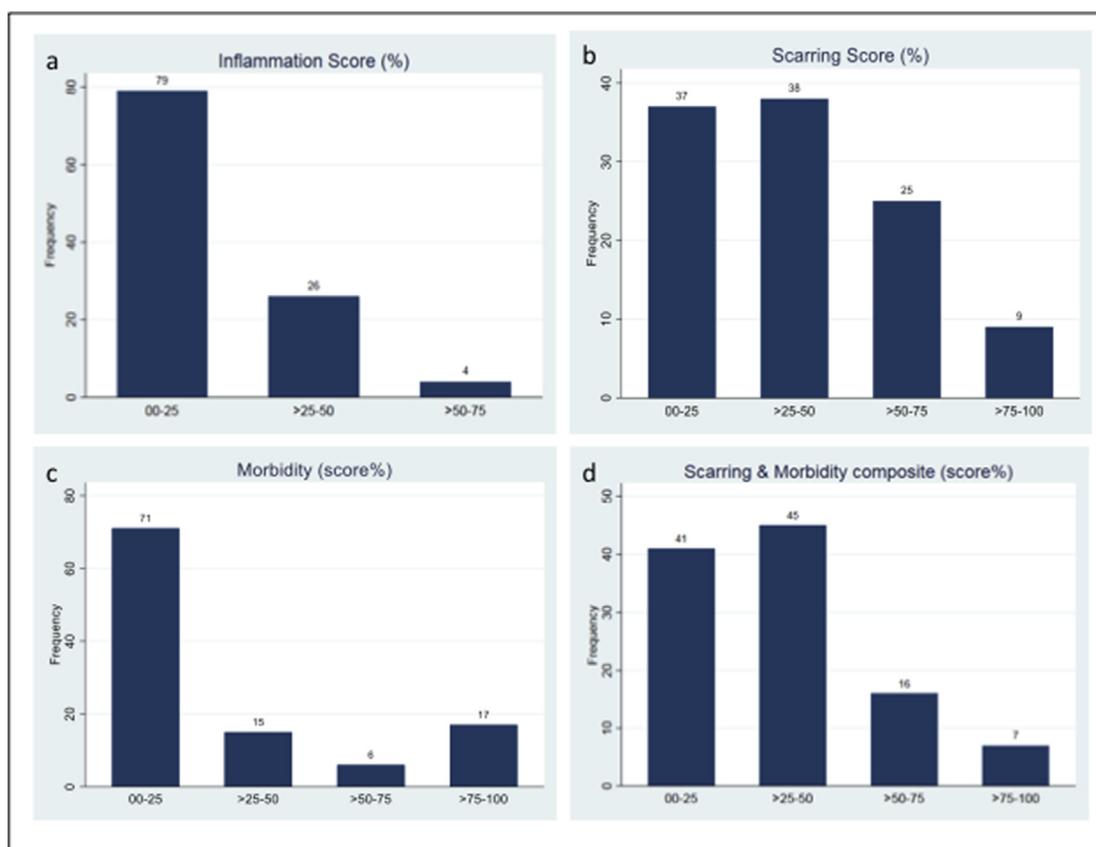
<sup>b</sup> Composite score for morbidity (effects of inflammation and scarring) comprises of a) corneal vascularisation, and b) corneal opacity.

<sup>c</sup> Overall composite score comprising of all elements in the scarring and morbidity composite scores.

**Table 4**  
Inter-component correlation, component-composite correlation, and internal consistency of overall composite scores of scarring and morbidity categories.

Components of score	Observations	Component-Composite Correlation <sup>a</sup>	Component-rest of components correlation <sup>b</sup>	Average inter-component covariance <sup>c</sup>	Cronbach's alpha <sup>d</sup>
Lower fornix symblepharon	109	0.718	0.585	470.846	0.827
Upper fornix symblepharon	109	0.761	0.677	491.416	0.818
Lower fornix depth percentage foreshortening	109	0.755	0.612	437.755	0.824
Upper fornix depth percentage foreshortening	109	0.713	0.596	487.069	0.826
Corneal vascularisation	109	0.832	0.729	404.900	0.798
Corneal opacity	109	0.763	0.617	428.765	0.824
<b>Overall Composite score</b>				<b>453.456</b>	<b>0.845</b>

<sup>a</sup> Correlation between the component and the composite score.  
<sup>b</sup> Correlation between the component and all the remaining components combined.  
<sup>c</sup> Average covariance when the component is dropped in calculation of composite score.  
<sup>d</sup> Chronbach's alpha value when the component is dropped in calculation of composite score.



**Fig. 1.** Frequency distribution of composite scores a) Inflammation score b) Scarring score c) Morbidity score d) Composite score combining scarring and morbidity scores.

*Comparison of the CC assessment tool with the Sotozono grading system for chronic SJS*

These were compared in a cohort of 31 patients with ocular involvement by SJS. Good correlation was found between the composite score of our CC assessment tool and the total Sotozono score [Pearson r 0.93 p < 0.001] [Fig. 2a]. No significant proportional bias was also observed when Bland-Altman plots were applied [Fig. 2b].

**Discussion**

This study has developed an assessment tool for CC. It has been

developed from previously described tools for ocular MMP and SJS (compared in Supplementary Table 1) and then modified following the calculation of inter- and intra-observer levels of agreement, so as to include only those components that can be reproducibly scored and which show no redundancy (the presence of one does not predict the presence or severity of another). Instead of using anatomical categories, this tool focuses on three functional categories of inflammation (an indicator of disease activity) together with scarring and morbidity (both indicators of disease damage) that result from CC. This is unlike most previous assessment tools that have focused on scarring alone [15]. These functional categories are important measures needed for both clinical assessments and for clinical trials. The different categories of

**Table 5**  
Examples of the use of the cicatrising conjunctivitis grading tool for individual patients.

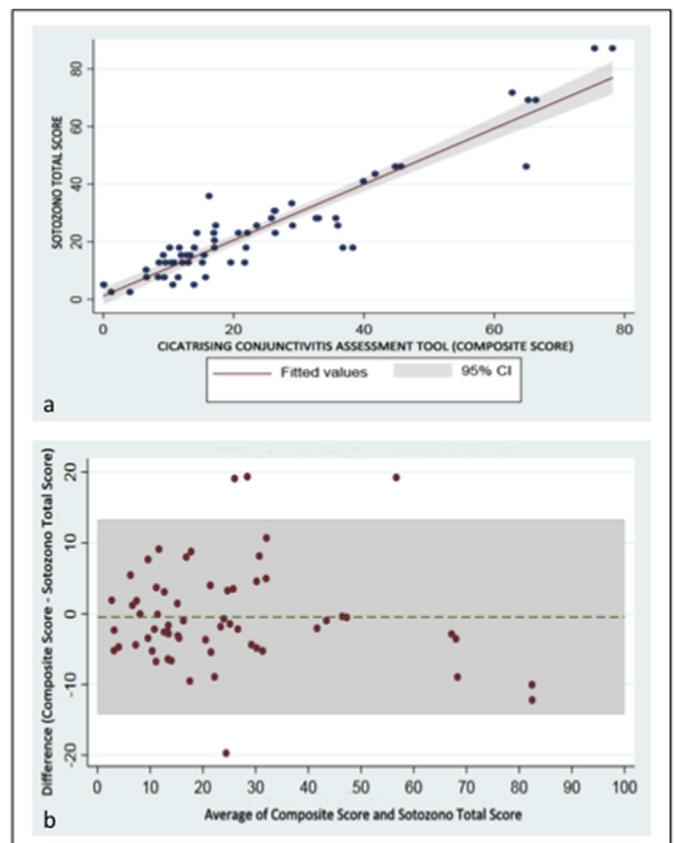
Patient study #1 Aged 86, Right Eye				
Categories	Components	Raw value	Transposition	Percentage
<b>Inflammation</b>	1 Inflammation	3	3/16 x 100	18.8
<i>Inflammation score</i> (total of percentage value)				18.8%
Limbitis: No				
<b>Scarring</b>	2 Lower fornix depth	2		80.2%
	3 Upper fornix depth	14		6.7%
	4 Lower symblepharon	1	1/2 x 100	50%
	5 Upper symblepharon	0	0	0%
<i>Scarring score</i> (total of percentage values/4)				34.2%
<b>Morbidity</b>	6 Corneal vascularity	3 peripheral 0 central	3/5 x 100	60%
	7 Corneal opacity	2 peripheral 0 central	2/9 x 100	22%
<i>Morbidity score</i> (total of percentage values/2)				41%
Distichiasis: No				
Keratinisation: No				
<b>Combined scarring &amp; morbidity score</b> (total of percentage values/6)				36.5%
Patient study #2 Aged 68, Right Eye				
Categories	Components	Value	Transposition	Percentage
<b>Inflammation</b>	1 Inflammation	9	9/16 x 100	56%
<i>Inflammation score</i> (total of percentage value)				56%
Limbitis: No				
<b>Scarring</b>	2 Lower fornix depth	4		63.2%
	3 Upper fornix depth	3		80.5%
	4 Lower symblepharon	1	1/2 x 100	50%
	5 Upper symblepharon	1	1/2 x 100	50%
<i>Scarring score</i> (total of percentage values/4)				60.9%
<b>Morbidity</b>	6 Corneal vascularity	4 peripheral 0 central	4/5 x 100	80%
	7 Corneal opacity	4 peripheral 1 central	9/9 x 100	100%
<i>Morbidity score</i> (total of percentage values/2)				90%
Distichiasis: No				
Keratinisation: No				
<b>Combined scarring &amp; morbidity score</b> (total of percentage values/6)				68.8%

Patient (hypothetical case with burned out disease no inflammation, with ankyloblepharon, central vascularisation and opacity Aged 68, Right Eye)

Categories	Components	Value	Transposition	Percentage
<b>Inflammation</b>	1 Inflammation	0/16		0%
<i>Inflammation score</i> (total of percentage values)				0%
Limbitis: No				
<b>Scarring</b>	2 Lower fornix depth	0		100%
	3 Upper fornix depth	0		100%
	4 Lower symblepharon	2	2/2 x 100	100%
	5 Upper symblepharon	2	2/2 x 100	100%
<i>Scarring score</i> (total of percentage values/4)				100%
<b>Morbidity</b>	6 Corneal vascularity	4 peripheral 1 central	5/5 x 100	100%

**Table 5 (continued)**

Categories	Components	Value	Transposition	Percentage
Patient (hypothetical case with burned out disease no inflammation, with ankyloblepharon, central vascularisation and opacity Aged 68, Right Eye)				
	7 Corneal opacity	4 peripheral 1 central	9/9 x 100	100%
<i>Morbidity score</i> (total of percentage values/2)				100%
Distichiasis: No				
Keratinisation				
- Bulbar conjunctiva: Yes				
- Lid margin/tarsal conjunctiva: Yes				
- Cornea: No				
<b>Combined scarring &amp; morbidity score</b> (total of percentage values/6)				100%



**Fig. 2.** Comparison of the studied clinical severity-grading tool with the Sotozono grading system for Stevens-Johnson Syndrome. a) Good correlation was found between the composite scores from our cicatrising conjunctivitis assessment tool and the total scores from the Sotozono grading system [Pearson  $r = 0.93$   $p < 0.001$ ]; b) no significant proportional bias between the two grading systems was also observed when Bland-Altman plots were applied.

the tool can be used independently. For example, conjunctival hyperemia is a marker for conjunctival inflammation, which is an important disease activity marker and a principal short-term target for therapy; inflammation is easily assessed at every visit. Published inflammation grading schemes have used entirely subjective evaluation of conjunctival hyperaemia [20,28], whereas ours described in this study is guided by a comparison to standard panel of photographs and has good inter-observer agreement. The only moderate-to-good intra-observer agreement for this component may be partly explained by differences in the degree of hyperemia between the first and second assessments.

Supplementary Table 1 summarizes the comparison of our tool to three other published tools for the assessment of different types of CC (other than trachoma) and shows that our tool has fewer categories and

components. The rationale for excluding some of the components used in other schemes is as follows. We chose the components for measurement based on our experience of parameters that were reliably measurable in our cross-sectional study [26]. Of the components used by Francis [17] we omitted: entropion because patients have often had multiple surgeries altering the lid position, trichiasis as lashes may have been recently epilated, lagophthalmos which is difficult to identify unless severe, Schirmer's test because patients with short fornices cannot have this measured, ocular rotations as we found poor inter- and intra-observer correlation in our analysis, punctate keratopathy as it is present in most cases and time consuming to assess, corneal epithelial defect/ulceration and infection as these are of short duration and their effects are measured by the degree of residual corneal opacification which we assessed, and lastly visual acuity. Visual acuity is an unreliable measure of the severity of CC because many of the patients have other causes of visual loss that may be hard to identify and quantify in the setting of loss of corneal translucency. By comparison with the Sotozono grading system [20], in addition to leaving out the above mentioned components, we left out mucocutaneous junction, meibomian gland disease, punctal and palisade of Vogt grading, as being potentially difficult to measure reliably, and unlikely to add to the ability to discriminate between cases using the reduced numbers of components that we finally selected. We also left out conjunctivalisation because of the difficulty our observers found in distinguishing this from corneal neovascularisation in our previous cross-sectional study [26]. The comparison of our CC assessment tool to the Sotozono grading system in SJS showed good correlation and agreement [Fig. 2], despite a 50% reduction in the number of components graded (from 14 to 7) in our tool.

Scarring is a functional category, which is a hallmark sign of ocular MMP, resulting in much of the disease morbidity, and as such has to date, been the only category measured in many MMP grading systems. Most of these systems rely upon subjective grading of scarring usually restricted to lower fornix shortening and symblepharon. The most widely used for clinical studies and trials were developed three decades ago. Mondino and Brown described a four stage grading system in which the percentage shrinkage of the lower fornix was qualitatively evaluated [8]. Foster subsequently described a different four stage qualitative system which included sub-epithelial fibrosis without fornix shortening (the earliest sign of CC) as Stage I, and differentiated 3 subsequent stages by the presence of fornix foreshortening (Stage II), presence or absence of symblepharon (Stage III), and ankyloblepharon (Stage IV) [9]. These methods can often be used to retrospectively grade disease from clinical records. However, Tauber and Foster recognised that for the assessment of progression of scarring, both of these methods were relatively insensitive in detecting disease progression. To remedy this, they described a qualitative system combining four grades of fornix foreshortening with four grades of horizontal involvement by symblepharon, resulting in a 10 stage system, further subdivided by the number of discrete symblephara when these are present [10]. However, this number of symblephara is not discrete but variable depending on the degree of traction on the lower lid. Although these grading systems have been the most widely used to date [15,16], their assessment of scarring are qualitative, depending on subjective assessment, and limited to the assessment of the lower conjunctival fornix. Studies have indicated that subjective measures to evaluate scarring are not as good as objective measures [12]. Recognising the limitations of these methods, various groups have since introduced devices to facilitate the objective evaluation of the conjunctival fornices [11,12,19,22,23]. Such devices include the FDM, which allow quantitative measurements of the lower, and in some devices, the upper forniceal depths. We have used an FDM which reproducibly measures inter- and intra-observer variability within  $\pm 1$  mm for the lower fornix, giving intra-observer agreement to within  $\pm 1$  mm for 86% of measurements compared to only 63% for the subjective measurements [12]. The 95% confidence limits for the upper fornix were  $\pm 2.5$  mm. Furthermore, this device has

also been used to define the upper and lower fornix depths in normal Caucasian [23] and South Asian [22] populations which are invaluable for detecting fornix depths that indicate early CC and which we have incorporated into our tool. Our published phenotype comparison cross-sectional study incorporates the use of both the FDM and the Tauber staging for the evaluation of scarring [26], the latter of which has been considered the 'gold standard' for the grading of conjunctival scarring [10]. In our study, we found good correlation between the lower fornix depth measurements and Tauber staging (Spearman's  $r = -0.874$ ,  $p < 0.001$ ) (Supplementary Fig. 2). However, this does not infer that the inter- and intra-observer agreements are as good for the Tauber system as described above. The FDM device is shortly to become commercially available as an inexpensive disposable device (<http://www.scopeophthalmics.com/fdm>, Scope Ophthalmics Ltd, United Kingdom); FDM availability has been a principal deterrent to its widespread adoption because of the expense of in-house production and inadvertent disposal of the device in clinics.

Our final CC assessment tool, together with ready reckoners and calculators, is provided in Supplementary Appendices 2 and 3. We think that the three categories of our tool will be used independently in different applications. For clinical use, the inflammation category assessment takes only a few minutes to grade. Being reproducible, we hope that this category might be used in all specialist clinics treating patients with anterior segment inflammatory diseases who wish to assess the responses to anti-inflammatory treatments. The scarring category assessment with the FDM and ready reckoner for the lower fornix also takes only a few minutes and will identify patients with scarring diseases of the conjunctiva. For example, they can be used in glaucoma clinics where conjunctival scarring is a common side effect of medication, in general ophthalmology clinics for the assessment of those with chronic conjunctivitis to identify those with significant scarring earlier in the course of disease, and in oculoplastics surgery clinics where patients with entropion due to CC are often initially misdiagnosed. The morbidity category assessment aims to identify patients who suffer from the sight-threatening complications of inflammation and scarring. Components in this category include an evaluation of the presence or absence of ocular surface keratinisation, distichiasis, and corneal pathologies (vascularisation and opacification). These measures convey an indication of the overall severity of ocular surface disease and visual prognosis. Such information can also be used for epidemiological studies, assessment of treatment outcomes, as well as a guide to when visual rehabilitative therapy is needed.

Ocular surface keratinisation deserves a special mention. The presence of keratin on the ocular surface (conjunctiva, lid margin, and cornea), has been shown in previous studies to be an important marker of the severity of dry eye and ocular surface damage [17,20,27,28]. In SJS the presence of lid margin keratinisation is a direct causative factor of poor visual prognosis [27,29]. The damage from lid margin keratinisation is thought to be through a vicious cycle of blink-related microtrauma to the ocular surface resulting in chronic inflammation and eventually scarring. Indeed, the presence of lid margin involvement may be an indication for specific treatments, both in the acute phase (e.g. amniotic membrane graft) [27] and chronic phase (e.g. mucous membrane graft) [29]. Thus, despite poor levels of inter- and intra-observer agreement for the ocular surface keratinisation grading found in our study, its identification has been retained in the final clinical assessment tool.

The key limitations of this study are that the examiners might have potentially recalled their previous measurements when carrying out the second examination in the part of the study that evaluated intra-observer repeatability. Also, although designed for use in any disease causing CC, the tool has only been evaluated in patients with MMP and SJS. Further studies are required to evaluate its usability in other causes of CC. The validation of this clinical assessment tool was performed using a cross-sectional design. A longitudinal design of inter-observer and intra-observer assessments on the same cohort of patients may give

information on the robustness of this clinical assessment tool in repeated measurements at different time points to monitor disease progression. However, as many CC conditions progress over long periods of time (months to years), it was beyond the scope of this study to perform such longitudinal validation studies.

Nevertheless, there are a number of strengths: (i) the adoption of the FDM to quantify fornix scarring which shows better inter- and intra-observer agreement than the widely used qualitative Tauber system [10,12], but with which it correlates well for the assessment of fornix foreshortening, (ii) the choice of components for the final tool by evaluation of those used in other studies [26] (iii) refinement of these components by rejecting those having poor inter- and intra-observer agreement (iv) the exclusion of redundancy between the remaining components (v) the comparison of this refined tool with the Sotozono grading system [20] showing similar case discrimination using fewer components. The principal utility of this CC assessment tool is the minimal number of components that have to be measured to provide an adequate evaluation of a patient with CC in a clinical setting, resulting in a tool that takes an experienced user only 10 min to perform or less, for measurement of individual categories of disease effects.

### Financial support

This study was funded by the Moorfields Eye Hospital and the UCL Institute of Ophthalmology National Institute of Research (NIHR) Biomedical Research Centre (BRC), United Kingdom (Ref: BMRC 045), anonymous donors to Moorfields Eye Charities, United Kingdom for research into cicatrising conjunctivitis, a Singapore National Eye Centre Health Research Endowment Fund, Singapore (Ref: R1430/13/2017), a SingHealth Foundation Grant, Singapore (Ref: SHF/FG707S/2017), and a SingHealth Duke-NUS Ophthalmology & Visual Sciences Academic Clinical Programme Award (Ref: 05/FY2019/P1/18-A33). None of the funders had any role in the design or conduct of this research. Part of JD's salary was paid by the NIHR BRC.

### Declaration of competing interest

None Declared.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jtos.2019.10.010>.

### References

- [1] Dart JK. The 2016 Bowman Lecture Conjunctival cures: scarring conjunctivitis 30 years on. *Eye (Lond)* 2017;31(2):301–32.
- [2] Radford CF, Rauz S, Williams GP, Saw VP, Dart JK. Incidence, presenting features, and diagnosis of cicatrising conjunctivitis in the United Kingdom. *Eye (Lond)* 2012;26(9):1199–208.
- [3] Taylor J, McMillan R, Shephard M, et al. World Workshop on Oral Medicine VI: a systematic review of the treatment of mucous membrane pemphigoid. *Oral surg. Oral Med. Oral pathol. Oral Radiol.* 2015;120(2):161–71. e20.
- [4] Sobolewska B, Deuter C, Zierhut M. Current medical treatment of ocular mucous membrane pemphigoid. *Ocul Surf* 2013;11(4):259–66.
- [5] Ahadome SD, Abraham DJ, Rayapureddi S, et al. Aldehyde dehydrogenase inhibition blocks mucosal fibrosis in human and mouse ocular scarring. *JCI Insight*; 1(12).
- [6] Zhong H, Sun G, Lin X, Wu K, Yu M. Evaluation of pirfenidone as a new post-operative anticarring agent in experimental glaucoma surgery. *Investig Ophthalmol Vis Sci* 2011;52(6):3136–42.
- [7] Mathewson PA, Williams GP, Watson SL, et al. Defining ocular surface disease activity and damage indices by an international delphi consultation. *Ocul Surf* 2017;15(1):97–111.
- [8] Mondino BJ, Brown SI. Ocular cicatricial pemphigoid. *Ophthalmology* 1981;88(2):95–100.
- [9] Foster CS. Cicatricial pemphigoid. *Trans Am Ophthalmol Soc* 1986;84:527–663.
- [10] Tauber J, Jabbur N, Foster CS. Improved detection of disease progression in ocular cicatricial pemphigoid. *Cornea* 1992;11(5):446–51.
- [11] Kawakita T, Kawashima M, Murat D, Tsubota K, Shimazaki J. Measurement of fornix depth and area: a novel method of determining the severity of fornix shortening. *Eye (Lond)* 2009;23(5):1115–9.
- [12] Williams GP, Saw VP, Saeed T, et al. Validation of a fornix depth measurer: a putative tool for the assessment of progressive cicatrising conjunctivitis. *Br J Ophthalmol* 2011;95(6):842–7.
- [13] Reeves GM, Lloyd M, Rajlawat BP, Barker GL, Field EA, Kaye SB. Ocular and oral grading of mucous membrane pemphigoid. 2012;250(4):611–8. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie.
- [14] Rowsey JJ, Macias-Rodriguez Y, Cukrowski C. A new method for measuring progression in patients with ocular cicatricial pemphigoid. *Arch Ophthalmol* 2004;122(2):179–84.
- [15] Lee BWH, Tan JCK, Radjenovic M, Coroneo MT, Murrell DF. A review of scoring systems for ocular involvement in chronic cutaneous bullous diseases. *Orphanet J Rare Dis* 2018;13(1):83.
- [16] Munyangango EM, Le Roux-Villet C, Doan S, et al. Oral cyclophosphamide without corticosteroids to treat mucous membrane pemphigoid. *Br J Dermatol* 2013;168(2):381–90.
- [17] Francis IC, McCluskey PJ, Walls RS, Wakefield D, Brewer JM. Ocular cicatricial pemphigoid. *Aust N Z J Ophthalmol* 1990;18(2):143–50.
- [18] Murrell DF, Marinovic B, Caux F, et al. Definitions and outcome measures for mucous membrane pemphigoid: recommendations of an international panel of experts. *J Am Acad Dermatol* 2015;72(1):168–74.
- [19] Schwab IR, Linberg JV, Gioia VM, Benson WH, Chao GM. Foreshortening of the inferior conjunctival fornix associated with chronic glaucoma medications. *Ophthalmology* 1992;99(2):197–202.
- [20] Sotozono C, Ang LP, Koizumi N, et al. New grading system for the evaluation of chronic ocular manifestations in patients with Stevens-Johnson syndrome. *Ophthalmology* 2007;114(7):1294–302.
- [21] Sharma N, Venugopal R, Maharana PK, et al. Multi-step grading system for evaluation of chronic ocular sequelae in patients with Stevens-Johnson syndrome. *Am J Ophthalmol* 2019.
- [22] Khan IJ, Ghauri AJ, Hodson J, et al. Defining the limits of normal conjunctival fornix anatomy in a healthy South Asian population. *Ophthalmology* 2014;121(2):492–7.
- [23] Jutley G, Carpenter D, Hau S, et al. Upper and lower conjunctival fornix depth in healthy white caucasian eyes: a method of objective assessment. *Eye (Lond)* 2016;30(10):1351–8.
- [24] Williams GP, Nightingale P, Southworth S, et al. Conjunctival neutrophils predict progressive scarring in ocular mucous membrane pemphigoid. *Investig Ophthalmol Vis Sci* 2016;57(13):5457–69.
- [25] Thylefors B, Dawson CR, Jones BR, West SK, Taylor HR. A simple system for the assessment of trachoma and its complications. *Bull World Health Organ* 1987;65(4):477–83.
- [26] Ong HS, Setterfield JF, Minassian DC, Dart JK, Mucous Membrane Pemphigoid Study G. Mucous membrane pemphigoid with ocular involvement: the clinical phenotype and its relationship to direct immunofluorescence findings. *Ophthalmology* 2018;125(4):496–504.
- [27] Kohanim S, Palioura S, Saeed HN, et al. Acute and chronic ophthalmic involvement in Stevens-Johnson syndrome/toxic epidermal necrolysis - a comprehensive review and guide to therapy. II. Ophthalmic disease. *Ocul Surf* 2016;14(2):168–88.
- [28] Francis IC, McCluskey PJ, Wakefield D, Branley MG. Medial canthal keratinisation (MCK): a diagnostic sign of ocular cicatricial pemphigoid. *Aust N Z J Ophthalmol* 1992;20(4):350–1.
- [29] Basu S, Sureka SP, Shanbhag SS, Kethiri AR, Singh V, Sangwan VS. Simple limbal epithelial transplantation: long-term clinical outcomes in 125 cases of unilateral chronic ocular surface burns. *Ophthalmology* 2016;123(5):1000–10.