Three-Year Outcomes of Cultured Limbal Epithelial Allografts in Aniridia and Stevens-Johnson Syndrome Evaluated Using the Clinical Outcome Assessment in Surgical Trials Assessment Tool

ALEX J. SHORTT,[a,b,c] CATEY BUNCE,[c] HANNAH J. LEVIS,[b] PETER BLOWS,[c] CAROLINE J. DORÉ,[d] AMANDA VERNON,[a,b] GENEVIEVE A. SECKER,[a] STEPHEN J. TUFT,[a,c] JULIE T. DANIELS[a,b]

Key Words. Stem cell • Cornea • Stem cell deficiency • Stem cell transplantation • Outcome measures

[a]Cells for Sight Transplantation and Research Programme and [b]Ocular Biology and Therapeutics Division, University College London Institute of Ophthalmology, London, United Kingdom; [c]Moorfields Eye Hospital National Health Service Foundation Trust, London, United Kingdom; [d]University College London Clinical Trials Unit, London, United Kingdom

Correspondence: Alex J. Shortt, M.R.C.Ophth., Ph.D., National Institute for Health Research Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital National Health Service Foundation Trust and University College London Institute of Ophthalmology, 162 City Road, London EC1V 2PD, United Kingdom. Telephone: 44(0)207-253-3411, ext. 4654; E-Mail: a.shortt@ucl.ac.uk

©AlphaMed Press

ABSTRACT

Limbal stem cell deficiency (LSCD) is an eye disorder in which the stem cells responsible for forming the surface skin of the cornea are destroyed by disease. This results in pain, loss of vision, and a cosmetically unpleasant appearance. Many new treatments, including stem cell therapies, are emerging for the treatment of this condition, but assessment of these new technologies is severely hampered by the lack of biomarkers for this disease or validated tools for assessing its severity. The aims of this study were to design and test the reliability of a tool for grading LSCD, to define a set of core outcome measures for use in evaluating treatments for this condition, and to demonstrate their utility. This was achieved by using our defined outcome set (which included the Clinical Outcome Assessment in Surgical Trials of Limbal stem cell deficiency [COASTL] tool) to evaluate the 3-year outcomes for allogeneic ex vivo cultivated limbal epithelial transplantation (allo-CLET) in patients who had bilateral total LSCD secondary to aniridia or Stevens-Johnson syndrome. The results demonstrate that our new grading tool for LSCD, the COASTL tool, is reliable and repeatable, and that improvements in the biomarkers used in this tool correlate positively with improvements in visual acuity. The COASTL tool showed that following allo-CLET there was a decrease in LSCD severity and an increase in visual acuity up to 12 months post-treatment, but thereafter LSCD severity and visual acuity progressively deteriorated. STEM CELLS TRANSLATIONAL MEDICINE *2014;3:265–275*

INTRODUCTION

The cornea is the clear window at the front of the eye. Its transparency depends on a complex relationship between collagen fibrils, extracellular matrix, and several specialized cell types. One of the cell populations, the limbal epithelial stem cells (LESCs), are responsible for generating and maintaining the transparent multilayered epithelium that covers the outer surface of the cornea forming a protective barrier against the external environment [1–7]. LESCs are located in a specialized niche called the limbus [8]. The limbus is the semitranslucent ring of tissue that forms the junction between the transparent cornea and opaque sclera.

A wide range of diseases can result in damage to the LESCs and/or their niche. If severe, this damage can lead to a deficiency in LESC function resulting in an inability to maintain an intact transparent epithelium on the surface of the eye. Clinically, this condition is known as limbal stem cell deficiency (LSCD). Instead of a healthy transparent corneal epithelium generated by functioning LESCs, patients with LSCD suffer from migration of peripheral conjunctival epithelial cells and blood vessels onto the corneal surface, which is associated with loss of vision and intermittent ulceration [1–7]. The leading causes of LSCD in the developed world are chemical/thermal injury, Stevens-Johnson syndrome (SJS), aniridia, and ocular cicatricial pemphigoid [1–7]. Following chemical injury, LESCs are destroyed by direct physical-chemical damage. In SJS, the mechanism of LESC damage is poorly understood, but it may occur as a result of the acute inflammatory episode simultaneous with orocutaneous ulceration or develop several years later as a result of chronic ocular surface disease. The cause of LSCD in aniridia is related to an abnormality of the *PAX6* gene. This gene plays a key role in the development of the anterior segment of the eye and mutations in a single allele result in a wide range of abnormalities. Interestingly, patients with aniridia have normal corneal epithelium at birth but begin to develop signs of LSCD in their 20s to 30s; these signs slowly progress to LSCD. The mechanism by which *PAX6* mutations result in LSCD is not known.

The severity of LSCD can vary. It may be classified as partial or total and as unilateral or bilateral. The treatment of severe LSCD has been revolutionized by the introduction of surgical

techniques for LESC transplantation. There are several different surgical options, all of which have been summarized recently by the Corneal Society (http://www.corneasociety.org) [9]. Broadly, there are two approaches: (a) direct transplantation of whole segments of healthy limbal tissue—these can be either autologous (from the fellow eye) or allogeneic (from a cadaveric donor) [10–16]—and (b) transplantation of a bioengineered cell sheet that is constructed from LESCs in a specialist culture facility. These cell sheets can be constructed from autologous (the fellow eye) or from allogeneic (a cadaveric donor) LESCs [17–20]. Direct limbal tissue transplantation was first reported in 1989 [12] and bioengineered LESC cell sheets in 1997 [17]. Since then, there have been ever increasing numbers of publications reporting the clinical outcomes for both approaches [6, 21]. A significant problem when interpreting data from these studies is the way in which outcomes have been measured and reported. Outcomes have generally been based on subjective grading, which is open to bias, or on attempts to stratify outcomes based on poorly defined criteria, the reliability of which had not been tested [19, 22–26]. This lack of objective outcome measures makes it difficult to compare results between studies and to determine which of the many surgical options has the best outcome in a given patient cohort. The increasing use of stem cell technology to treat these patients has resulted in additional pressure from regulatory authorities and health care funders to demonstrate their safety, efficacy, and cost-effectiveness using defined, objective, and reliable outcome criteria. At present, no such outcomes exist.

The aims of this study were to (a) design and test the reliability of a grading tool for quantifying the severity of LSCD, (b) define a set of core outcome measures for use in treatments for this condition, and (c) demonstrate the utility of these core outcomes by reporting the 3-year results for allogeneic ex vivo cultivated limbal epithelial transplantation (allo-CLET) in patients who had bilateral total LSCD secondary to aniridia or SJS.

MATERIALS AND METHODS

Design and Validation of a Grading System

This study adhered to the tenets of the Declaration of Helsinki. Ethical approval was obtained from the institutional Medical Research Ethics Committee, and informed consent was obtained from all patients prior to treatment. A list of clinical signs useful in the diagnosis and assessment of LSCD severity was created from outcome measures used in previously published studies in this field. Nine clinicians and scientists with experience in diagnosing and treating limbal stem cell deficiency anonymously selected four key clinical signs of LSCD from this list for use in a new grading tool: corneal epithelial haze, superficial corneal neovascularization, corneal epithelial irregularity, and corneal epithelial defect. A standardized grading plate was produced for each of these parameters with severity graded from 0 (normal) to grade 3 (severe) (Fig. 1; supplemental online Figs. 1–3). The use of separate plates allowed independent grading of each parameter in corneas manifesting multiple abnormalities.

Clinical photographs were formatted using a standardized technique. The protocol required a minimum of two images of the whole cornea, with one high ($\times16$) magnification image of the central cornea and one image under cobalt blue illumination following instillation of one drop of fluorescein 2%. The four photographs were arranged into a single large image using Adobe Illustrator CS4 (Adobe Systems Inc., San Jose, CA, http://www.adobe.com)

(supplemental online Fig. 4). Two observers masked to the clinical history then graded each image against the standardized grading plates. The scores for each of the four clinical signs of LSCD were summated to give a score of overall severity (global score, range 0 to 12). This system was designated the Clinical Outcome Assessment in Surgical Trials of Limbal stem cell deficiency (COASTL) tool. It is a composite ordinal grading scale.

The reliability of the COASTL tool was tested using images from 26 patients with varying degrees of LSCD from a range of etiologies. Two observers graded the same set of images in random order on two occasions 28 days apart. Repeatability (intraobserver agreement) was assessed using weighted κ to assess between the respective observers' scores at different time points [27, 28]. Reproducibility (interobserver agreement) was assessed using the individual observers' initial scores, which were again assessed using weighted κ because the scale was ordinal. In all cases of weighted κ, we used the following quadratic weighting: $1 - \{(i - j)/(k - 1)\}^2$, where i and j index the rows and columns of the ratings by the two raters, and k is the maximum number of possible ratings. Because there was one patient who contributed both eyes to the analysis, we treated eyes as independent (but checked that main results were robust by sensitivity analyses without this one observation).

Grading of Allo-CLET Outcomes

Fourteen eyes of 13 patients with total bilateral LSCD with aniridia (10 eyes of 9 patients) or SJS (4 eyes of 4 patients) were treated by allo-CLET. The culture and transplantation techniques have been described previously [29]. Patients received postoperative oral prednisolone for 4 weeks and oral cyclosporin or mycophenolate for 6 months. Corneal photographs were taken before surgery and at intervals up to 3 years following allo-CLET. These were graded using the COASTL tool. The effect of treatment on the individual parameters of LSCD and on the global grading score was evaluated by plotting the mean change in score from baseline over time. We arbitrarily defined success as a 25% reduction in the global score and using this definition of success created a Kaplan-Meier survival curve for each patient group.

Assessment of Biomarker Impact and Relevance

To confirm the impact and relevance of the four biomarkers (corneal epithelial haze, superficial corneal neovascularisation, corneal epithelial irregularity, and corneal epithelial defect) each was correlated with visual acuity. Data were collected from the aforementioned 14 eyes of the 13 patients that underwent allo-CLET. Changes in each of the biomarkers and the global (total) score were correlated with changes in visual acuity using Spearman rank correlation.

In Vivo Confocal Microscopy and Impression Cytology

In vivo confocal microscopy (IVCM) with the Rostock Corneal Module and Heidelberg Retina Tomograph II (Heidelberg Engineering Gmbh, Heidelberg, Germany, http://www.heidelbergengineering.com) was performed on five normal volunteers and 31 patients with LSCD. Eleven of these 31 patients subsequently underwent allo-CLET and IVCM was then repeated at intervals following surgery determined by patient tolerance and clinical status. Optical sections were collected from at least two locations in the central cornea. Impression cytology was also performed before and after surgery in the 11 allo-CLET patients. Samples were collected using Biopore membranes (Millicell-CM 0.4 μm PICM 012550; Millipore Corp.,
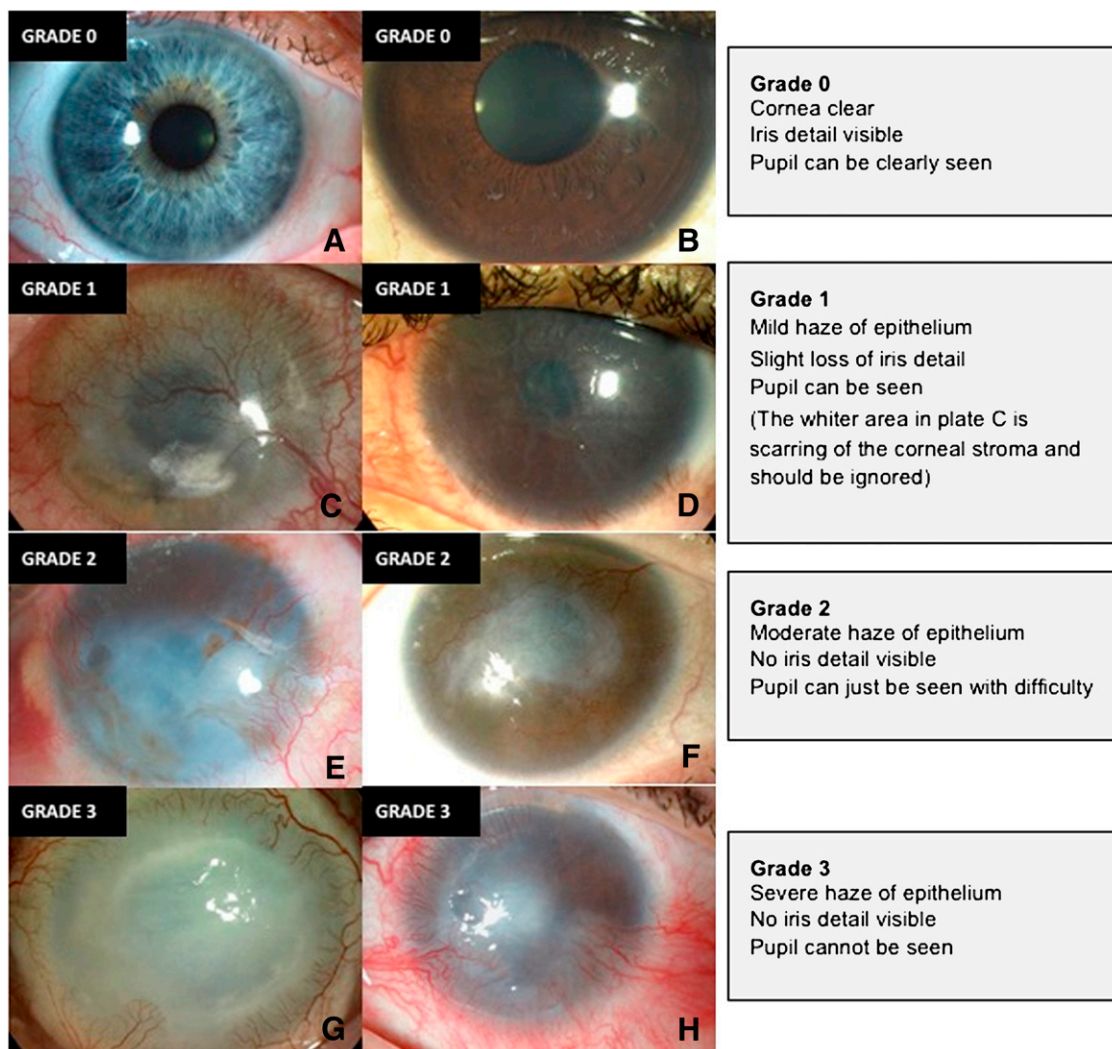
Figure 1. Standardized grading plate used to grade corneal epithelial haze as normal (grade 0, no signs) **(A, B)**, mild (grade 1) **(C, D)**, moderate (grade 2) **(E, F)**, or severe (grade 3) **(G, H)**.

Bedford, MA, http://www.millipore.com) and were immunostained with monoclonal antibodies to CK3 (Clone AE5, Chemicon; Millipore Corp.) and CK19 (AB15463; Abcam, Cambridge, U.K., http://www.abcam.com) to determine the phenotype of cells populating the corneal surface [30, 31], and interpreted by a consultant pathologist.

Images from healthy individuals were used to construct two plates demonstrating the range of appearances of different levels in the conjunctival and corneal epithelium. IVCM images from the 31 LSCD patients and the 11 who had allo-CLET were compared against the normal plates by two masked observers who were asked to classify each image to one of the following five groups: normal corneal morphology, normal conjunctival morphology, mixed populations of corneal and conjunctival morphology, normal epithelial cells but phenotype undistinguishable, and few/no cells with an epithelial phenotype visible. Between 5 and 20 representative images were evaluated from each patient (365 images in total).

Visual Outcomes

Visual acuities were recorded as the log of the minimum angle of resolution (logMAR) with acuities of count fingers at 1 m = 2.0,

hand movements at 1 m = 2.3, perception of light = 2.75, and no perception of light = 3.0 [32]. Postoperative visual outcome was assessed at 6, 12, 24, and 36 months. Visual outcome data were presented in three ways: scatter plots of pre- versus postoperative acuity at various time points, mean gain in acuity (lines of logMAR acuity) plotted against time, and the percentage of patients who gained or lost one or more lines of logMAR acuity at each time point.

RESULTS

Repeatability and Reproducibility of the COASTL Tool for Grading LSCD

The sum of the scores assigned to each of the four key signs of LSCD in the 26 patients with LSCD had a wide range that reflected the different severities in the population sample (as assessed using data gathered from the first set of observations from observer 1) (Fig. 2A). Figure 2B shows the distribution of scores for individual key clinical signs. There is an even distribution except for grade 1 epithelial defects (supplemental online Fig. 3), little evidence of
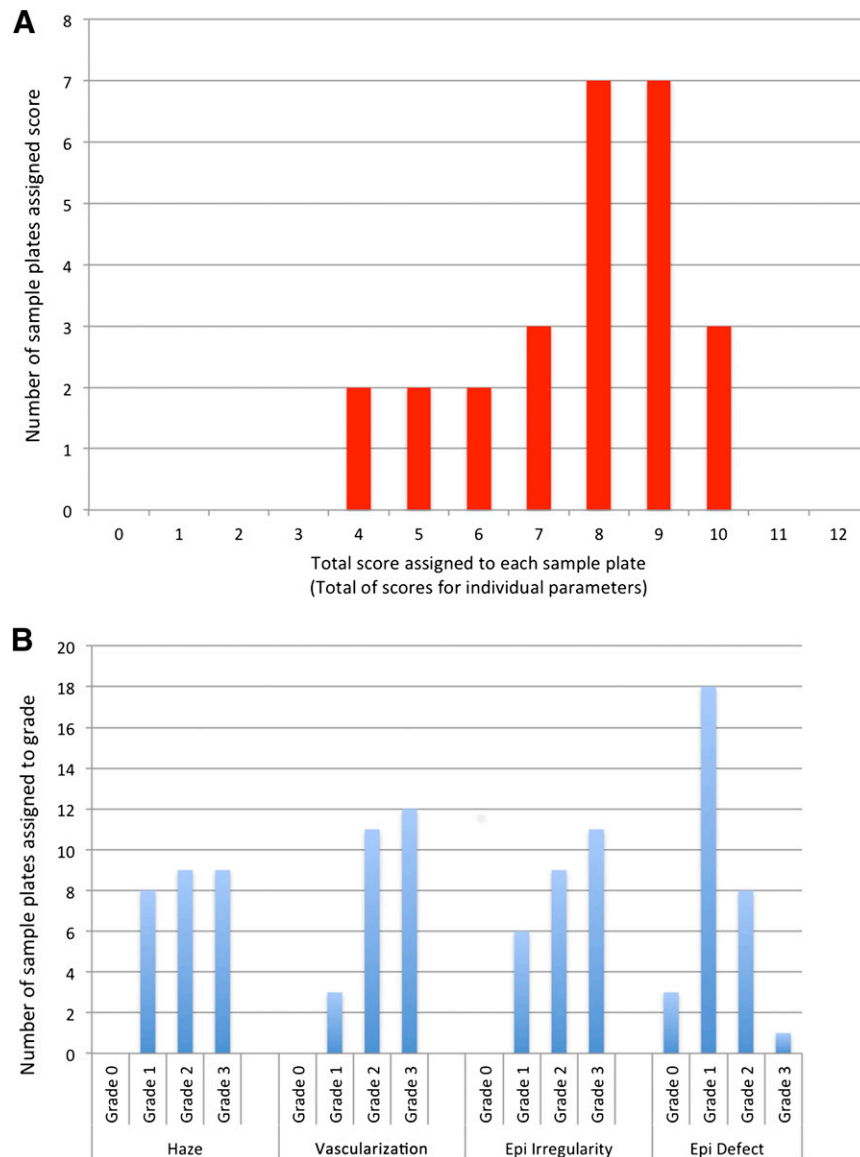
Figure 2. Graphs used to assess distribution of grading scores. **(A):** Distribution of global scores (the sum of the scores assigned to each of the four key signs of limbal stem cell deficiency [LSCD]) for 26 patients with limbal stem cell deficiency. Scores were distributed over a wide range on the scale, reflecting the range of LSCD severities in the population sample. **(B):** Distribution of scores within individual clinical parameters. Scores again show an even distribution with the exception of grade 1 epithelial defects. Abbreviation: Epi, epithelial.

floor or ceiling effects, and investigators did not report instances in which they felt their grading had been censored.

Repeatability was assessed by calculating weighted κ scores for intraobserver agreement (supplemental online Table 1). κ scores of between 0.61 and 0.8 indicate good agreement, and values of greater than 0.8 indicate very good agreement [26]. Observer 1 was an ophthalmic grading technician who had extensive experience of grading retinal disorders but no prior experience of corneal disease. Observer 1 had high levels of agreement and weighted κ scores confirming very good agreement for three of four of the individual parameters (epithelial haze, vascularization, and epithelial defects). The agreement for epithelial irregularity and for the global score was good. Observer 2 was a clinician experienced in the assessment and treatment of LSCD. Observer 2 had high levels of agreement between test and retest scores, both for the individual parameters of LSCD

and for the global score. The weighted κ scores for observer 2 showed very good agreement for all parameters and for the global score.

Reproducibility was assessed by calculating weighted κ scores for interobserver agreement using the first set of grading data generated by each observer (supplemental online Table 1). The agreement and weighted κ values scores indicated that there was very good agreement between observers for all of the individual parameters of LSCD and for the global score.

Application of Grading System to a Cohort of Allo-CLET Allografts

Ten patients with aniridia and four with SJS successfully underwent allo-CLET. Systemic immunosuppression with oral ciclosporin or mycophenolate was administered for a median of
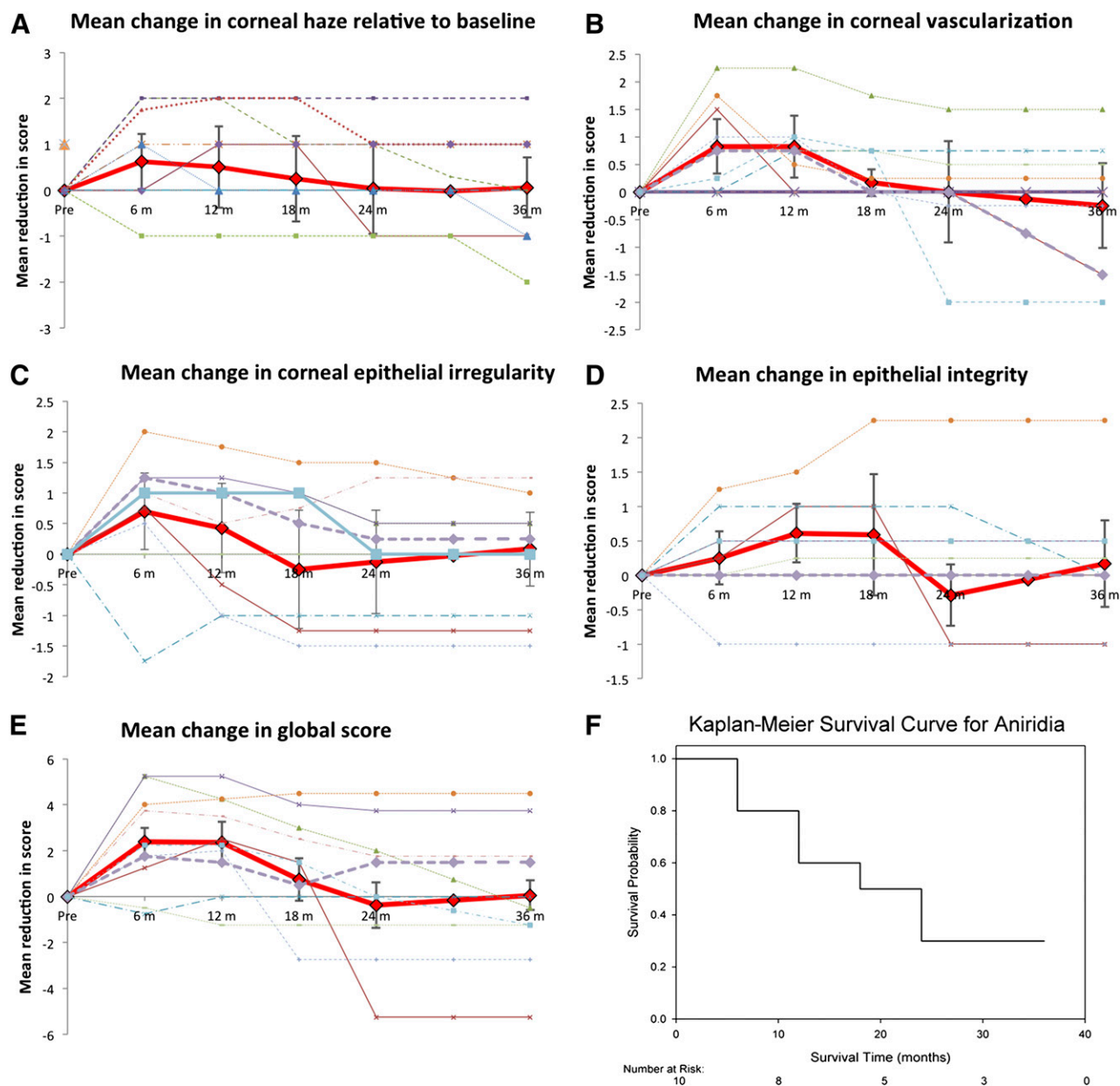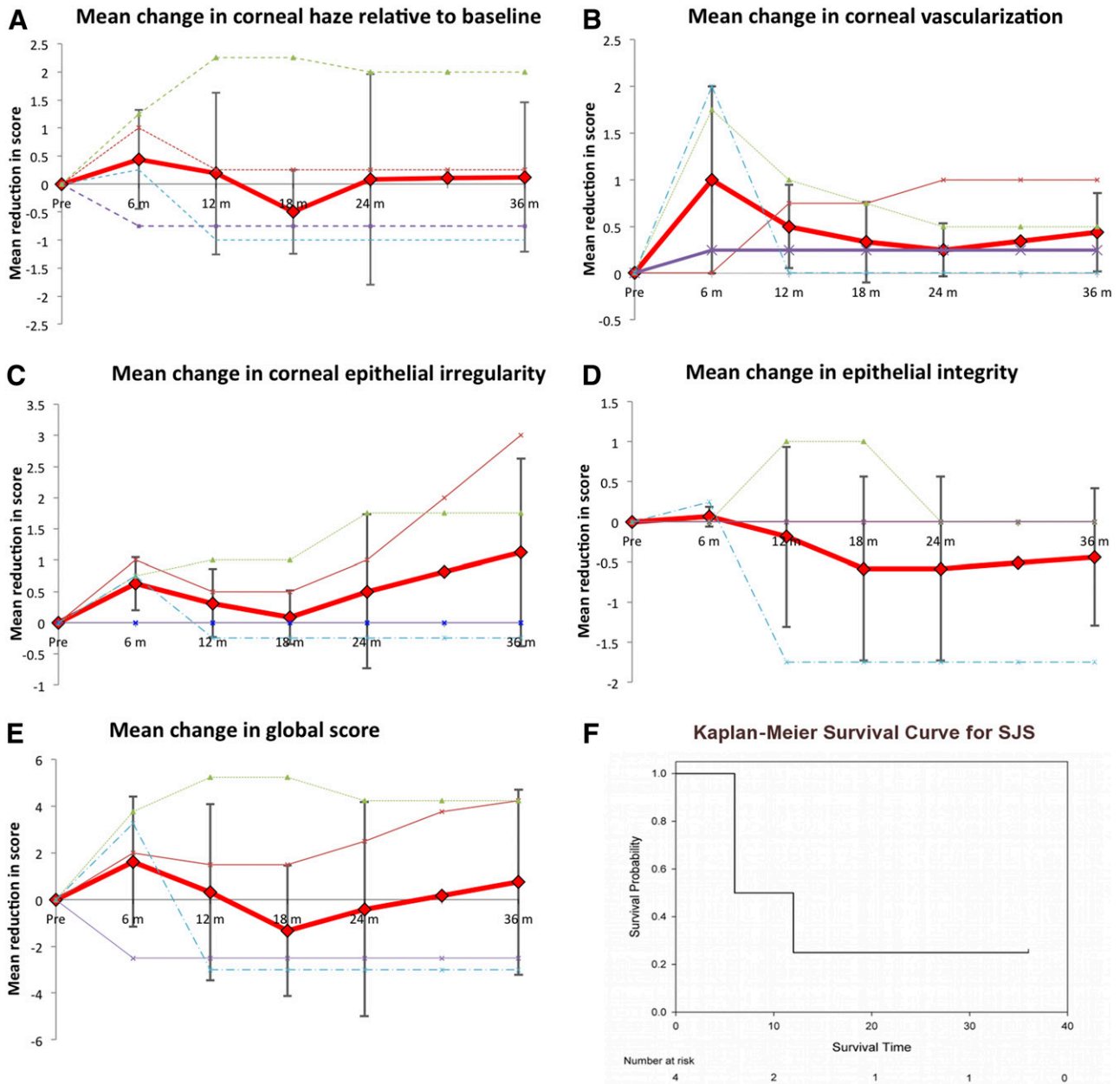
Figure 3. Outcomes of ex vivo cultivated limbal allografts for aniridia. **(A–D):** Effect of treatment on scores for corneal haze **(A)**, vascularization **(B)**, corneal epithelial irregularity **(C)**, and integrity (epithelial defects) **(D)**. In these graphs the solid red line represents the mean for all patients, the error bars indicate the 95% confidence intervals for the mean, and the dashed lines represent the results for individual patients. **(E, F):** Kaplan-Meier survival curve **(F)**, where success or survival was defined as a 25% improvement in the global score **(E)** versus the preoperative global score.

6 months then stopped. The individual scores for corneal haze, vascularization, corneal epithelial irregularity and integrity, as well as the global score (total of these), are presented in Figures 3 (aniridia) and 4 (SJS). In aniridia, there was a substantial improvement in mean scores for all four clinical parameters for up to 12 months following surgery. Signs of LSCD were recurring by 18 months, and by 24 months, patients on average had grading scores equivalent to their preoperative scores. The Kaplan-Meier survival curve shows that the probability of a sustained benefit beyond 2 years was only 25% (Fig. 3F). Although there were only a low number of patients with SJS, the results showed an initial improvement in all parameters at 6 months, but between 6 and 18 months,

signs of stem cell deficiency recurred. Interestingly, the global score in SJS improved steadily from 18 months onward due to a reduction in epithelial irregularity and decreased epithelial haze.

Assessment of Biomarker Impact and Relevance

There was a significant positive correlation between improvements in corneal epithelial haze (Spearman ρ 0.496, $p < .0001$), superficial corneal neovascularisation (Spearman ρ 0.397, $p = .002$), corneal epithelial irregularity (Spearman ρ 0.530, $p < .0001$), the global (total) score (Spearman ρ 0.562, $p < .0001$) and improvement in visual acuity (supplemental online Fig. 6). There was no significant

Figure 4. Outcomes of ex vivo cultivated limbal allografts for Stevens-Johnson syndrome. **(A–D):** Effect of treatment on scores for corneal haze **(A)**, vascularization **(B)**, corneal epithelial irregularity **(C)**, and integrity (epithelial defects) **(D)**. In these graphs, the solid red line represents the mean for all patients, the error bars indicate the 95% confidence intervals for the mean, and the dashed lines represent the results for individual patients. **(E, F):** Kaplan-Meier survival curve **(F)**, where success or survival was defined as a 25% improvement in the global score versus the preoperative global score **(E)**. Abbreviation: SJS, Stevens-Johnson syndrome.

correlation between the presence of corneal epithelial defects and a change in visual acuity (Spearman ρ 0.074, $p < .586$). Epithelial defects cause significant pain and discomfort; therefore, this biomarker was included in our outcome set regardless.

In Vivo Confocal Microscopy and Impression Cytology Findings as Biomarkers in LSCD and Following Treatment

Initial classification of IVCM images from LSCD patients was based on a comparison with images of conjunctiva and cornea from normal volunteers. Independent observers agreed on only 20% of the

365 images. The main reason was that in LSCD the predominant phenotype was a thin sheet of small cells, usually a mono- or bilayer, with hyper-reflective nuclei but with no other cellular detail visible. This did not resemble anything seen in normal conjunctiva or cornea (Fig. 5A, 5B). In particular, intercellular junctions are prominent in normal cornea and conjunctiva but were completely absent in this LSCD phenotype. We labeled this as "nonstratifying epithelium." Impression cytology performed on these eyes showed that these cells had a spindle shape and expressed CK19 but not CK3 (Fig. 5C, 5D). In eyes in which confocal images did show clear intercellular boundaries, the observers could not distinguish
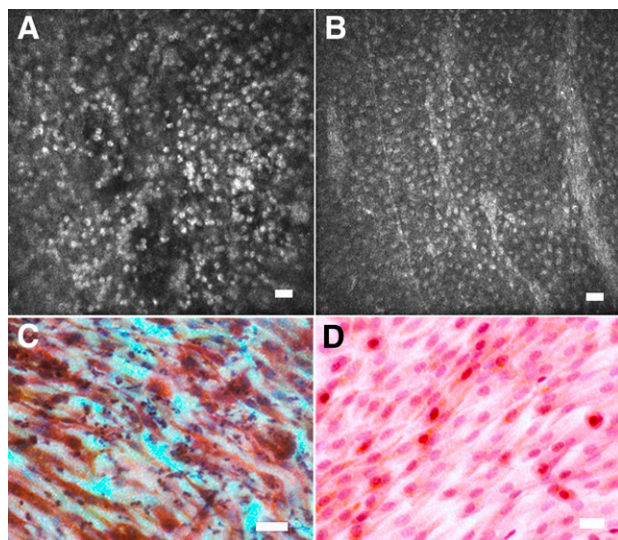
Figure 5. Nonstratifying epithelial phenotype observed in in vivo confocal microscopy and impression cytology observed in several patients with limbal stem cell deficiency. Shown are confocal microscopy **(A, B)** and corresponding impression cytology **(C, D)** images from the central cornea of two patients with limbal stem cell deficiency prior to treatment. **(A, B):** Confocal microscopy shows a thin sheet of small cells, usually a mono- or bilayer, with hyperfluorescent nuclei, but no other cellular detail is visible. **(C, D):** Correlation with impression cytology performed on the same patients directly after confocal scans showed that these cells had a spindle-like shape and expressed CK19 (red) but not CK3 (brown) **(C, D)**. This phenotype may represent conjunctival cell migration onto the corneal stroma but failure of the cell layer to form cell junctions and differentiate into a mature multilayered epithelium. Scale bars = 25 μm.

between corneal or conjunctival cells. We concluded that confocal microscopy could not be used to distinguish conjunctival and corneal epithelial enotypes. We therefore devised a revised classification based on LSCD alone (supplemental online Fig. 5). Images were classified as follows: (a) no epithelial cells visible, interpreted as an epithelial defect, ulcer, or an acellular scar on the ocular surface; (b) nonstratifying epithelium one or two cells thick, with hyper-reflective nuclei but absent intercellular junctions; or (c) stratified epithelium with clear intercellular boundaries indicative of normal epithelial differentiation. These could be conjunctival or corneal cells.

With this system, the observers agreed on 98% of the 365 images from 31 patients. Each of the 31 patients was assigned an overall score (based on the score most commonly recorded by the observer for that patient's images). We then assessed agreement between overall scores using κ and found almost perfect agreement with a κ of 0.91 (SE 0.11).

This modified system of classifying confocal phenotypes was then applied to patients undergoing allo-CLET. IVCM images of sufficient quality for analysis were available for 8 eyes before allo-CLET and 10 eyes following surgery. In addition, impression cytology of sufficient quality for analysis was available from 14 eyes before allo-CLET and 11 eyes following surgery (supplemental online Table 1). Before surgery, 6 of the 8 eyes that underwent IVCM showed a thin layer of hyper-reflective cells without identifiable cell boundaries. One patient had stratified differentiated epithelium and one had a mixed phenotype. Patients with a nonstratifying epithelial phenotype on IVCM had a correlating conjunctival morphology and CK19-positive (+ve) phenotype on impression cytology consistent

with this as shown in Figure 5. Three patients were examined 1 year after surgery, and they showed a multilayered differentiated epithelium that was mainly CK3+ve but also some areas of CK19+ve cells. Two of 4 eyes examined at 3 years following surgery showed the nonstratifying epithelial phenotype, while the other two displayed a differentiated epithelium with both CK3 and CK19 expression.
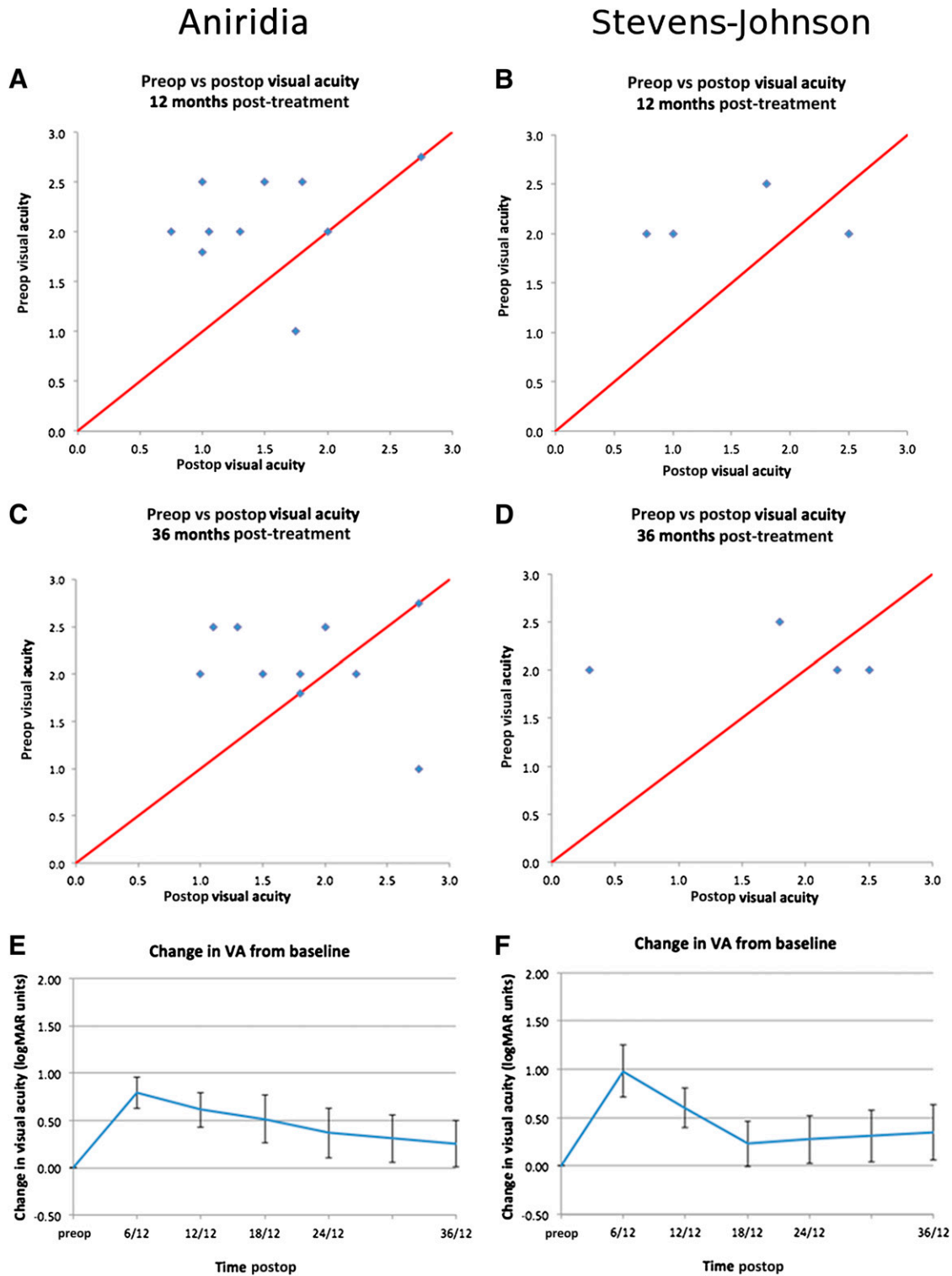
Visual Acuity

The preoperative visual acuity was 6/60 (Snellen acuity 20/200) or worse in all eyes undergoing allo-CLET. Scatter plots of pre- versus post-treatment acuity showed an improvement in visual acuity in 79% of eyes at 6 months, 71% at 12 months, 64% at 18 months, and 57% at both 24 and 36 months (Fig. 6A–6C). Visual acuity data were inspected for normality by plotting distribution at each time point. No significant skew was identified; therefore, mean visual acuity is reported. There was an initial improvement in the mean visual acuity of 0.86 logMAR units at 6 months (aniridia 0.78 logMAR, SJS 0.98 logMAR). The visual gain gradually reduced over the subsequent 30 months, but at 36 months, there was a stable mean gain of 0.28 logMAR units (0.26 in aniridic and 0.35 in SJS eyes) (Fig. 6D–6F). The percentage of eyes gaining one line of log-MAR acuity was 57% overall at 6 months (aniridia, 50%; SJS, 75%) and 29% at 36 months (aniridia, 30%; SJS, 25%).

DISCUSSION

The variability of biological systems means that clinical signs can rarely be simplified to exact numerical values, and in medicine, this is reflected in the use of therapeutic ranges and grading scales [33]. The laboratory and surgical techniques developed to treat LSCD continue to evolve, but an inability to objectively compare treatment outcomes has limited progress. There is currently no method to track the long-term fate of transplanted corneal epithelial cells, and surrogate measures are the only alternative at present. Regulatory bodies increasingly require proof of efficacy for new techniques, and a subjective assessment of success is insufficient. In this study, we have validated a set of objective measures to describe treatment outcomes for LSCD and demonstrated the use of these in reporting the 3-year outcomes of allo-CLET for aniridia and SJS.

In virtually all studies of limbal stem cell transplantation to date, the clinical outcome has been assessed subjectively by the investigating clinician. This is clearly open to significant measurement and reporting bias. Some studies have attempted to reduce bias by using an agreed set of outcome measures or assigning a score based on the overall interpretation of the observer [19, 22, 23, 25, 26]. However, the weakness of these studies is that the parameters assessed were poorly defined and the classification systems were not validated. An exception is the study by Sotozono et al., which used a reference plate of clinical signs to grade the ocular surface changes in SJS [24]. In the present study, we have reduced potential bias by using two independent masked observers to grade four specific indicators of stem cell deficiency. The use of different photographic magnifications and fluorescein staining to highlight areas of epithelial breakdown was specifically designed to enable grading of four defined and key signs of LSCD.

The utility of the COASTL tool developed for this study was evaluated by measuring its reliability. Grading systems that lack reliability are susceptible to random measurement error.

Aniridia

Stevens-Johnson

Figure 6. Visual acuity data for patients with aniridia and Stevens-Johnson syndrome treated with ex vivo cultivated limbal allografts. Values are means, and error bars indicate the 95% confidence intervals. Abbreviations: Postop, postoperative; Preop, preoperative; VA, logMAR visual acuity.

The intraobserver agreement (repeatability) was extremely good for both observers. Interobserver agreement (reproducibility) measures how similar or different the scores assigned by the two observers to the same cornea are. There was substantial agreement between observers. These demonstrate that the

COASTL tool is a reliable method of obtaining objective outcome data for surgical trials of limbal stem cell deficiency. To confirm that the biomarkers assessed by the COASTL tool (haze, vascularization, etc.) are clinically relevant, we investigated their correlation with visual acuity. A clinical improvement (i.e., a reduction) in

corneal haze, vascularization, and epithelial irregularity correlated significantly with an improvement in visual acuity, confirming the appropriateness of these biomarkers for measuring outcomes in trials of LSCD treatments. The exception to this was corneal epithelial integrity. Changes in this biomarker did not correlate significantly with changes in visual acuity. Nonetheless, it was included in the COASTL tool because the inability to maintain an intact corneal epithelium is a critical sign of LSCD, and furthermore, there is little doubt that epithelial defects result in significant pain and discomfort.

We used impression cytology with immunostaining for CK3/CK19 and IVCM as indices of biological success of these treatments. We hypothesized that IVCM could replace impression cytology in the assessment LSCD; however, this was not the case. Instead we conclude that the two techniques provide complementary data. Confocal microscopy alone could not accurately distinguish between corneal and conjunctival cell phenotypes, whereas impression cytology could, based on CK3 or CK19 expression. However, IVCM gave more information about the degree of epithelial stratification and the presence of cell-cell junction formation, indicating normal epithelial differentiation. A notable finding in some patients with LSCD was a central corneal epithelial thickness of only one to two cell layers formed of elongated spindle-shaped CK19+ve cells that appeared to be migrating conjunctival epithelial cells. To the best of our knowledge, this phenotype has not previously been described.

For studies of refractive eye surgery procedures, there are a set of standardized outcome measures that should be adhered to and reported [34]. No such standardized set of visual acuity outcome measures exists for LSCD treatments, probably because it is a challenge to interpret the wide range of visual acuities seen as well as the potential gain in vision following treatment. Also, pre-existing comorbidity, such as macular hypoplasia and nystagmus in aniridic patients and stromal corneal scarring in patients with SJS, limit potential visual improvement. We showed that pre- versus postoperative scatter plots are an effective method of evaluating this outcome. To the best of our knowledge, this is the first time that longitudinal long-term follow-up data on visual acuity has been presented. Another measure, the percentage of patients gaining five letters or more (one line of logMAR acuity) was also useful in determining the percentage of patients who had a functionally significant gain in acuity.

Having developed the COASTL tool and defined a set of outcome measures including standardized visual acuity plots, we applied this outcome tool to a series of allo-CLET procedures performed for aniridia or SJS. The outcomes clearly show an initial successful reduction in the signs of LSCD. However, in both aniridia and SJS, the signs of disease gradually returned and, by 24 months, reached pretreatment levels. The IVCM and confocal microscopy findings suggest that this is due to the recurrence of partial LSCD initially, as indicated by the CK3/19 mosaic and, in some cases, by the recurrence of total LSCD, as indicated by the nonstratifying epithelial phenotype. The initial improvement in clinical signs and cell phenotype was associated with an improvement in visual acuity, which slowly regressed over 36 months in both conditions, but it is notable that despite recurrence of clinical signs the visual benefit persists in 30% of aniridic and 25% of SJS patients at 36 months. This suggests that the treatment can have beneficial effects on visual function and, potentially, on quality of life, even in the presence of a recurrence of the signs of LSCD. This raises the issue of how a treatment "success" should be defined or what

the primary outcome measure for future trials of this therapy should be.

There are a wide variety of surgical treatments for transplanting LESCs [9]. Published data suggest that the procedure of choice for patients with long-standing unilateral alkali injuries is an ex vivo cultivated limbal autograft (auto-CLET) [19]. Evidence to guide decisions on the treatment of other disease groups is poor [35]. There is still controversy as to whether transfer of ex vivo expanded cells offers an advantage over direct tissue transfer techniques for patients with bilateral disease. Until recently, long-term outcome data for keratolimbal allografts (KLAL) showed universally poor results [36], but a more potent immunosuppressive regimen combining tacrolimus and mycophenolate with a short course of oral steroids can significantly prolong KLAL graft survival [37, 38]. Another novel way of improving outcomes of these allogeneic direct tissue transfer techniques is to combine living related conjunctival limbal autograft and KLAL (the Cincinnatti Procedure) thus giving superior outcomes to KLAL outcomes alone [38]. The closest comparable data to our data on patients with aniridia was reported by Biber et al. [38], who reported that at a mean follow-up of 43.4 months, 33% had an improved ocular surface and 75% of eyes had an improvement in vision of at least one line at the last follow-up. This present study found that 22% of aniridia eyes had a persistent improvement in clinical signs and 30% had a persistent gain of one line of logMAR acuity at 36 months. Apart from these data, it is difficult to compare the two sets of outcomes, highlighting the potential benefits of a standardized outcome reporting system.

The discrepancy between the outcomes reported by Biber et al. [38] and the less favorable outcomes in the present study may be accounted for by differences in immunosuppressive regimes. In the former study, two immunosuppressive agents (tacrolimus and mycophenolate) were administered indefinitely, whereas in the present study, one agent (cyclosporin or mycophenolate) was administered for 6 months only. The decision to use this regimen was based on evidence that donor epithelial cells could not be detected on the surface of the cornea beyond 8 months [39]. We hypothesized that for the 6 months following treatment, transplanted cells provide a suitable environment for innate cells to restore corneal epithelial homeostasis. We felt that beyond this period there was unlikely to be a significant benefit from immunosuppression that would justify the risk of potentially life-threatening complications. The alternative hypothesis for the mechanism of action of this treatment is that transplanted cells do indeed survive in the long term. There are data from animal models of the disease to suggest this may be the case [40]. The heterogeneous outcomes between studies using a variety of differing immunosuppressive regimes suggests that effective long-term immunosuppression could potentially improve outcomes. It is possible, however, that the difference could also be accounted for by inherent differences in the surgical procedures used.

From the patients' perspective, quality-of-life assessments that quantify the physical, emotional, and socioeconomic aspects of treatment for LSCD are an important outcome measure [41]. The quality of life for patients who have had conventional limbal transplants has been reported by Miri et al. [42], and that for patients who have received cultured LESCs has been reported by Di Girolamo et al. [43] and Kolli et al. [25]. A limitation of the present study is that these data were not collected. This was in part due to the fact that there is no quality-of-life assessment tool specifically designed for or validated for use in this condition. We are

currently developing such a tool in collaboration with patients and will use this tool in future studies.

CONCLUSION

The work of the Corneal Society in standardizing the nomenclature for ocular surface reconstructive techniques is to be welcomed [9]. The data presented in this manuscript highlight how a set of standardized outcome measures can be used to document outcomes. The Corneal Society is currently facilitating a consultation among the LSCD research community on adopting such a standardized system for reporting outcomes so that comparisons between studies and techniques can be made.

AUTHOR CONTRIBUTIONS

A.J.S., S.J.T., and J.T.D.: conception/design, provision of study material or patients, collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript; C.B.: collection and/or assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript; H.J.L. and P.B.: data analysis and interpretation, final approval of manuscript; C.J.D.: data analysis and interpretation, manuscript writing, final approval of manuscript; A.V. and G.A.S.: provision of study material or patients, collection and/or assembly of data, final approval of manuscript.

REFERENCES

1   Daniels JT, Dart JK, Tuft SJ et al. Corneal stem cells in review. Wound Repair Regen 2001;9:483–494.

2   Grueterich M, Espana EM, Tseng SC. Ex vivo expansion of limbal epithelial stem cells: amniotic membrane serving as a stem cell niche. Surv Ophthalmol 2003;48:631–646.

3   Holland EJ. Epithelial transplantation for the management of severe ocular surface disease. Trans Am Ophthalmol Soc 1996;94:677–743.

4   Kim JY, Djalilian AR, Schwartz GS et al. Ocular surface reconstruction: Limbal stem cell transplantation. Ophthalmol Clin North Am 2003;16:67–77.

5   Kinoshita S, Adachi W, Sotozono C et al. Characteristics of the human ocular surface epithelium. Prog Retin Eye Res 2001;20:639–673.

6   Shortt AJ, Secker GA, Notara MD et al. Transplantation of ex vivo cultured limbal epithelial stem cells: A review of techniques and clinical results. Surv Ophthalmol 2007;52:483–502.

7   Tseng SC. Concept and application of limbal stem cells. Eye (Lond) 1989;3:141–157.

8   Shortt AJ, Secker GA, Munro PM et al. Characterization of the limbal epithelial stem cell niche: Novel imaging techniques permit in vivo observation and targeted biopsy of limbal epithelial stem cells. STEM CELLS 2007;25:1402–1409.

9   Daya SM, Chan CC, Holland EJ et al. Cornea Society nomenclature for ocular surface rehabilitative procedures. Cornea 2011;30:1115–1119.

10   Meallet MA, Espana EM, Grueterich M et al. Amniotic membrane transplantation with conjunctival limbal autograft for total limbal stem cell deficiency. Ophthalmology 2003;110:1585–1592.

11   Santos MS, Gomes JA, Hofling-Lima AL et al. Survival analysis of conjunctival limbal grafts and amniotic membrane transplantation in eyes with total limbal stem cell deficiency. Am J Ophthalmol 2005;140:223–230.

12   Kenyon KR, Tseng SC. Limbal autograft transplantation for ocular surface disorders. Ophthalmology 1989;96:709–722; discussion 722–723.

13   Daya SM, Ilari FA. Living related conjunctival limbal allograft for the treatment of stem cell deficiency. Ophthalmology 2001;108:126–133; discussion 133–134.

14   Tsubota K, Satake Y, Kaido M et al. Treatment of severe ocular-surface disorders with corneal epithelial stem-cell transplantation. N Engl J Med 1999;340:1697–1703.

15   Ilari L, Daya SM. Long-term outcomes of keratolimbal allograft for the treatment of severe ocular surface disorders. Ophthalmology 2002;109:1278–1284.

16   Solomon A, Ellies P, Anderson DF et al. Long-term outcome of keratolimbal allograft with or without penetrating keratoplasty for total limbal stem cell deficiency. Ophthalmology 2002;109:1159–1166.

17   Pellegrini G, Traverso CE, Franzi AT et al. Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. Lancet 1997;349:990–993.

18   Di Iorio E, Ferrari S, Fasolo A et al. Techniques for culture and assessment of limbal stem cell grafts. Ocul Surf 2010;8:146–153.

19   Rama P, Matuska S, Paganoni G et al. Limbal stem-cell therapy and long-term corneal regeneration. N Engl J Med 2010;363:147–155.

20   Shortt AJ, Tuft SJ, Daniels JT. Ex vivo cultured limbal epithelial transplantation. A clinical perspective. Ocul Surf 2010;8:80–90.

21   Baylis O, Figueiredo F, Henein C et al. 13 years of cultured limbal epithelial cell therapy: A review of the outcomes. J Cell Biochem 2011;112:993–1002.

22   Nakamura T, Takeda K, Inatomi T et al. Long-term results of autologous cultivated oral mucosal epithelial transplantation in the scar phase of severe ocular surface disorders. Br J Ophthalmol 2011;95:942–946.

23   Nakamura T, Sotozono C, Bentley AJ et al. Long-term phenotypic study after allogeneic cultivated corneal limbal epithelial transplantation for severe ocular surface diseases. Ophthalmology 2010;117:2247–2254.e2241.

24   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;114:1294–1302.

25   Kolli S, Ahmad S, Lako M et al. Successful clinical implementation of corneal epithelial stem cell therapy for treatment of unilateral limbal stem cell deficiency. STEM CELLS 2010;28:597–610.

26   Baradaran-Rafii A, Ebrahimi M, Kanavi MR et al. Midterm outcomes of autologous cultivated limbal stem cell transplantation with or without penetrating keratoplasty. Cornea 2010;29:502–509.

27   Altman DG. Practical Statistics for Medical Research. London, U.K.: Chapman and Hall, 1991.

28   Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. Oxford, United Kingdom: Oxford University Press, 2008.

29   Shortt AJ, Secker GA, Rajan MS et al. Ex vivo expansion and transplantation of limbal epithelial stem cells. Ophthalmology 2008;115:1989–1997.

30   Tole DM, McKelvie PA, Daniell M. Reliability of impression cytology for the diagnosis of ocular surface squamous neoplasia employing the Biopore membrane. Br J Ophthalmol 2001;85:154–158.

31   Donisi PM, Rama P, Fasolo A et al. Analysis of limbal stem cell deficiency by corneal impression cytology. Cornea 2003;22:533–538.

32   Lange C, Feltgen N, Junker B et al. Resolving the clinical acuity categories "hand motion" and "counting fingers" using the Freiburg Visual Acuity Test (FrACT). Albrecht Von Graefes Arch Klin Exp Ophthalmol 2009;247:137–142.

33   Cantrill HL. The diabetic retinopathy study and the early treatment diabetic retinopathy study. Int Ophthalmol Clin 1984;24:13–29.

34 Stulting RD, Dupps WJ Jr, Kohnen T et al. Standardized graphs and terms for refractive surgery results. Cornea 2011;30:945–947.

35 Shortt AJ, Tuft SJ. Ocular surface reconstruction. Br J Ophthalmol 2011;95:901–902.

36 Cauchi PA, Ang GS, Azuara-Blanco A et al. A systematic literature review of surgical interventions for limbal stem cell deficiency in humans. Am J Ophthalmol 2008;146:251–259.

37 Liang L, Sheha H, Tseng SC. Long-term outcomes of keratolimbal allograft for total limbal stem cell deficiency using combined immunosuppressive agents and correction of ocular surface deficits. Arch Ophthalmol 2009;127:1428–1434.

38 Biber JM, Skeens HM, Neff KD et al. The Cincinnati procedure: Technique and outcomes of combined living-related conjunctival limbal allografts and keratolimbal allografts in severe ocular surface failure. Cornea 2011;30:765–771.

39 Daya SM, Watson A, Sharpe JR et al. Outcomes and DNA analysis of ex vivo expanded stem cell allograft for ocular surface reconstruction. Ophthalmology 2005;112:470–477.

40 Djalilian AR, Mahesh SP, Koch CA et al. Survival of donor epithelial cells after limbal stem cell transplantation. Invest Ophthalmol Vis Sci 2005;46:803–807.

41 Fitzpatrick R, Fletcher A, Gore S et al. Quality of life measures in health care. I: Applications and issues in assessment. BMJ 1992;305:1074–1077.

42 Miri A, Mathew M, Dua HS. Quality of life after limbal transplants. Ophthalmology 2010 117:638, 638.e1–638.e3.

43 Di Girolamo N, Bosch M, Zamora K et al. A contact lens-based technique for expansion and transplantation of autologous epithelial progenitors for ocular surface reconstruction. Transplantation 2009;87:1571–1578.

See www.StemCellsTM.com for supporting information available online.