OPO OPHTHALMIC AND PHYSIOLOGICAL OPTICS
THE JOURNAL OF THE COLLEGE OF OPTOMETRISTS

Ophthal. Physiol. Opt. 2009 **29:** 573–583

Review Article

Efficacy of different dry eye treatments with artificial tears or ocular lubricants: a systematic review

Michael J. Doughty and Sara Glavin

Department of Vision Sciences, Glasgow-Caledonian University, Glasgow, G4 OBA Scotland, UK

Abstract

Purpose: To objectively review the outcome of clinical studies where rose bengal stain (RB) has been used as an outcome measure to assess the efficacy of artificial tears (AT) in patients with dry eye.

Methods: From peer-reviewed articles published between 1947 and 2008, information was sought on dry eye status, as reported using a grading scheme, after use of RB as a diagnostic test, before and after use of a specific regimen of artificial tears or ocular lubricants for approximately 30 days. Mean baseline scores and post-treatment scores were calculated, along with the net change and the percentage change in the RB scores.

Results: From a total of 33 suitable data sets, published between 1985 and 2006, the group mean pre-treatment RB score was 4.25 ± 1.55 (±S.D.), which decreased to 2.84 ± 1.24 after 30 days of treatment. This represented a net change of −1.43 (95% CI of −1.04 to −1.45). For use of traditional AT (saline, hypromellose, etc), the net change was −0.95, it was −1.33 for use of carbomer (polyacrylic acid) gels and −2.10 for hyaluronic acid (HA) products. These changes represented net improvements of 25.9 ± 18.4%, 38.0 ± 20.7% and 41.8 ± 16.3% respectively. The greater change with HA was not associated with a lower final outcome score, but with higher pre-treatment scores.

Conclusions: Based on RB grading schemes used by numerous different clinicians over many years, treatment of dry eye with artificial tears or ocular lubricants can be expected to improve the condition of the exposed ocular surface. Assuming no improvement without treatment, a 30 days treatment period can be projected to produce an overall improvement of around 25%, but with no unambiguous statistical differences between product types.

Keywords: artificial tears, dry eye symptoms, keratoconjunctivitis sicca, rose bengal, Sjögren's syndrome

Introduction

Dry eye, which has many different causes, was likely once considered simply as a nuisance to some patients, and managed with various forms of palliative treatments (Volker-Dieben *et al.*, 1987). Most dry eye treatments are available as general sales list (GSL) products or P medicines and so all UK-registered optometrists can sell and supply these products to their patients (Doughty, 2007). However, by current perspectives, dry eye is considered not only to be a condition worthy of substantial attention in terms of patient well-being, but that there can be considerable patient benefit from active and timely intervention (Reddy *et al.*, 2004). With such open access to numerous types of artificial tear and ocular lubricant products, it is logical to consider not only what the difference might be in terms of product types but also

Received: 4 December 2008
Revised form: 13 March 2009
Accepted: 26 April 2009

Correspondence and reprint request to: Professor Michael J. Doughty. Tel.: +44 141 331 3393; Fax: +44 141 331 3387.
E-mail address: m.doughty@gcal.ac.uk

doi: 10.1111/j.1475-1313.2009.00683.x

what evidence can be used to demonstrate that the treatment has actually improved the dry eye condition.

The rose bengal (RB) test has been used for many years as an integral and fundamental part of the diagnosis of dry eye (reviewed in Doughty *et al*., 2007). With the emergence of the concept that dry eye is a complex disease, much attention has been paid to the overall utility and reliability of a range of clinical tests that might be used to assist a clinician in making a diagnosis of 'dry eye' (Pflugfelder *et al*., 2000; Bron *et al*., 2003; Nichols *et al*., 2004; Smith *et al*., 2008). If RB staining is considered to be an indicator of the ocular surface damage caused by dry eye, then a reduction in the RB score following treatment can be used as an indicator of an improvement in the condition (health) of the exposed ocular surface. It is likely that the RB test has been that most commonly selected to assess the efficacy of dry eye treatments in systematic assessments or clinical trials, and it is the data from such studies that have formed the basis for comparing different products. This is so that they can be marketed for use, or continue to be marketed when other alternatives become available.

For such studies, a patient with some form of dry eye disease (regardless of its aetiology) would likely have been diagnosed on the basis of a positive RB test, with a Schirmer test also likely to be used. The interest in this study was to assess how much this staining might change after a patient had then undergone a period of treatment with commonly available ophthalmic preparations that could be classified as artificial tears or ocular lubricants. It is not the goal of the current analyses to review the use of tests for diagnosis of the condition, nor to generally compare the use of different tests before or after the diagnosis. The focus of this systematic review is on the utility of RB staining in the assessment of commonly-used interventions to manage dry eye, namely the patient-driven use of artificial tears or ocular lubricants. It was of interest to know how predictable any change in RB staining might be and, with selection of studies using a similar period of treatment, how fast any such predictable changes might actually be. A secondary goal was to assess whether or not there were repeatable differences in this presumed beneficial effect, based on comparisons between different studies using the same or different tear replacement or lubricant products. Thirdly, such an analysis was undertaken to generally assess the overall consistency of reporting methods in these published studies.

Methods

This study is a literature-based review with the primary goal being to find published studies, on human subjects, which included data on RB staining before and after treatment of dry eye. Ophthalmic Literature (Institute of Ophthalmology, London) and the electronic MedLine (PubMed; http://www.ncbi.nlm.nih.gov) were used as primary sources, followed by reading of identified articles and relevant references cited within, and then also cross-checking the related articles listed in PubMed. Key words included rose bengal, artificial tears, conjunctiva and dry eye as well as the main ingredients of these types of products (e.g. hypromellose, carbomer and hyaluronic acid). This treatment could be either with commercially available dry eye products or preparations with nominally the same ingredients. No language limitation was imposed on the search, i.e. any article published in a foreign language was considered. In brief, publications were considered suitable for inclusion in the analysis largely if they provided some form of numerical scores for the RB staining before and after a specific period of treatment. No study report was excluded on the basis of size (i.e. number of participants in a trial). The treatment period was selected as being that most commonly reported on, namely close to 1 month (see Results). The basis of the scoring system should be indicated to be specifically or broadly similar to that accredited to Van Bijsterveld (1969), with a maximum of 9 points. The data were taken as reported, without any adjustments. Suitable studies should also include some detail of the cohort (e.g. in terms of age, although a few exceptions were included), some detail on the product used (e.g. as identified by brand name and/or its principal ingredients) and how often it was used, and for what period of time.

Data were entered into spreadsheets in SYSTAT v. 11.0 (Systat, Evanston, IL, USA) for compilation of descriptive statistics, tests for data normality (the default Shapiro–Wilk test, S–W) and the generation of graphical outputs. The overall goal of the analyses was to compare RB scores, before and after treatment for different interventions. Such within-subjects differences are presented using a paired *t*-test without Bonferroni adjustment, since only single comparisons were being made (although in all cases, the outcome of these paired *t*-tests were checked using non-parametric tests such as a Wilcoxon signed rank test). Some comparisons, to see if the overall outcomes were different, were made with multiple ANOVA. However, comparisons between discrete sets of data for different treatment effects (between-subjects differences) were generally made using the non-parametric rank order Mann–Whitney *U*-test chosen so as to allow for the fact that the data sets were unequally sized. A level of statistical significance was set at $p < 0.05$. The S.D. of the net changes were used to calculate a 95% confidence interval (CI).

Results

General characteristics of published articles reporting on the use of artificial tears in patients with dry eye

Overall, just over 500 publications were identified where RB eyedrops had been used on human subjects, but mostly as part of the diagnosis of some type of ocular surface damage. A similar number of papers were also identified as reporting on the effects of various types of artificial tears or lubricants on the human eye over the period covering 1947 through 2008, some of which also used RB staining as part of the diagnosis of the dry eye. However, only a relatively small number of these (41) actually included suitable numerical scores for RB staining before and after a specified period of use of some type of treatment that could be reasonably described as artificial tears or ocular lubricants (see below for details). If there was considerable uncertainty in the scoring system used for the RB test, then the paper was excluded (5). In addition, some studies (17) were not suitable for objective analyses because only relative or proportionate data were provided (e.g. numbers of patients showing improvement or reduced staining), or only descriptive terms were used for the staining (e.g. 'intermediate changes'). In some articles (12), while the protocol reported for the clinical study included RB staining, no actual data were included in a form which could be used for a structured analysis of the type chosen for the present study. In a few examples (4), as will be outlined below, only a change in RB score was provided. These were suitable for some analyses, although, as will be outlined (see Discussion), the ideal scenario would be to have baseline scores with some information of the circumstances under which these baseline assessments were made. In other examples, still not considered suitable for inclusion, there was simply insufficient detail on the actual scores and/or the RB scoring (grading) system used to allow for a before and after treatment comparison to be made. Lastly, it can be noted that some publications, especially older ones, clearly include RB staining as part of the diagnosis of the dry eye but do not include any details of its use in follow-up of the treated patients.

On the broad rationale that the effects of any intervention would be time-dependent, a follow-up period of 30 days (or close to this) was chosen. In a few examples (3) where the published data extended to longer periods of repeated follow-up, the outcomes were noted and, if possible, data were then estimated for a 28 day (1 month) period, e.g. from graphical presentations of data assuming that the intermediate effect was proportionately placed along the lines presented. Some studies (5) only reported on the outcome from a much longer period and so were not suitable for any estimations of an effect after 1 month. However, in many of the publications ultimately considered suitable for these analyses, specific data were actually provided for the specified follow-up period, e.g. in tables. Studies (7) that only extended to 1 or 2 weeks were also excluded (see Discussion). The present analyses were restricted to studies providing specific data (see above) on the outcome of intervention with some form of artificial tears or ocular lubricants. These were broadly categorised into traditional products (hypromellose etc), carbomer (polyacrylic acid)-based products and those containing hyaluronic acid. Studies assessing the effects of anti-inflammatory drugs (including NSAIDs, corticosteroids and cyclosporine A) were not considered.

Data assembly and analysis was based on the following strategy since two different approaches have been used in presentation of data from clinical studies on RB staining of the external eye. Firstly, the average scores for RB staining at baseline and at the follow up assessment were entered into spreadsheets in SYSTAT v. 11 (Systat).These were entered along with a code for the type of product used. From these sets of scores, group mean values could be calculated, as well as the magnitude of the net change in RB scores for any particular cohort. In addition, relative changes in RB scores were calculated as a percentage change from baseline, e.g. if the score was reduced by 1 from 5 to 4, then this would be a 20% reduction. In addition, this percentage change was used as an assessment of the efficacy of the intervention, and now presented in positive terms, i.e. a 20% improvement in RB scores.

Overall summary of data sets

Over the time period from 1947 to the end of 2008, a total of just 33 suitable data sets could be identified. In some instances, more than one data set was obtained from a single publication; either where a comparison was being made between products, or because two different cohorts were being assessed in the same study. Fourteen of these publications had the study participants using what will be termed traditional or older artificial tears. Such products were either the primary treatment or were being used as a comparator or reference product in an assessment that was mainly undertaken to assess other types of interventions (including anti-inflammatory drugs). These simple or older forms of artificial tears ranged from saline eyedrops (Shimmura *et al.*, 1995; Tananuvat *et al.*, 2001) to similar and presumably isotonic formulations containing polymers such as hydroxyethylcellulose (HEC) (Vadillo Gonzalo *et al.*, 1985); hypromellose (also known as hydroxypropylmethylcellulose or HPMC) (Prause, 1986; Toda *et al.*, 1996; Iester *et al.*, 2000; Chang *et al.*, 2005); hypromellose with dextran

(Simon Castellvi *et al*., 1989); polyvinyl alcohol (also known as PVA) (Nelson and Farris, 1988; Vico *et al*., 2005); or a combination of PVA with povidone (also known as polyvinylpyrrolidone) (Laibovitz *et al*., 1993; McDonald *et al*., 2002; Avunduk *et al*., 2003). One study simply indicated the use of 'artificial tears' (Tsifetaki *et al*., 2003) and this could indicate either that traditional products were used or that a range of different and unspecified products were actually used; the former perspective was taken for these analyses. A total of seven data sets were identified with suitable information on the use of some type of carbomer (polyacrylic acid)-based formulation (Laroche *et al*., 1991; Lepri *et al*., 1991; Sullivan *et al*., 1997; Troiano *et al*., 1997; Chiambaretta *et al*., 2004; Ostuni *et al*., 2005). Finally, 10 sets of data were found on the use of hyaluronate-based polymer formulations. Some of these were experimental (non-marketed) products (Nelson and Farris,1988; Shimmura *et al*., 1995; Troiano *et al*., 1997), while others were similar to, or modified from, commercially-available products (Condon *et al*., 1999; Iester *et al*., 2000; Papa *et al*., 2001; Aragona *et al*., 2002a,b; McDonald *et al*., 2002; Vico *et al*., 2005).

Overall characteristics and strategy for considering outcomes of dry eye treatment studies

The 33 data sets included evaluation of a total of 1293 individuals before and after treatment, with the number in any particular study ranging from just 8 to 139 (mean 40 subjects/study data set). For 24 of these data sets, age data were provided such that an average age for each cohort (or close to this) could be gleaned. These average ages ranged from 44 to 65 years (group mean 56 years), so indicating that most of the studies did indeed assess individuals considered to be characteristic of those with dry eye surface disease. Based on the patient group profile, it was considered likely that the other nine studies also evaluated older subjects or included a reasonable proportion of older subjects. Omission of adequate age data occurred for reports using traditional artificial tears (3), studies on carbomer products (2) and on hyaluronic acid (4). As indicated earlier, the strategy for data selection was to find reports where the follow-up period was close to 28 days (1 month). For the 33 data sets, the mean follow-up time period was 29 days (range 21–35 days, median 30 days). Of the 33 data sets, 29 of them provided details on RB scores before and after treatment, with the additional four studies only providing data on the change in RB score (with no details of baseline or final RB scores). While most publications provide the raw RB scores in tabular or graphic form, some provide baseline data and then present the treatment outcome in terms of a percentage change from baseline. In such cases, a calculation could be made to estimate the post-treatment RB scores.

Effects of treatments with different types of artificial tears and ocular lubricants

The net reduction in RB scores after 30 days of use of any type of artificial tears or ocular lubricants was from a median value of 4.40 to a median value of 2.40 (*Figure 1*). The range of pre-treatment (baseline) and post-treatment (treated) RB scores was similar (as indicated by the height of the box plot ±25% inter-quartile) range and there were no outliers. The group mean ± S.D. RB baseline scores were 4.23 ± 1.55 and the treated RB scores were 2.84 ± 1.24. A paired *t*-test indicated that this reduction in RB scores was highly statistically significant ($p < 0.001$), with the mean difference between each set of data being a change in RB score of −1.43 (estimated 95% CI of between −1.04 and −1.81). The percentage improvement in RB scores was 33.4 ± 19.1%. This overall analysis indicates that treatment with any type of artificial tears or ocular lubricants over a 30 day period can be expected to reduce the severity of ocular surface damage, at least as based on RB scores. However, this outcome is not necessarily the case as some studies (3) reported no improvement or only very small reductions on RB scores after treatment. The next stage of the analyses was then to more specifically consider whether it could be shown if there were consistent differences in the outcome of different types of treatment. It should be stressed that this analysis only considers one treatment period, namely that of approximately 30 days. The data were selected on this basis so as to standardise the comparisons. Assuming that any treatment would be time-dependent, the outcome of these analyses should be interpreted in terms of the rate of improvement, i.e. if the treatment produced greater effects, then this reflects a faster rate of improvement, and *vice versa*.

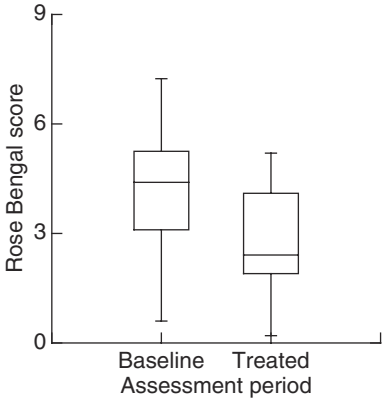

Figure 1. Box plot to show change in RB staining before (baseline) and after 30 days (treated) of use of any form of artificial tears or ocular lubricants.

For the 30 day use of traditional or older artificial tear or lubricant products, the median baseline score in the 13 studies reporting suitable data was 4.2. This was reduced to 2.4 after 30 days (*Figure 2*). Again, the interquartile interval across the different studies was broad, but the data were uniformly distributed with no outliers. However, the group mean baseline scores changed by a lesser amount from 4.03 ± 1.88 to 3.09 ± 1.47, a net change of just −0.95 (95% CI −0.45 to −1.45). Notwithstanding, this reduction in RB scores was statistically significant ($p = 0.001$). This overall outcome was not changed if the study reporting on use of 'artificial tears' was excluded.

Analysis of the seven studies reporting use of carbomer gels indicated a more variable outcome, at least as assessed on the reduction in RB scores (*Figure 3*). The RB scores at baseline appeared reasonably homogeneous with a median value of 3.7, and as a result of treatment the median value was reduced to 1.90. However, as can be seen from the box plot, there were two distinct outliers indicating much less and also a much greater effect. As a result, the group mean RB scores changed from 3.53 ± 0.82 to 2.20 ± 1.05, for a net change of −1.33 (95% CI of −0.42 to −2.24) that was only just statistically significant (Wilcoxon signed rank test, $p = 0.027$).

Analysis of the data from 10 studies reporting on the use of hyaluronic acid-based products is shown in *Figure 4*. For these studies, the median baseline score was the highest at 4.90 and this was reduced to 3.00 after 30 days treatment. The group mean scores were 5.01 ± 1.16 at baseline and 2.92 ± 0.97 after treatment, there were no outliers, and the net change was −2.10 (95% CI of between −1.41 and −2.78). This reduction in RB scores by hyaluronic acid based products was highly statistically significant ($p < 0.001$).

Comparisons between treatments with different types of artificial tears and ocular lubricants

The above analyses (*Figures 2–4*) indicate some differences in the treatment efficacy of different types of artificial tear products or ocular lubricants in being able to change the RB staining. The next step was to make more specific comparisons to see how robust (or otherwise) any such apparent differences in efficacy might be. With the change (reduction) in RB scores being the most commonly reported outcome measure, comparisons based on this can be made between product types (*Figure 5*). As can be seen, treatments with traditional artificial tears or ocular lubricants (AT) have yielded similar results to treatment with carbomer-based products (GEL) while more substantial changes in RB scores have been achieved with the use of hyaluronic acid based products (HA). The numbers of reported studies included in this analysis is slightly greater than earlier (*Figures 2*–4) since studies that just reported the change in RB scores can now also be included. The net reduction in RB scores (group mean values ±S.D.) for

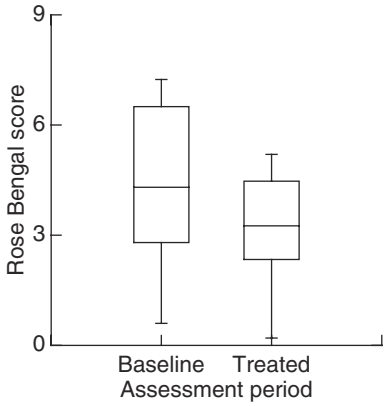

Figure 2. Box plot to show change in RB staining before (baseline) and after 30 days (treated) of use of traditional or older artificial tears or ocular lubricants.

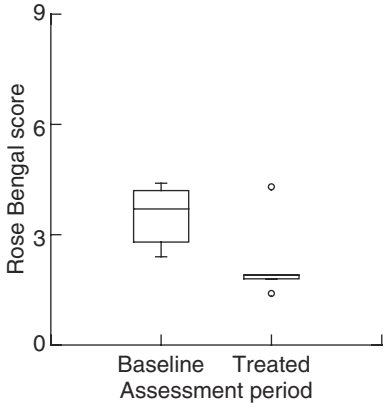

Figure 3. Box plot to show change in RB staining before (baseline) and after 30 days (treated) of use of carbomer gel-based products.

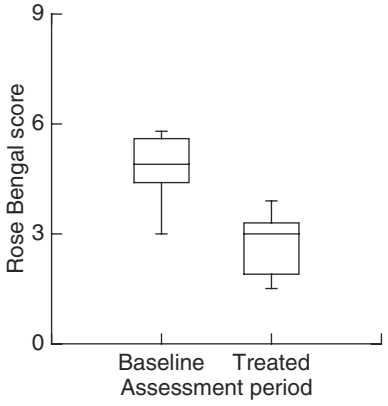

Figure 4. Box plot to show change in RB staining before (baseline) and after 30 days (treated) of use of hyaluronic acid-based products.

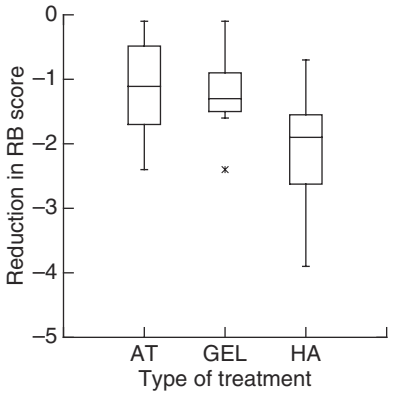

Figure 5. Box plot to show differences in effect of various types of treatment as given by the net reduction in RB scores after 30 days. AT, traditional artificial tears; GEL, carbomer-based products; HA, hyaluronic acid-based products.

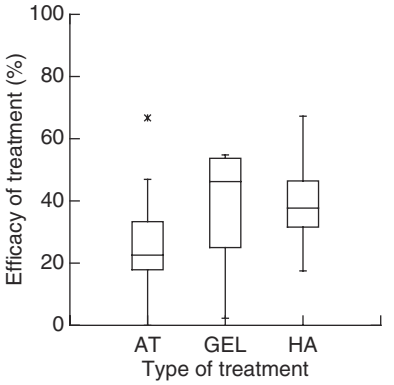

Figure 6. Box plot to show differences in effect of types of treatment according the percentage improvement in RB scores after 30 days (treatment efficacy %). AT, traditional artificial tears; GEL, carbomer-based products; HA, hyaluronic acid-based products.

the treatments were now −1.08 ± 0.806 for traditional artificial tears, −1.23 ± 0.73 for carbomer gels and −2.11 ± 0.91 for hyaluronic acid based products. There was no difference between the use of traditional artificial tears and carbomer gels ($p = 0.3610$), but there was a substantial difference between use of traditional artificial tears and the hyaluronic acid based products ($p = 0.006$). This result would almost be expected since the net change in RB score in the HA studies was almost twice that reported for the use of traditional artificial tears. It should be noted however that the baseline data in the HA studies is also much higher (*Figure 4*, and see *Figure 7* later). There was a smaller but still just statistically significant difference between hyaluronic acid based products compared to carbomer gels ($p = 0.037$). A multiple ANOVA, comparing outcomes using the different treatments revealed no statistical difference between the three options ($p > 0.1$).

An alternative method of examining efficacy is to compare the treatment vs baseline scores on a percentage basis and this analysis is given in *Figure 6*. The overall result has similar characteristics to that found if the reduction in RB scores is used as the measure of treatment effect, but the overall result fails to reveal any statistical significance between treatments. The group mean (± S.D.) percentage improvements in RB scores were 25.9 ± 18.4% for traditional artificial tears (AT), 38.0 ± 20.7% for carbomer-based products (GEL) and 41.8 ± 16.3% for hyaluronic acid-based products (HA). There was no statistically significant difference (Kruskall–Wallis test) between traditional artificial tears and carbomer gels ($p = 0.137$) or carbomer gels vs hyaluronic acid based products ($p = 0.906$). The apparent difference between use of traditional artificial tears and HA products just failed to reach statistical significance ($p = 0.059$). As with the comparisons for changes in absolute scores, multiple ANOVA, comparing outcomes using the different treatments revealed no statistical difference in the relative changes in RB scores between the three treatment options ($p > 0.1$).

Two further analyses are presented on the pre- and post-treatment characteristics of the study groups (*Figures 7* and *8*). *Figure 7* shows a comparison of the baseline RB scores for the different studies. The box plot clearly shows that the cohorts studied appear to differ rather substantially in terms of the overall severity of the dry eye disease prior to treatment. While studies using traditional artificial tears have assessed a very wide range of presentations of dry eye, as based on severity of RB staining, those used to assess the efficacy of carbomer gels were lower and had a narrower range. Studies assessing the effects of use of hyaluronic acid based products clearly involved patients with more severe dry eye. As partially analysed earlier, the studies using traditional artificial tears had a mean baseline RB score of 4.1 (range 0.60–6.58), those assessing carbomer gels 3.53 (range 2.4–4.7)

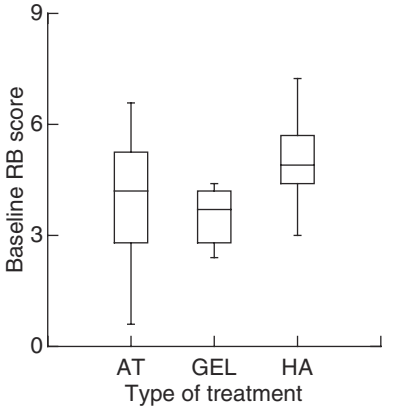

Figure 7. Box plot comparing baseline RB scores for the different studies according to treatment type. AT, traditional artificial tears; GEL, carbomer-based products; HA, hyaluronic acid-based products.

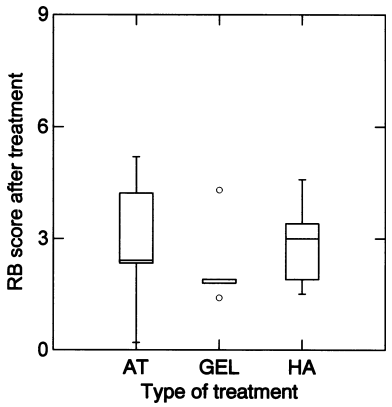

Figure 8. Box plot to show outcome, as based on RB scores after treatment, for the different studies according to the type of product used. AT, traditional artificial tears; GEL, carbomer-based products; HA, hyaluronic acid-based products.

and this baseline score was 5.01 (range 3.00–7.24) for hyaluronic acid studies. The post-treatment scores, according to the treatment type, are given in *Figure 8*. The final outcome measures are thus similar being 3.08 (range 0.2–5.20) for traditional artificial tears, 2.20 (range 1.4–4.30) for carbomer, and 2.92 (range 1.5–4.59) for hyaluronic acid products. Indeed, the median value for the post-treatment RB scores in the HA studies was the highest (and thus the absolute outcome after 30 days of treatment would be the worst).

Discussion

The main reason for undertaking this analysis was simply because this type of data could not be located in a peer-reviewed publication related to dry eye treatments. The RB test was selected for these analyses because of the long history of its use and so providing the maximum chance of obtaining enough published data to see if any consistent differences could be seen between products. It is beyond the scope of this paper to consider other diagnostic tests for dry eye but analyses so far (unpublished) indicate that there is even less published data on the use of other tests (as opposed to assessing symptoms), and an even greater inter-study variability in outcomes after dry eye treatments. While the analyses presented here indicate that a degree of caution is needed in trying to make any comparisons between the apparent efficacy of artificial tear preparations, the data obtained with RB seems to be the most consistent. It is important to note however that the outcome depends on whether changes in absolute scores or relative changes in the RB scores are chosen. The analyses indicate that it is the absolute scores that need to be compared for different products, rather than just noting the magnitude of the change or difference. However, from a patient management perspective, it seems more logical to consider the relative improvement that has occurred (see below).

The main goal for the current analyses was to objectively and systematically assess whether there was a consistent or predictable outcome from intervention in management of dry eye patients. For most patients with dry eye, initial intervention will take the form of the use of some type of artificial tears. Alternative patient management strategies might also include physical interventions such as punctal plugs, eye shields (goggles) or even overnight eyelid closure. In some patients, there could also be the prescription of oral medications to increase lacrimal output and/or reduce inflammation of the lacrimal glands or ocular surface. None of these alternative options were considered further as part of this review, mainly because there are not enough data (including rose bengal scores) available in the literature to undertake a systematic analysis across multiple studies. The objective of the present review was to concentrate on the use of commercially available artificial tear products in patients with clinically significant signs of dry eye. Many of the types of products evaluated in this review are currently marketed in the UK. These include traditional artificial tears (saline, hypromellose etc), carbomer gels, as well as some newer hyaluronic acid-based products (Doughty, 2008).

Over some 50 or so years of use, it can be noted that different clinicians have either considered whether the RB staining was present or not, or felt that the staining could also be described using a range of general terms such as moderate or intermediate or intense. Examples can even be found where the number of discrete stained spots were counted (Török, 1985). However, as based on a grading scheme that was used at least 40 years ago (Van Bijsterveld, 1969), the RB staining can be assigned a numerical grade. In terms of descriptions given in peer-reviewed publications, it can be surmised that discrete punctate staining could be given a grade of 1, while confluent intense staining would be grade 3, as compared to no staining which would be given a score of zero. By adding up the scores from cornea and the visible bulbar conjunctiva either side of the cornea, a maximum score of 9 could be recorded (Van Bijsterveld, 1969). By comparing patients with Sjögren's disease to a reference cohort, a cut of 3.5 (on the scale of 9) was advocated to be able to identify those with true dry eye (Van Bijsterveld, 1969). Such a dry eye patient would presumably then be designated as having a positive RB test. This designation should be carefully distinguished from patients who might only have some symptoms of dry eye (Doughty *et al.*, 1997).

The extent of staining after use of RB was used as the end point in these analyses. It was chosen as the end point simply because it appears to be the only test for the condition of the ocular surface that has been

consistently used for both diagnosis and follow-up of dry eye patients in systematic assessments of the apparent efficacy of the type of dry eye treatments considered here. Regardless of opinions as to the overall utility or reliability of the RB test (Korb, 2000; Khan-Lim and Berry, 2004; Smith *et al*., 2008), very few studies have been carried out to compare RB staining with other possible options. This applies especially to assessing the outcome of the effects of artificial tears or ocular lubricant therapies in dry eye patients. However, assessment with fluorescein has been noted to give a remarkably different result to assessment with RB (Shimmura *et al*., 1995): a fluorescein assessment was taken to indicate a very substantial improvement following use of artificial tears while no improvement was seen in RB scores. Further studies are clearly needed to resolve such marked discrepancies. Furthermore, despite the modest number of reported studies using RB over a 20 year period and the somewhat ambiguous nature of the overall outcome, it would seem prudent to continue with the use of RB in dry eye treatment efficacy studies. Advocates of any alternative to RB stain (e.g. lissamine green) should take note that there is far less published information on how treatment with artificial tears or ocular lubricants might alter (improve) any other staining of the conjunctiva or cornea. If, as reported in depth in these analyses, it is not easy to distinguish between different types of artificial tears, then it is simply not possible to make a credible argument that any other stain is superior to rose bengal. While lissamine green might again be being considered for routine use (Versura *et al*., 2006), there is a lack of similar analyses and work still to be done to establish normal values for lissamine green staining and its use as a measure for treatment outcomes.

For this study, every effort was made to try to be as consistent as possible in the comparisons made, with RB being used as the diagnostic test for treatment efficacy. Obviously, the key aspect of this is the RB grading system itself. As noted earlier, the strategy adopted was to identify data sets that used an RB scoring system that was at least similar to that used in 1969 by Van Bijsterveld. Its existence and its 'classical' use is noted by Lemp (1995), and with a score of between 0 and 3 applied to each side of the exposed conjunctiva and also to the cornea, for a maximum of 9 points. It needs to be noted that almost no detail is actually provided in the 33 studies included here of the RB formulation used nor how it was used (see Doughty *et al*., 2007). However, after a careful search of the literature, a comment has to be made that there does not appear to have been a systematic study of the level (intensity) of staining that should be assigned to each grade. Equally importantly, a study could not be located where the intra-subject repeatability of the staining has been assessed (allowing, of course for some appropriate interval to elapse between test and re-test) nor on the agreement between different clinicians on assignments of a particular grade. This situation appears to exist to this day, even with variants of the Van Bijsterveld scheme. For example, Lemp (1995) proposed an 18 point scheme which only assesses the conjunctiva. The exposed conjunctival surface is divided into 6 zones and a score of 0–3 assigned for each. The assignment of grading is facilitated by the inclusion of a set of boxes with a few, a moderate number and a high number of dots. A further development of the RB scoring has been proposed in which the exposed conjunctiva is treated as a unit but now a 6 point grading scheme is used (0–5). In this case, the user is helped in the assignment of grading by a series of diagrams with increasing density of dots (Bron *et al*., 2003). The increase in the number of dots in transition from one grade to another is based on a logarithmic scale. However, as with the Lemp (1995) proposed scheme, no specific comparisons appear to have been undertaken to compare with the Van Bijsterveld (1969) scheme. It should also be noted that the number of dots given in the Oxford scheme (Bron *et al*., 2003) is not obviously based on clinical evidence, e.g. a study in which the actual number of RB stained dots on a series of patients were counted (Török, 1985).

Another area of difficulty in trying to use even RB data to compare artificial tear products is that there are differences in the baseline (pre-treatment) scores. This difference, in itself, is a likely reason (and an unacceptable bias) as to why the hyaluronate-based products can be shown to provide a slightly superior outcome to use of other products if changes in absolute RB scores are used for the analyses. Until there is better standardization of entry criteria for assessment of dry eye treatments, then it will continue to be difficult to make robust comparisons or even claims about the efficacy of one artificial tear product compared to any other. In future assessments on suitably sized cohorts, a goal should be to assess how much any treatment effect might depend on the baseline severity of the condition, especially when comparing different treatment options.

The earliest published article that could be found that met the inclusion criteria was, quite surprisingly, only from nearly 25 years ago (Vadillo Gonzalo *et al*., 1985). This included data on an assessment of the effects of 30 days treatment with a hydroxyethylcellulose (HEC)-based artificial tears. While even the earliest studies found do not obviously stipulate why a 30 day follow-up period was chosen, it is assumed that this was one of convenience. Overall, however, for most of the published studies considered suitable for analysis, the period of follow-up for use of the artificial tears was around 30 days. This was therefore selected as an approximate treatment period that would be used for further com-

parisons. Since a 30 day period of follow-up was used in the earliest study that could be located, subsequent publications where the period of follow up was substantially less that this (even if RB scores were provided) were excluded, i.e. those reporting on outcomes after just 1 or 2 weeks of intervention. The rationale for exclusion of these papers was based both on the overall outcome for longer periods of follow-up (unpublished analyses) and because of a general perspective that a real improvement in the condition of the dry eye-damaged ocular surface seemed unlikely with such a short period of treatment (even if symptoms were substantially improved). Based on graded conjunctival impression cytology samples, evidence has been presented that a month of sodium hyaluronate-based eyedrops produced no better effect than the use of saline eyedrops, but that there was significant improvement in the hyaluronate treated conjunctival cell scores if used for longer periods (Aragona *et al*., 2002b). It should be noted however that there is a dearth of published data on the outcome of inconsistent use of artificial tears or ocular lubricants, either assessed on patient symptoms, by rose bengal staining or other techniques such as impression cytology (Blades and Doughty, 2000; Aragona *et al*., 2002b).

This literature search and analyses reveal that there are various issues that need to be considered about standardization of RB pharmaceuticals and their use, use of grading schemes, disease severity in patients selected for study, and patient compliance in any treatment study. The data selected for these analyses might be considered too fragile (i.e. non-robust) to be of any use. However, these analyses do show that the regular use of artificial tears or ocular lubricant can apparently improve the condition of the exposed ocular surface. This conclusion should apply regardless of whether RB is considered to be a vital stain in the traditional sense (i.e. it stains devitalized cells or dead cells) or whether it is considered that the ability of RB to stain cells is somehow also dependent on whether the cell surface is protected by a coating of mucus (Feenstra and Tseng, 1992). While the latter concerns over mechanism have yet to be demonstrated for cells on the surface of the eye, one overall outcome of the present analyses is that a treatment with artificial tears can be expected to reduce the extent or severity of RB staining. Such an overall conclusion is perhaps somewhat speculative, although there do not appear to be detailed published studies of this type that actually report on the outcome of RB staining assessments without any intervention with some form of artificial tears or ocular lubricants. While it has also yet to be established if the active principle (or any other ingredients of artificial tears) alters or interferes with rose bengal staining of conjunctival cells *in situ*, a certain magnitude of reduction in RB staining appears to be predictable. All data were taken from studies with a similar treatment regimen and treatment period. Further studies can now be undertaken, especially in relation to whether there are any predictable artificial tear-rose bengal interactions that might influence the staining process.

Overall, while the magnitude of the net reduction in RB staining (after a constant 4 weeks of treatment) might be dependent on the severity of the RB staining prior to treatment (i.e. on the severity of the ocular surface disease), there appears to be little difference between types of treatment. The overall net reduction in RB staining after 4 weeks of treatment was 33% across the 33 studies. The data was however rather heavily skewed in some treatments (see *Figure 5*), so it seems more reasonable to use a net treatment effect of a 25% improvement in RB staining with 1 month of treatment. With this as a perspective, along with better standardization of entry criteria in terms of RB staining, it would be most useful to establish if this trend is actually continued, i.e. whether or not a further 25% improvement would occur over a second month of treatment.

Acknowledgements

The authors have no proprietary interests in any of the products or types of products covered in this review. The authors would like to thank Dr Santosh Khanal (formerly of Glasgow Caledonian University) and Dr Robert Harvey (Ophthalmology Unit, Paisley Hospital) for useful discussions, insight and commentary on earlier drafts of these analyses.

References

Aragona, P., Di Stefano, G., Ferreri, F., Spinella, R. and Stilo, A. (2002a) Sodium hyaluronate eye drops of different osmolarity for the treatment of dry eye in Sjögren's syndrome patients. *Br. J. Ophthalmol*. **86,** 879–884.

Aragona, P., Papa, V., Micali, A., Santocono, M. and Milazzo, G. (2002b) Long term treatment with sodium hyaluronate-containing artificial tears reduces ocular surface damage in patients with dry eye. *Br. J. Ophthalmol*. **86,** 181–184.

Avunduk, A. M., Avunduk, M. C., Varnell, E. D. and Kaufman, H. E. (2003) The comparison of efficacies of topical corticosteroids and nonsteropidal anti-inflammatory drops on dry eye patients: a clinical and immunocytochemical study. *Am. J. Ophthalmol*. **136,** 593–602.

Blades, K. and Doughty, M. J. (2000) Comparison of grading schemes to quantitative assessments of nucleus-to-cytoplasmic ratios for human bulbar conjunctival cells collected by impression cytology. *Curr. Eye Res*. **20,** 335–340.

Bron, A. J., Evans, V. E. and Smith, J. A. (2003) Grading of corneal and conjunctival staining in the context of other dry eye tests. *Cornea* **22,** 640–650.

Chang, Y.-H., Lin, H.-J. and Li, W.-C. (2005) Clinical evaluation of the traditional Chinese prescription Chi-Ju-Di-Huang-Wan for dry eye. *Phytother. Res.* **19,** 349–354.

Chiambaretta, F., Pouliquen, P., Menerath, J. M., Pilotaz, F., Rebika, H. and Rigal, D. (2004) Efficacite et tolerance d'un gel de carbomere fluide versus un gel de carbomere classique lors dur traitement du syndrome sec. *J. Fr. Ophtalmol.* **27,** 130–135.

Condon, P. I., McEwen, C. G., Wright, M., Mackintosh, G., Prescott, R. J. and McDonald, C. (1999) Double-blind, randomised, placebo-controlled, cross-over, multicenter study of determine the efficacy of a 0.1% (w/v) sodium hyaluronate (Fermavisc) in the treatment of dry eye syndrome. *Br. J. Ophthalmol.* **83,** 1121–1124.

Doughty, M. J. (2007) The scope and options on use of 'entry level ophthalmic drugs' by UK-based optometrists. *Optom. Today* **47,** 56–60.

Doughty, M. J. (2008) Artificial tears and ocular lubricants. In: *Drugs, Medications and the Eye*, 17th edn (ed. M. J. Doughty), Smawcastellane Information Services, Helensburgh, Scotland.

Doughty, M. J., Fonn, D., Richter, D., Simpson, T., Caffery, B. and Gordon, K. (1997) A patient questionnaire approach to estimating the prevalence of dry eye symptoms in patients presenting to optometric practices across Canada. *Optom. Vis. Sci.* **74,** 624–631.

Doughty, M. J., Lee, C.-A., Ritchie, S. and Naase, T. (2007) An assessment of the discomfort associated with the use of rose bengal 1% eyedrops on the normal human eye – a comparison with saline 0.9% and a topical ocular anaesthetic. *Ophthal. Physiol. Opt.* **27,** 159–167.

Feenstra, R. P. G. and Tseng, S. C. G. (1992) What is actually stained by rose bengal? *Arch. Ophthalmol.* **110,** 984–993.

Iester, M., Orsoni, G. J., Gamba, G., Taffara, M., Mangiafico, P., Guiffrida, S. and Rolando, M. (2000) Improvement of the ocular surface using hypotonic 0.4% hyaluronate acid drops in keratoconjunctivitis sicca. *Eye* **14,** 892–898.

Khan-Lim, D. and Berry, M. (2004) Still confused about rose bengal? *Curr. Eye Res.* **29,** 311–317.

Korb, D. R. (2000) Survey of preferred tests for diagnosis of the tear film and dry eye. *Cornea* **19,** 483–486.

Laibovitz, R. A., Solch, S., Andriano, K., O'Connell, M. and Silverman, M. H. (1993) Pilot trial of cyclosporine 1% ointment in the treatment of keratoconjunctivitis sicca. *Cornea* **12,** 315–323.

Laroche, L., Arata, M., Brasseur, G., Lagoute, F., Le Mer, Y., Lumbroso, P., Mercane, M., Normand, F., Rigal, D., Roncin, S., Royer, J. and Adrien, J. L. (1991) Traitement de la sécheresse oculaire par gel lachrymal: étude multicentrique randomisée. *J. Fr. Ophtalmol.* **14,** 321–326.

Lemp, M. A. (1995) Report of the National Eye Institute/Industry workshop on clinical trials in dry eyes. *CLAO J.* **21,** 221–232.

Lepri, A., Salvatori, E., Gloria, E. M., Bartalena, L., Marcocci, C., Bruno-Bossion, G., Bagazzi, F., Pichera, A. and Luciani, G. (1991) A new artificial tear gel preparation in the treatment of mild Graves' ophthalmolopathy. In: *The Lacrimal System* (eds O. P. Van Bijsterveld, M. A. Lemp and D. Spinelli). Kugler & Ghedini Publ., Amsterdam, pp. 51–55.

McDonald, C. C., Kaye, S. B., Figueiredo, F. C., Macintosh, G. and Lockett, C. (2002) A randomised, crossover, multicenter study to compare the performance of 0.1% (w/v) sodium hyaluronate with 1.4% (w/v) polyvinyl alcohol in the alleviation of symptoms associated with dry eye syndrome. *Eye* **16,** 601–607.

Nelson, J. D. and Farris, R. L. (1988) Sodium hyaluronate and polyvinyl alcohol artificial tear preparations. A comparison in patients with keratoconjunctivitis sicca. *Arch. Opthalmol.* **106,** 484–487.

Nichols, K. K., Mitchell, G. L. and Zadnik, K. (2004) The repeatability of clinical measurements of dry eye. *Cornea* **23,** 272–285.

Ostuni, P., Battista, M. E. C. and Furlan, A. (2005) Efficacia di Carbopol 974P (Siccafluid) nel trattamento della cheratoconjunctivite sicca di grado moderato-severo in 60 casi con S. di Sjögren primitive. *Reumatismo* **57,** 119–124.

Papa, V., Aragona, P., Russo, S., Di Bella, A., Russo, P. and Milazzo, G. (2001) Comparison of hypotonic and isotonic solutions containing sodium hyaluronate on the symptomatic treatment of dry eye patients. *Ophthalmologia* **215,** 124–127.

Pflugfelder, S. C., Solomon, A. and Stern, M. E. (2000) The diagnosis and management of dry eye: a twenty-five-year review. *Cornea* **19,** 644–649.

Prause, J. U. (1986) Treatment of keratoconjunctivitis sicca with Lacrisert. *Scand. J. Rheumatol.* **61**(Suppl)**,** 261–263.

Reddy, P., Grad, O. and Rajagopalan, K. (2004) The economic burden of dry eye. A conceptual framework and preliminary assessment. *Cornea* **23,** 751–761.

Shimmura, S., Ono, M., Shinozaki, K., Toda, I., Takamura, E., Mashima, Y. and Tsubota, K. (1995) Sodium hyaluronate eyedrops in the treatment of dry eyes. *Br. J. Ophthalmol.* **79,** 1007–1011.

Simon Castellvi, J. M., Verges, C., Camins, J. L. and Pita Salorio, D. (1989) Tratimento del syndrome de ojo seco. Analisis comparative. *Arch. Soc. Esp. Oftalmol.* **56,** 185–192.

Smith, J., Nichols, K. K. and Baldwin, E. K. (2008) Current patterns in the use of diagnostic tests in dry eye evaluation. *Cornea* **27,** 656–662.

Sullivan, L. J., McCurrach, F., Lee, S., Taylor, H. R., Rolando, M., Marechal-Courtois, C., Creuzot-Garcer, C., Easty, D. L., Karabatsas, C., Hoh, M. B., Faschinger, C. and Laroche, L. (1997) Efficacy and safety of 0.3% carbomer gel compared to placebo in patients with moderate-to-severe dry eye syndrome. *Ophthalmology* **104,** 1402–1408.

Tananuvat, N., Daniell, M., Sullivan, L. J., Yi, Q., McKelvie, P., McCarty, D. J. and Taylor, H. R. (2001) Controlled study of the use of autologous serum in dry eye patients. *Cornea* **20,** 802–806.

Toda, I., Shinozaki, N. and Tsubota, K. (1996) Hydroxypropyl methylcellulose for the treatment of severe dry eye associated with Sjögren's syndrome. *Cornea* **15,** 120–128.

Török, M. (1985) Hydroxypropylmethylhaltige Augentropfen in der Behandlung des "Sicca" Syndroms. *Klin. Monatsbl. Augenheilkd.* **187,** 60–62.

Troiano, P., Cavallotti, B., Oldani, A., Iraci, M., Galli, L. and Miglior, M. (1997) The preservation of the ocular surface during chronic administration of hypotensive eyedrops. *Orbit* **16**(Suppl)**,** 87–90.

Tsifetaki, N., Kitsos, G., Paschides, C. A., Alamanos, Y., Eftaxias, V., Voulgari, P. V., Psilas, K. and Drosos, A. A. (2003) Oral pilocarpine for the treatment of ocular symptoms in patients with Sjögren's syndrome: a randomised 12 week controlled study. *Ann. Rheum. Dis.* **62,** 1204–1207.

Vadillo Gonzalo, M. M., Abad, J. M., Verges Roger, C. and Pita Salorio, D. (1985) Effectos terapeuticos del tear gard en el sindrome de ojo seco. *Arch. Soc. Esp. Oftalmol.* **48,** 299–302.

Van Bijsterveld, O. P. (1969) Diagnostic tests in the sicca syndrome. *Arch. Ophthalmol.* **82,** 10–14.

Versura, P., Frigato, M., Mulé, R., Malavolta, N. and Campos, E. C. (2006) A proposal of new ocular items in Sjögren's syndrome classification criteria. *Clin. Exp. Rheumatol.* **24,** 567–572.

Vico, E., Querda, A., Benitez-del-Castillo, J. M., Fernandez, C. and Garcia-Sanchez, J. (2005) Estudio comparative entre el hialuronato sodico al 0,15% y el alcohol polivinilico como tratimento para el ojo seco. *Arch. Soc. Esp. Oftalmol.* **80,** 387–394.

Volker-Dieben, H. J., Kok-Van Alphen, C. C., Hollander, J., Kruit, P. J. and Batenburg, K. (1987) The palliative treatment of dry eye. *Doc. Ophthalmol.* **67,** 221–227.