Biomarkers In Dry Eye Disease

Michael A Lemp,1 Benjamin D Sullivan2 and Leslie A Crews2

1. Chief Medical Officer, TearLab Corporation, California and Clinical Professor of Ophthalmology, Department of Ophthalmology, Georgetown University and Department of Ophthalmology, George Washington University, Washington DC; 2. Chief Scientific Officer, TearLab Corporation, California; 3. Consultant, TearLab Corporation, California, US

Abstract

Dry eye disease is a multifactorial chronic disorder of the ocular surface that affects up to 100 million people worldwide. During the pathogenesis of dry eye, impaired function of the lacrimal and meibomian glands results in hyposecretion of aqueous tear fluid, coupled with increased evaporation and instability of the tear film, which becomes increasingly hyperosmolar in character at more severe stages of disease. One critical issue in the field is that the commonly used clinical signs and symptoms for the diagnosis or classification of dry eye often do not correlate with one another. This underscores the need to systematically evaluate current tests and highlights the importance of developing new markers of disease progression for use as endpoints in clinical trials of diagnostic devices and potential therapeutics. In this review, we examine the utility and limitations of commonly used signs and symptoms of dry eye disease and comment on newer biomarkers and analytical devices that show promise for future diagnostic and therapeutic use.

Keywords

Dry eye disease, biomarkers, tear film, hyperosmolarity, symptoms, objective, subjective

Disclosure: This study was supported by TearLab Corporation (San Diego, CA). The authors wish to indicate the following financial interests or in ophthalmology:
- Consultant (MAL, LAC), Equity (MAL, BDS), Patents (BDS), Employee (BDS).

Received: 3 March 2012 Accepted: 16 March 2012 Citation: European Ophthalmic Review, 2012,6(3):157–63 DOI: 10.17925/ EOR.2012.06.03.157

Biomarkers – surrogates or substitutes for variables of clinical disease – are often used in the assessment of disease states, particularly in clinical trials of therapeutics and diagnostic devices.1 The terms (bio)markers, surrogates, endpoints, outcomes and others have been used to describe a metric that is either objective or subjective, which accurately reflects the characteristics of disease. In medicine, common examples of this are the use of serum cholesterol or blood pressure for cardiovascular disease, or in ophthalmology intraocular pressure for glaucoma. A surrogate outcome has been defined as a “laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives”. The value of such a metric lies in how accurately it can capture important aspects of the disease. Such measures are typically used clinically in diagnosis, assessment of disease severity and response to therapy.

A hallmark of DED is the highly variable symptom profile at different stages of dry eye, with signs frequently conflicting with the objective markers used to quantify ocular surface status. In addition, although most patients with DED complain of symptoms of irritation or visual disturbance, not all patients experience overt symptoms. In a recent study, approximately 30% of subjects with clear objective evidence of DED based on a composite severity scale derived from seven commonly used clinical tests for dry eye2 were asymptomatic according to their Ocular Surface Disease Index (OSDI) questionnaire scores.23 As a consequence, symptoms alone are insufficient to facilitate proper management of the disease. Rather, it is necessary for the clinician to take a global view of the clinical signs and symptoms without relying too heavily on any single measure of discomfort.

Over the last decade, concepts concerning DED have shifted and are continuing to be developed and refined as understanding of the
mechanisms of the pathogenesis of the disease increases. The DEWS report advanced the concept that there are various risk factors for the development of the disease.\(^2\) These risk/initiating factors are summarised in Figure 1. Regardless of the initiating factor underlying an individual patient’s development and progression of disease, the final common clinical expression of dry eye is a dysfunctional tear film that is unstable and hyperosmolar in nature.\(^1\) These two defining characteristics of the disease contribute to inflammation of the ocular surface, ultimately leading to evidence of cellular damage.\(^3\) Several pro-inflammatory cascades are activated by tear hyperosmolality and there is an increase in apoptotic cell death in the ocular surface of DED patients.\(^4\) Reflex stimulation of the lacrimal gland is a compensatory mechanism designed to lower tear osmolality and can be operative transiently and asymptomatically between eyes. However, chronic reflex tearing can lead to an overexpression of autoantigens in the lacrimal glands and neurogenic inflammation within the gland, ultimately compromising aqueous tear production.\(^5\) These factors contribute to a vicious cycle of mutually re-enforcing events contributing to increased inflammation, tear film hyperosmolality and instability, and exacerbation of disease severity. As noted in the DEWS Report, increased tear osmolality is thought to be the central pathogenic mechanism leading to damage of the ocular surface in DED.\(^6\)

There are two main mechanistic subtypes of DED: aqueous tear deficient dry eye (ADDE), in which there is a decrease in lacrimal gland secretion that produces aqueous tear fluid and evaporative dry eye (EDE), which is most commonly caused by meibomian gland dysfunction (MGD). This latter subtype has been documented in great detail in the recently published Meibomian Gland Dysfunction Workshop report.\(^7\) The more common of the two subtypes is MGD, but as the disease progresses and increases in severity, compensatory mechanisms (reflex tearing, increased blinking and thickening of the lipid layer) tend to involve the other subtype (ADDE). This ultimately leads to the manifestation of a mixed or hybrid type of DED, the principal type seen in moderate to severe cases.\(^8\)

Diagnosis and management of DED has been a source of frustration to clinicians for a number of reasons; one of which is a lack of correlation between signs and symptoms.\(^9\) While this is undoubtedly attributable, in part, to psychological factors and personality type, recent evidence of sensory changes on the ocular surface have pointed to a hyperalgesia observed in early disease and a decrease in corneal sensation in more severe disease.\(^9\) Additionally, the most commonly used tests in clinical practice frequently do not give clear information as to the status of dry eye, particularly in early or mild stages of disease.\(^9\) A common clinical presentation is that of a patient with significant symptoms of irritation without obvious evidence of disease as measured by current commonly used clinical tests.

This lack of correlation between signs and symptoms has led to problems in several areas such as the accurate diagnosis of early or mild disease and a conspicuous lack of a single metric to assess disease severity and response to treatment. This has been made apparent by the difficulties in identifying a single test as a primary endpoint for clinical trials. This condition is necessary to meet the US Food and Drug Administration’s requirement for an improvement in both a sign and symptom for clearance of new therapeutic agents for dry eye. This is a very high standard that presents a barrier in the process of gaining approval of new therapeutics in dry eye and is a major reason that there have been no new therapeutic drugs approved in the US since 2003, despite many Phase II and III trials. Since the lack of correlation between conventional tests and symptoms is well-known,\(^11\)\(^,\)\(^23\)\(^,\)\(^24\)\(^,\)\(^27\) there is increasing interest within both the FDA and pharmaceutical companies in the development of biomarkers of DED.

New research has begun to elucidate the factors contributing to this problem and may lead the way to improved diagnosis, severity assessment and disease management for the clinician and better endpoints for more effective drug development in clinical trials. The central characteristics of DED, namely tear instability and hyperosmolality, may well hold the keys to unlock this mystery.

### Objective and Subjective Clinical Tests

An example of an ideal biomarker is the measurement of human chorionic gonadotrophin to indicate the presence of an implanted embryo during pregnancy. This protein is a small, highly charged molecule that is essentially absent in the circulation of non-pregnant females and can rise in concentration by almost four orders of magnitude over a sustained period of time. It is as close to a binary indicator as biologically possible, making it very clear to evaluate whether the subject is pregnant or not. Currently there is no single biomarker of such quality for DED. Most importantly, unlike binary conditions, such as pregnancy or presence of an infection, dry eye exists along a continuum of severity.\(^15\) Patients transition from intermittent, mild disease to chronic severe disease, with subtypes defined by the involvement of various pathological mechanisms to differing degrees at each stage. Further, while the traditional markers of corneal staining and Schirmer test may be sufficient for diagnosing severe subjects, they have been shown to be inadequate at quantifying the impact of therapeutic intervention,\(^16\) as evidenced by the numerous failed clinical trials that used these markers as primary endpoints.\(^24\) New data also suggest that punctate staining may have little to do with apparent damage to the cornea and more to do with relative levels of mucin expression.\(^17\) Therefore, both researchers and clinicians must be
Table 1: Various Clinical Biomarkers of Dry Eye Disease Severity

<table>
<thead>
<tr>
<th>Clinical Test of Severity</th>
<th>Units</th>
<th>Parameter Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaporimetry</strong></td>
<td>g/cm²/sec</td>
<td>Evaporation rate</td>
</tr>
<tr>
<td>Schirmer Test</td>
<td>mm</td>
<td>Tear production</td>
</tr>
<tr>
<td>Optical scattering</td>
<td>OSI</td>
<td>Visual stability</td>
</tr>
<tr>
<td>TBUT</td>
<td>seconds</td>
<td>Tear stability</td>
</tr>
<tr>
<td>Interferometry</td>
<td>μm</td>
<td>Lipid layer thickness</td>
</tr>
<tr>
<td>Blink Rate</td>
<td>blinks/minute</td>
<td>Tear stability (indirect)</td>
</tr>
<tr>
<td><strong>Molecular Tests</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tear Osmolarity</td>
<td>mOsm/L</td>
<td>Tear concentration and stability</td>
</tr>
<tr>
<td>MMP-9</td>
<td>μg/mL</td>
<td>Inflammation</td>
</tr>
<tr>
<td>Cytokine profiles (IL-6, IL-8, etc.)</td>
<td>pg/mL</td>
<td>Inflammation</td>
</tr>
<tr>
<td>Lactoferrin/Lysozyme</td>
<td>mg/mL</td>
<td>Tear production</td>
</tr>
<tr>
<td><strong>Subjective Tests</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corneal staining</td>
<td>grade 0–4</td>
<td>Damage (mucin deficiency)</td>
</tr>
<tr>
<td>Conjunctival staining</td>
<td>grade 0–4</td>
<td>Tissue damage</td>
</tr>
<tr>
<td>MGD grading</td>
<td>grade 0–28</td>
<td>Gland obstruction and inflammation</td>
</tr>
<tr>
<td>Symptoms (OSDI)</td>
<td>score 0–100</td>
<td>Discomfort</td>
</tr>
</tbody>
</table>

TBUT = tear breakup time; MGD = meibomian gland dysfunction; OSDI = ocular surface disease index.

Tests of DED severity can be broadly categorised as either objective or subjective, with the objective markers broken down as either functional tests or those that report the molecular constituents of the tear film (Table 1). In an ideal world, the quality of a biomarker would be neatly characterised by a broad dynamic range, high sensitivity to therapeutic intervention, linearity (i.e. probability distribution at different severities), low biological variance over time, and minimal interoperator variance during readout. However, very little referent data exist for the majority of biomarkers associated with disease severity.

Functional Tests

Evaporimetry

Given the increased understanding of dry eye as a consequence of MGD, tear evaporation rate should be directly related to disease severity. However, technical challenges to actually performing this test in a clinical setting, coupled with very large variation in referent values, have precluded its practical implementation. Most of the devices developed to measure tear evaporative rates are of custom design and incorporate a sealed goggle apparatus with inflow and outflow channels. The differences in water content of the inflow air with the outflow are measured. Of the seminal studies in this regard, some investigators found that evaporation rate increased in DED, while others found that the evaporation rate decreased. Specifically, measurement of water evaporation is not the same as the evaporation rate from the tear film – more water may evaporate from a normal subject in a given time than a dry eye subject because in a normal subject there is more fluid available to evaporate, while the overall rate of evaporation is lower than in dry eye subjects. When measuring fluid in a closed chamber, to understand the rate at which fluid is evaporating, one must compensate for the ambient temperature, humidity, ocular physiology, interpalpebral space over time, and other factors. Recent developments have incorporated the use of optical interference or thermography to estimate evaporation in real time. These more recent data suggest that the dry eye features a significantly increased evaporation rate compared with normal subjects.

Schirmer Test

This is a test that aims to measure aqueous tear production, wherein a strip of filter paper is inserted over the inferior lid margin and the length of filter paper that is wetted is measured in millimetres after five minutes. Frequently a topical anaesthetic is instilled prior to measurement to reduce the irritating effects of the strip on the conjunctiva, which can trigger reflex tearing. It has been thought that this might represent ‘basal’ tearing, however, it has been shown that while tearing is lessened by decreasing some of the sensory input, stimuli from the lid margin and lashes are unaffected. Therefore, most clinicians perform this test without anaesthetic. There are several variations on this methodology such as the phenol red test, in which a colour-impregnated thread is substituted for the strip. One limitation of this test is that it is a measure only of aqueous tear production and does not capture metrics pertaining to EDE. Additionally, it is susceptible to wide intra-subject, day-to-day and visit-to-visit variation. More consistent readings are attainable with increasing severity of disease, as the test scoring has a lower boundary limited by a measurement of zero mm. A recent study has demonstrated that in mild/moderate DED, Schirmer test results lack sensitivity and are essentially a ‘random number generator’. This test tends to improve in accuracy in moderate-to-severe disease in which most subjects have a mixed or hybrid form of the disease.

Tear Breakup Time

Tear instability is a hallmark of the dry eye state. It affects vision and contributes to the loss of homeostatic control of the lacrimal functional unit, giving rise to the variability of many clinical tests used to assess this disease. Tear instability is, therefore, of particular interest and accurate measurement of this factor could provide information that reflects a principal axis of the disease. The tear breakup test has been used as a measure of tear instability for decades. Most commonly, it is measured after the instillation of a defined volume of fluorescein, followed by assessment of the rapidity with which the tear film begins to break up after a blink. Values of less than 10 seconds have been recommended as diagnostic thresholds, however, more recently the use of smaller volumes (5 μL) of instilled 2 % fluorescein with a referent value of 5 seconds has been recommended. Unfortunately this test suffers from limitations such as poor reproducibility, low performance accuracy and the fact that other events such as the presence of surface irregularities will influence results.

Attempts to standardise the assessment of tear instability have centred on the development of automated visual capture of serial images using video keratography to eliminate the subjective aspects of tear instability measurement. Several different algorithm-driven programmes have been described. While more rapid inter-blink tear breakup times (TBUTs) have clearly been demonstrated in patients with DED, the range of normative values and relationships over the spectrum of DED severity are yet to be elucidated.

Interferometry

The use of interference ring patterns has been of interest for some years in assessing dry eye. Various devices have been developed to
measure the thickness of the lipid layer of the tear film, and correlations between a thinned lipid layer and dry eye disease have been described.\(^\text{14}\) Most of these devices depend on observer judgment of colour patterns to estimate thickness of the lipid layer. Recently an automated device has been introduced in which pattern recognition is determined by software and correlations between a thin lipid layer and patient symptomatology have been reported.\(^\text{42}\) However, no published studies are available as yet that establish the reproducibility of these devices, correlations between these tests and other measures of DED severity, or response to treatment. This technology, in a more developed form, might prove to be useful in the diagnosis of MGD.

### Optical Quality Tests

Since the tear layer plays an important role in the optical quality of the eye’s visual pathway, alternations in the tear film should be relevant to the pathogenesis of DED. Variations in the optical qualities of the tear film have been reported. These include changes is visual acuity correlated to central corneal surface disruption and degradation of the visual image during the interblink interval in patients with DED. This latter characteristic – a consequence of an unstable tear film – has been dubbed ‘functional visual acuity’, and instruments aiming to capture the degradation of the visual image over time, after a blink, are in development. Other visual parameters affected by an unstable tear film include contrast sensitivity, Hartmann–Shack wavefront sensing and double-pass optical analysis.\(^\text{15,16}\) These measurements may contribute sensitive clinically relevant markers in the coming years.

### Tear Constituent Tests

Changes in the composition of the tear film have been identified in the development of DED. Decreases in certain tear proteins, e.g. tear lysozyme and tear lactoferrin, have long been known to occur and have been utilised as diagnostic tests. Although these have been in use in previous years, there are currently no widely utilised commercial assay systems for these markers. It is not known why these tests are no longer commercially available. Most probably this is due to the technical challenges of performing these tests, the lack of referent values, and a lack of widespread adoption. It should be noted that as their putative relationship to DED is to the aqueous tear component and is likely to respond to anti-inflammatory therapy, this relationship is a consequence of an unstable tear film. MMP-9 correlates so well with staining, it is not clear that the marker provides any additional information than ocular surface staining alone. MMP-9 activation is associated with tissue damage from inflammation, which is a prominent feature of dry eye in moderate to severe cases. Several studies have demonstrated the rise of MMP-9 in DED. In one study measuring tear levels of MMP-9, the protein was elevated across four levels of increasing disease severity.\(^\text{17}\) However, the actual expression levels for grades 1 and 2 were significantly lower than in grades 3 and 4. Of interest, the categorisation of subjects in the study was primarily based on corneal staining. Since other studies have shown that less than 50 % of patients with mild-to-moderate disease have corneal staining (vide infra), it may be that MMP-9 correlates very well with staining, but not necessarily DED. MMP-9 levels in the tears have been shown to be elevated in ocular surface diseases other than DED.\(^\text{17}\) The recent commercialisation of an office test for MMP-9 (Rapid Pathogen Screening [RPS] InflammaDry DetectorTM) which is qualitative may prove useful in identifying patients with a significant inflammatory component and is likely to respond to anti-inflammatory therapy, however, it should also be noted that RPS excludes the difficult to diagnose mild-to-moderate subjects (with OSDI between 1–13) in order to achieve the performance reported in its labelling. Given that MMP-9 correlates so well with staining, it is not clear that the marker gives any additional information than ocular surface staining alone.

In this era of advancing technology, broad studies at the proteomics level are becoming more affordable and feasible to carry out in patient populations. Thus, further research in these directions may identify other protein candidates that represent clinically useful markers.

### Quantitative Tests

#### Ocular Surface Staining

The definition of DED expressed in the DEWS report emphasises damage to the ocular surface.\(^\text{18}\) For many years, damage has been

---

**Biomarkers In Dry Eye Disease**
assessed by the use of vital dyes which increase visibility of areas of discontinuity on the corneal and conjunctival surfaces. Fluorescein solutions are common in current clinical use to assess corneal disruptions, and preceding this was rose bengal staining and more recently lissamine green has been implemented to visualise these areas of the conjunctiva. Ocular surface staining is used to diagnose DED, assess its severity and as a clinical endpoint in clinical trials for pharmaceutical effect.

It should be noted, however, that precisely what damage is revealed by staining methods and whether distinct clinical patterns of staining reflect different changes to the ocular surface, remain undefined. Traditionally, positive rose bengal staining was thought to be due to a breakdown in the structure of the cell wall of conjunctival surface cells, however, experimental studies revealed that discontinuities in the mucin coating of the cells corresponded to the patterns of staining observed. Fluorescein staining of the corneal surface is thought to reflect a breakdown in the cell barrier to aqueous solutions, but the import of differences in staining patterns (punctate versus confluent) is unclear.

Several grading systems for assessing severity of ocular surface staining have been developed. These include an older van Bijsterveld system, the National Eye Institute/Industry Report system and the Oxford system. The latter two are the most commonly used in clinical trials. They differ in the subdivisions of the corneal and conjunctival surfaces and the progression of severity (e.g. the Oxford progression is geometric).

Recent evaluation of corneal and conjunctival staining in DED, as measured using a composite index of severity over the entire range of disease, demonstrated that in mild-to-moderate disease less than 50% of subjects demonstrated any corneal staining. A slightly higher fraction of subjects displayed evidence of conjunctival staining, however, much higher percentages would be expected if staining were a precise indicator of the presence of disease. This exemplifies the problems that can arise if one is using the presence of ocular surface staining for the diagnosis of disease. This demonstrates the danger in using a single one of the currently used metrics for diagnosis. While staining provides an important piece of information, it is not necessarily informative for overall disease assessment. In summary, ocular surface staining remains a useful data point for assessing DED, but alone can be misleading. As with most of the tests, staining measurements become more informative in more advanced stages of disease but lack specificity in mild/moderate disease.

Meibomian Gland Dysfunction Grading

MGD is the most common form of DED and the most common subtype of DED. Its prevalence increases with more severe disease when it is part of a mixed form of the disease. Recognising its presence and assessing its severity is crucial not only for accurately diagnosing disease and assessing severity but in formulation of a treatment plan, which will differ from that for pure ADDE. There is general agreement that the quality of the expressed secretion from the meibomian glands is the most important component of this test. A commonly used scale is that reported by Foulks and Bron. This area is discussed in much greater detail in the recently published Report of the Tear Film and Ocular Surface Society (TFOS) Workshop of Meibomian Gland Dysfunction.

Tear Osmolarity

Hyperosmolar tears are regarded as the central pathogenetic mechanism of dry eye. Hyperosmolarity, regardless of symptomatic status, has been shown to induce inflammation, including an upregulation of IL-1β, TNF-α, MMP-9 and activation of c-Jun N-terminal kinase (JNK), extracellular signal-regulated kinase (ERK) and mitogen-activated protein kinase (MAPK) pathways. Additionally, hyperosmolarity has been shown to activate caspase-3 and cytochrome c pathways that result in spontaneous epithelial cell death. Hyperosmolarity causes decreased intracellular connection, loss of microfilae, cell membrane disruption and increased desquamation. An increase in the concentration of tears also has functional consequences, wherein a hyperosmolar tear film has been directly associated with fluctuating vision. Most importantly, data suggest that an increase in osmolarity directly reduces the ability of the cornea and lid to maintain proper boundary lubrication, thereby increasing friction during each blink and likely manifesting in long term wear and lid epitheliopathies evident in DED.

In the recently published guideline, 2011 American Academy of Ophthalmology Preferred Practice Patterns on Dry Eye Syndrome, it is noted that “tear osmolarity has been shown to be a more sensitive method of diagnosing and grading the severity of dry eye compared to corneal and conjunctival staining, TBUT, Schirmer test, and meibomian gland grading.” Tear osmolarity was also found to be a superior correlate to a composite disease severity index as compared with the other most common tests for DED. In addition, tear osmolarity showed the lowest variability over time and the highest sensitivity to therapeutic intervention as compared with corneal staining, TBUT, Schirmer tests and meibomian gland grading. In effect, tear osmolarity gives a direct insight to the health of the ocular surface while predicting response to therapy before symptoms or the other signs have resolved.

Despite the extensive literature detailing the effects of hyperosmolarity on the ocular surface, many clinicians have not had first hand experience interpreting results, which can lead to confusion and incorrect conclusions surrounding its clinical utility. For diagnosis, the most important difference between osmolarity and other tests is that the unstable tear film and its effects on compensatory mechanisms require that both eyes be tested to make a proper determination. Once that is completed, a positive determination is made if either of the eyes are >308 mOsm/L or the difference between the eyes is ≥8 mOsm/L. Since a normal homeostatic tear film is isotonic with plasma osmolarity (290 mOsm/L), either the elevation of osmolarity or significant differences between eyes are indicators of tear film instability and an unhealthy ocular surface. For instance, a mildly symptomatic subject may exhibit a two-eye reading of 304/295 mOsm/L (OD/OS). While the two-eye maximum is clearly elevated from 290 mOsm/L, it is below the 308 mOsm/L threshold. However, given that the low analytical variance (±4 mOsm/L) makes the number statistically equivalent to 308 mOsm/L, coupled with presence of symptoms and a ≥8 mOsm/L difference between the eyes, these findings suggest a positive diagnosis.

Subjective Tests

Symptoms (Ocular Surface Disease Index)

It is frequently stated that DED is a symptomatic disease. This implies that without assessing symptoms the diagnosis of DED is inaccurate.
While most patients with DED do complain of irritative symptoms, as noted above, not all do. A recent analysis of 299 subjects both with and without DED revealed that only about 70 % of patients with evidence of DED by other clinical tests complained of symptoms as measured by the OSDI scoring. Symptoms alone, although they have been thought to be more representative of the disease as a whole than any individual signs,

Perhaps the most vexing aspect of the use of symptoms in DED is their lack of correlation with signs. A newer way of looking at disease may lead to a better understanding of their relationship.

Relationship Between Individual Tests of Dry Eye Disease

Data presented at the 2012 meeting of the Association for Research in Vision and Ophthalmology meeting revealed that across a general patient population, there was no correlation between subjective and objective tests. Each of the common tests was compared to each of the other signs and symptoms and the average r² for each test was consistently low, as follows: TBUT (0.12), Schirmer test (0.08), corneal (0.15) and conjunctival staining (0.16), meibomian grading (0.10), osmolarity (0.04) and OSDI (0.11). As such, each type of measurement was independent of the other tests, such that there was no predictability based on the presence of any single result. For instance, moderate staining levels did not provide any information as to whether the subject should have a low breakup time or low tear production. This result carries significant implications for clinical practice, but more importantly, for clinical trials. It was not determined whether the lack of correlation was statistically driven (i.e. the variation in each marker over time was great enough to preclude correlation) or whether the information given by each marker was independent because each reveals evidence about distinctly different patho-physiological mechanisms of the disease.

Composite Index of Disease Severity

Given the statistical complexity inherent to the signs and symptoms of DED, there has been little guidance as to how to synthesise multiple data points into a single coherent index of overall disease severity. The traditional Delphi approach reprinted in the DEWS report outlines a 1–4 grading scheme, with one being mild or episodic dry eye and four being severe constant and disabling disease. Unfortunately, the Delphi schema does not supply a quantitative method for dealing with conflicts within individual patients, such as when a patient exhibits low breakup time but no apparent staining or symptoms. Conflicting data from a battery of clinical signs is more the norm than the exception. When coupled with a lack of consensus on any single gold standard test for dry eye, there is a need to quantitatively and objectively determine disease severity for clinical trials.

In creating a composite scale, clinical signs are combined by converting each data point into a common severity basis such as a scale of 0–4, or 0–1 scheme, adding them together to generate a final severity score, then parsing the total into normal, mild/moderate and severe categories. A similar but attenuated method was used in a seminal OSDI paper, as a means to check the performance of the symptom questionnaire. To create their index, the authors mixed Schirmer test, lissamine green staining and a subjective facial expression scale, with cutoffs for scoring determined by a literature review (e.g. Schirmer normal values were >10 mm, mild/moderate values between 6 to 10 mm and severe values <5 mm in length). For each sign, normal results were given one point, mild/moderate two points and severe three points. The summed index was graded as ≤3 normal, 4–7 mild/moderate and then 8–9 as severe. While straightforward, there is a series of problems with this approach. Most importantly, when comparing signs against the total score, there is an inherent risk of ordinate bias against the composite abscissa. Little data is available about the facial expression scale used in the paper, but unsurprisingly, two measures of symptoms (OSDI and facial expressions) were shown to be very well correlated (r=0.69). Especially when a composite is formed from a small number of variables, these biases can be significant. As a result, the OSDI showed a sensitivity of 60 % against the physician rating, but jumped up to 80 % sensitivity against the composite score with the facial expression as part of the abscissa. A second limitation of this approach is that to facilitate simplicity, the categories are discrete. A person with a 6/9 in lissamine green staining is given the same input to the total severity as someone with a 2/9 in staining or a 10 mm Schirmer result. This resulting lack of resolution typically creates spurious false negatives or false positives in sensitivity and specificity analysis, which is unacceptable in regulatory environments.

A more recent attempt to create an objective, unbiased composite index outlined a continuous method for determining DED severity. In this approach, the standard procedure of relying upon clinical expertise to assign thresholds to each of seven quantitative measures was followed by fitting either a linear or exponential function to the clinical breakpoints, depending on strength of fit. These functions were inverted and allowed raw clinical data to be mapped on a common 0–1 scale without discretisation (converting continuous differential equations into discrete ones). The next step, adding the individual dimensions together to form a final severity score, was accomplished by first rotating each dimension (equivalent to weighting each data point) by the amount of independent information the measure contributed to the composite. This independent components analysis, unlike the direct linear addition approach used in previous studies, attempted to equalise correlation risk against the composite abscissa (point on the x-axis). Control data sets (random data with no clinical meaning) showed that the expected correlation was on the order r≈0.1, which indicated the successful removal of bias for each clinical measure, that is, a sign that may provide the same and thus redundant, information as another sign. This was in part due to the fact that there were many more inputs to the composite scale than in earlier attempts, so that any one measurement could not dominate the x-axis. Finally, the bases were geometrically combined to form a single normalised composite index from 0–1, with 0 having no evidence of disease and one the greatest. The resulting continuous index revealed the gross failure of threshold-based classification schemes for clinical trials (e.g. a patient must have N out of M tests above a threshold to qualify as a dry eye patient in the study), wherein basic thresholds failed to properly classify 63 % of the mild/moderate subjects recruited for analysis.

The results of measuring the relationship between the individual tests and the composite severity scale revealed that only one test – tear osmolarity – demonstrated a linear relationship to the composite severity score. This should not be too surprising since tear osmolarity is viewed as the central pathogenic mechanism resulting in damage to the ocular surface in DED. Thus, although it has been shown that osmolarity may be the ideal global biomarker correlating to disease
severity — whether the dry eye is evaporative, aequous deficient, or a hybrid — it does not alter the aetiology of the disease and the other biomarkers remain useful in that role.

Conclusion
The authors have reviewed the major candidates for biomarkers which will accurately reflect disease effects over the entire range of DED severity. There are a number of candidates that demonstrate either good utility for subtypes of DED (i.e. ADDE or EDE) and others are still in early stages of testing and might prove to be of value with further development. Currently, however, only tear osmolarity has been found to serve as a clinically useful biomarker over a wide range of disease severity. The increasing use of biomarkers may enable clinicians to more effectively recognize disease in patients, particularly at mild/moderate or early stages and facilitate the evaluation of new treatment modalities for improved disease management.