Comparison of Two Methods for Composite Score Generation in Dry Eye Syndrome

Craig See,^{1,2} Richard A. Bilonick,^{3,4} William Feuer,² and Anat Galor^{1,2}

¹Miami Veterans Administration Medical Center (VAMC), Miami, Florida

²Bascom Palmer Eye Institute, University of Miami, Miami, Florida

³Department of Ophthalmology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

⁴Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania

Correspondence: Anat Galor, Miami VAMC, 1201 NW 16th Street, Miami, FL 33125; agalor@med.miami.edu.

Submitted: April 2, 2013

Accepted: August 5, 2013

Citation: See C, Bilonick RA, Feuer W, Galor A. Comparison of two methods for composite score generation in dry eye syndrome. *Invest Ophthalmol Vis Sci*. 2013;54:6280–6286. DOI: 10.1167/iovs.13-12150

PURPOSE. To compare two methods of composite score generation in dry eye syndrome (DES).

METHODS. Male patients seen in the Miami Veterans Affairs eye clinic with normal eyelid, corneal, and conjunctival anatomy were recruited to participate in the study. Patients filled out the Dry Eye Questionnaire 5 (DEQ5) and underwent measurement of tear film parameters. DES severity scores were generated by independent component analysis (ICA) and latent class analysis (LCA).

RESULTS. A total of 247 men were included in the study. Mean age was 69 years (SD 9). Using ICA analysis, osmolarity was found to carry the largest weight, followed by eyelid vascularity and meibomian orifice plugging. Conjunctival injection and tear breakup time (TBUT) carried the lowest weights. Using LCA analysis, TBUT was found to be best at discriminating healthy from diseased eyes, followed closely by Schirmer's test. DEQ5, eyelid vascularity, and conjunctival injection were the poorest at discrimination. The adjusted correlation coefficient between the two generated composite scores was 0.63, indicating that the shared variance was less than 40%.

CONCLUSIONS. Both ICA and LCA produced composite scores for dry eye severity, with weak to moderate agreement; however, agreement for the relative importance of single diagnostic tests was poor between the two methods.

Keywords: dry eye syndrome, composite score, latent class analysis, independent components analysis

Dry eye syndrome (DES) is a prevalent condition, with symptoms that negatively impact the ability to work and function.^{1–5} DES is a leading cause of visits to optometry and ophthalmology clinics, and DES medications account for approximately \$1.9 billion in US sales annually.^{6–8} Despite its significant morbidity and cost implications, the diagnosis of DES remains a problem for the clinician and researcher alike. This is largely due to the lack of a gold standard definition for the disease, which may arise from the complexity of the underlying disease process.

The Dry Eye Workshop panel defined DES as “a multifactorial disease of the tears and ocular surface that results in symptoms of discomfort, visual disturbance, and tear film instability with potential damage to the ocular surface.” Classically implicated factors include insufficient tear production (aqueous deficiency), meibomian gland dysfunction (lipid tear deficiency), and anatomical considerations. Several measured parameters describe different facets of disease, and have been found to correlate poorly with symptoms and with each other.^{9,10} To add to the complexity, some parameters provide information on one facet of disease, whereas others provide information on multiple facets. For example, although meibum quality mostly assesses lipid health, tear breakup time (TBUT) is a collective parameter influenced by the status of the aqueous, lipid, and mucous.¹¹

With the multifactorial pathology in mind, Sullivan et al.¹² created a composite score that considered multiple factors in its

generation. The authors used independent component analysis (ICA),¹³ a method commonly used in signal processing, to arrive at such a score. Although there is a definite need for a reliable DES composite score, it is not clear if ICA is the optimal approach to use in its generation. ICA methodology has not been frequently applied to the problem of diagnostic tests and has not been previously applied to problems in ophthalmology. Latent class analysis (LCA), on the other hand, is a technique often applied to imperfect diagnostic tests in the absence of a gold standard. It has been applied in ophthalmology to estimate the sensitivity and specificity of laboratory and clinical tests for trachoma in the absence of a gold standard.¹⁴

The purpose of this study was to create composite DES scores using both ICA and LCA approaches, to compare the results, and to examine the benefits and limitations of each method. This information is important for clinicians and researchers, as it increases knowledge on the potential usefulness of composite scores in the evaluation of dry eye status.

METHODS

Study Population

The Veterans Dry Eye Study is an observational study examining tear function in male US military veterans. Patients were recruited from the Miami Veterans Affairs Medical Center

(VAMC) eye clinic irrespective of their tear function status. Patients were not eligible to participate if they were younger than 50 years; had anterior segment abnormalities, such as a pterygium or corneal edema; wore contact lenses; used any ocular medication, with the exception of artificial tears/topical cyclosporine; had HIV, sarcoidosis, or a collagen vascular disease; or had ocular surgery within the preceding 3 months.

Data Collected

Each participant filled out a validated DES questionnaire (the five-item dry eye questionnaire [DEQ5]).¹⁵ The ocular surface examination consisted of assessment of conjunctival injection (range, 0–4), tear osmolarity (measured once in each eye) (TearLAB, San Diego, CA), TBUT (measured twice in each eye and averaged per eye) (range, 0–15), corneal staining (punctate epithelial erosions, range, 0–5),¹⁶ Schirmer's strips with anesthesia, and morphologic and qualitative eyelid and meibomian gland information. Morphologic information collected included the degree of eyelid vascularity (0, none; 1, mild engorgement; 2, moderate engorgement; 3, severe engorgement)¹⁷ and the presence of inferior eyelid meibomian orifice plugging (0, none; 1, less than one-third lid involvement; 2, between one-third and two-thirds involvement; 3, greater than two-thirds lid involvement). Meibum quality was graded on a scale of 0 to 4 (0, clear; 1, cloudy; 2, granular; 3, toothpaste; 4, no meibum extracted).¹⁸ Data were entered into a standardized database.

Statistical Analysis

We applied both independent components analysis (ICA), which maximizes independence among latent components, and latent class analysis (LCA), which assumes the population is a mixture of unobserved subpopulations (i.e., latent classes, which must be described and identified by the researcher). Severity of cases was assigned by a weighted model (ICA) and a posterior probability of dry eye (LCA). Relative performance of the diagnostic tests was determined by weighting factors (ICA; higher is better) and by constructing receiver operator characteristic (ROC) curve based on the predicted class from the LCA and computing the corresponding areas under the

curve (AUCs). The more severe observed score of the two eyes was used for all diagnostic tests.

Independent Components Analysis

For the ICA method, test results were normalized on the range (0,1), with 0 set to the least and 1 to the most severe value in the sample. Note that this approach differs slightly from the previously reported ICA method, which normalized to the consensus of an expert panel.¹² This was done as there was a concern that our population of men aged 55 and older would have different “normal” values. As there is no age-adjusted nomogram for tear parameters, we elected to normalize values based on the severity relative to the rest of the cohort. To test the effect of transformation on the outcome, we performed a secondary analysis by normalizing the data to its SD.

We performed an ICA analysis using a custom method designed by Sullivan et al.¹² known as Ocular Surface Severity Score (OS³). ICA is a method that transforms multivariate data into components. As the name suggests, ICA attempts to maximize the statistical independence between components. The method assumes that observations arise from a linear mixture of unobserved “source signals” and solves for the matrices that convert between sources and observed data. The matrix that transforms sources into observations, A , is called the *mixing matrix*.

To calculate weights for the OS³ composite, we adapted code from OS³ (Sullivan BD, written communication, 2011) and computed weights using GNU Octave (provided in the public domain at <http://www.gnu.org/software/octave/>). In summary, this method uses a fast infomax algorithm to solve for A .¹⁹ To create the weights, we calculated the mean of each row of A , normalized it, and took the multiplicative inverse of each entry. The resulting vector contained the weight for each diagnostic test. Composite scores were then computed from the sum of squares of parameters with their respective weights.

Latent Class Analysis

Latent class analysis was used to determine whether the population of subjects could be described as a mixture of two

FIGURE 1. Path diagram for latent class (mixture) model illustrating one continuous measurement X_1 and one ordinal measurement X_2 for one class. Each measurement is an indicator for a latent mean. For the ordinal measurement, X_2 it is assumed there is an underlying continuous variable χ_2 that is normally distributed with mean of zero and SD of 1. The observed X_2 depends on the value of the unobserved χ_2 (in a nonlinear fashion, as indicated by the wavy line) in relation to thresholds that determine the probability of each ordinal category.

FIGURE 2. Estimated response probabilities for ordinal measurements and means and SDs for continuous measurements conditioned on latent class. Probabilities and means are shifted to worse indications of disease for the diseased latent class compared with the healthy latent class. The diseased latent class was estimated at approximately 42% of the sampled population. For the osmolarity by symptoms plot, the location of the bivariate means are indicated by a red “1” for the diseased class and a blue “2” for the healthy class.

subpopulations, and if so, whether the subpopulations could be identified as healthy and diseased. To avoid problems with the clustering of observations due to censoring, Schirmer measurements were square root transformed (to reduce positive skewness) and then transformed to six ordinal categories with the highest category for measurements at or above 30 mm. Similarly, TBUT measurements were square root transformed and then transformed to seven ordinal categories, with the highest category for measurements at or above 15 seconds. Unavoidably, some information was lost in order to mitigate the unwanted effects of censoring. Examination of the distributions of DEQ5 and osmolarity revealed them to be approximately symmetric. They were treated as continuous variables and standardized to a mean of zero and SD of 1. The remaining discrete ordinal measurements (conjunctival injection, corneal staining, eyelid vascularity, meibomian orifice plugging, and meibum quality) all tended to be positively skewed.

Figure 1 shows the LCA path diagram for the simplified case of one quantitative continuous measurement and one ordinal measurement. Ordinal measurements were handled using a threshold model that assumes an underlying continuous scale that is normally distributed with a mean of 0 and SD of 1. Within each class, the residual errors among the nine measurements are assumed normally distributed and uncorrelated.

The LCA was implemented as a hierarchical structural equation model using the OpenMx^{20,21} package (Version 1.2.0-1926; OpenMx project, Charlottesville, VA) for the R statistical language and environment software (Version 2.15.1; R Foundation for Statistical Computing, Vienna, Austria).²² The model was estimated 40 times using starting values that were randomly selected from plausible ranges for each model parameter to ensure convergence to the maximum likelihood solution.

Figure 2 shows the resulting two subpopulations subsequently identified as diseased (Class 1) with prior probability of

FIGURE 3. Path diagram for the bi-normal model for determining the ROC curve and the corresponding AUC for ordinal measurements. The observed ordinal measurements denoted by R are assumed to be generated from a latent continuous scale denoted by μ which is assumed to be normally distributed. The mean of μ depends on the latent class. For the healthy class, the mean of the continuous scale μ_H has a fixed mean M_H of zero and a fixed standard deviation σ_H of 1. For the diseased class, the mean of the continuous scale μ_D has mean M_D and a standard deviation of σ_D which must be estimated from the data. The relationship between the latent continuous variable μ and the observed ordinal rating score R is nonlinear and this is indicated by a wavy line. If there are 4 ordinal classes, then there will be 3 thresholds denoted by C and these are the same for both latent classes. The thresholds determine which ordinal class R is observed depending on the value of continuous latent variable μ . The probability of observing a particular rating R will differ between the two latent classes if M_D differs from zero and/or σ_D differs from 1. For a given dry eye parameter, the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) are functions of M_D and σ_D .

0.422 and healthy (Class 2) with prior probability of 0.578. The prior probabilities can be interpreted as the proportion of subjects in each latent class. If a patient was randomly selected from our VA population, we would predict their latent class to be healthy before measuring any of their dry eye parameters, because the prior probability of belonging to the healthy class is greater than 50%. The response probabilities conditioned on latent class for all of the ordinal variables (including Schirmer and TBUT) are shown in Supplementary Table S1, and the means and standard deviations for the bivariate distributions for DEQ5 and osmolarity are shown in Supplementary Table S2.

For the ordinal measurements, the probability shifts to being more severe in what we identified as the diseased class. Similarly, for the continuous measurements, both means shift to more severe for the diseased class. The posterior probability

TABLE 2. ICA Calculating the Importance (Weight) of the Nine Measured Ocular Surface Parameters

Dry Eye Parameter	
Osmolarity	20.86
Eyelid vascularity	10.26
Meibomian orifice plugging	10.11
Symptoms, DEQ5	9.18
Schirmer's test	8.77
Corneal staining	8.41
Meibum quality	8.00
Conjunctival injection	7.77
TBUT	6.25

for being in the diseased class (the probability of being in the diseased class conditional on the observed responses for a patient) was computed using Bayes Theorem, substituting the estimated prior and conditional probabilities. Each patient was assigned a latent class based on his modal posterior probability.

Using the assigned class membership, ROC curves were developed and the AUC estimated for each of the nine measured parameters. This was done as a gauge of the importance of each measured DES parameter and the LCA classification. An AUC of 1.0 for a parameter would indicate that it was individually able to generate each patient's LCA classification as healthy or diseased, while an AUC of 0.5 would indicate that the parameter's association with LCA classification was no better than chance.

To accomplish this for the ordinal measurements, the bi-normal model²³ was implemented as structural equation models using thresholds. The corresponding path diagram is shown in Figure 3. The model assumes that each ordinal response has an underlying continuous scale whose mean differs by class. Each underlying continuous distribution is assumed to be normally distributed. For the healthy class, the mean M_H is fixed to zero and the SD σ_H is fixed to 1. For the diseased class, the mean M_D and the SD σ_D are estimated from the data. The thresholds C_i are constrained to be the same for both classes and must be estimated. Using the bi-normal model, the ROC curve is given by:

$$ROC(t) = \Phi(a + b\Phi^{-1}(t))$$

with

$$a = \frac{M_D - M_H}{\sigma_D} = \frac{M_D}{\sigma_D}$$

and

$$b = \frac{\sigma_H}{\sigma_D} = \frac{1}{\sigma_D}$$

where t is 1 – specificity and Φ is the cumulative unit-standard normal distribution function. The AUC is then given by

TABLE 1. Demographic Characteristics of the Study Population

Demographic		%	<i>n</i>	Demographic	Min	Max	Mean	SD
Race	White	69	170	Age, y	55	95	68.7	8.9
	Black	29	71					
Ethnicity	Other	2	6					
	Hispanic	27	66					
	Non-Hispanic	73	181					

n, number in each group; Min, minimum; Max, maximum.

TABLE 3. AUC for Each Dry Eye Parameter Based on the LCA

Dry Eye Parameter	Bi-Normal Parameters		AUC		
	a	b	Lower Bound	Estimate	Upper Bound
TBUT	−2.213	1.095	0.555	0.932	1.000
Schirmer's test	−1.488	0.677	0.750	0.891	0.982
Corneal staining	0.979	0.621	0.708	0.797	0.863
Meibomian orifice plugging	−0.441	0.609	0.540	0.647	0.759
Meibum quality	0.430	0.692	0.539	0.638	0.713
Osmolarity	NA	NA	0.411	0.572	0.732
Symptoms, DEQ5	NA	NA	0.404	0.565	0.725
Eyelid vascularity	0.067	0.610	0.408	0.523	0.624
Conjunctival injection	0.048	0.295	0.411	0.518	0.624

$$AUC = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right)$$

This approach avoids treating the ordinal categories as if they were quantitative, continuous measurements and avoids the inherent downward bias in the AUC estimates that would otherwise occur.

RESULTS

Study Population

Clinical data from 247 men with complete tear film parameter information were included in this analysis. Demographic information for the patients is summarized in Table 1. The

demographic characteristics were reflective of the Miami VA population, with 69% of patients self-identifying as white and 27% self-identifying as of Hispanic ethnicity.

Relative Importance of DES Parameters to Independent Components Analysis

ICA was used to calculate the importance (weight) of each of the nine measured ocular surface variables (Table 2). Osmolarity was found to carry the largest weight, followed by eyelid vascularity and meibomian orifice plugging. Conjunctival injection and TBUT carried the lowest weights. In a comparative analysis, normalizing the data by SD yielded a narrower range of weights and only moderately correlated composites ($R^2 = 0.397$) when compared with the primary ICA analysis.

Relative Importance of DES Parameters to Latent Class Analysis

The estimated AUCs based on the LCA are shown in Table 3. The confidence intervals for the AUCs are also shown but cannot be used to compare AUCs because the dependency due to being computed on the same patients is not accounted for. The results show that TBUT was best at discriminating healthy from diseased, followed closely by Schirmer's test. DEQ5, eyelid vascularity, and conjunctival injection were the poorest at discrimination, operating only slightly above chance. The LCA ranking for discriminating healthy versus diseased bears little or no resemblance to the ICA results.

Agreement Between Two Methodologies

A comparison was made between the ICA composite scores and the LCA posterior probabilities computed for each subject. The results are graphically displayed in Figure 4. As one can see from the graph, the relationship between the two parameters is not entirely linear. Although the slope coefficients relating LCA scores to ICA scores were highly statistically significant ($P \ll 0.00001$), the R^2 was only 0.399, indicating a weak to moderate linear relationship.

DISCUSSION

This study presents DES composite score generation using two different statistical methods: ICA and LCA. Both ICA and LCA produced composite scores for dry eye severity, with weak to moderate agreement ($R^2 = 0.399$); however, agreement on the importance of single diagnostic tests was poor between the two methods.

FIGURE 4. Regression of the LCA logit score (log odds) on the ICA composite score, and ICA composite score on the LCA logit score. The posterior probabilities P were first transformed to logit scores (i.e., log odds): $\text{logit} = \log(P/(1 - P))$ in order to help linearize the relationship. Probabilities greater than 0.9997 were set to 0.9998 to avoid positive infinite values and probabilities less than 0.000162 were set to 0.000161 to avoid negative infinite values. The logit scores were regressed on the ICA composite scores (CS-ICA) and then the ICA composite scores (CS-ICA) were regressed on the logit scores with the following expectations: $\text{logit} = -15.25 + 35.82 \text{ CS-ICA}$ and $\text{CS-ICA} = 0.4168 + 0.01114 \text{ logit}$. The residual SEs were 3.928 for logit regressed on CS-ICA, and 0.06927 for CS-ICA regressed on logit.

In evaluating the merits of each approach, it is important to recognize that both make specific assumptions. ICA assumes statistical independence of the latent variables and makes the assumption of no random measurement error. Furthermore, the model assumes the data are metric (continuous and quantitative), whereas most DES measurements are discrete and/or ordinal. LCA typically makes the assumption of local independence (i.e., once a certain class is assigned, the variables within that class are assumed statistically independent). However, the effect of the latter assumption was evaluated. The correlations among the measured parameters, to begin with, were low. Assignment to latent classes had little effect on the observed correlations (data not shown). Finally, an important part of using LCA is to identify the nature of the discovered classes; that is, to characterize significant and relevant attributes of each population based on the pattern of differences among the measured parameters for each class.

Considering their history, mathematical complexity, and underlying assumptions, each method has its strengths and weaknesses. A strength of LCA is that it has a more established track record for characterizing diagnostic tests in the absence of a gold standard.¹⁴ For our patients, based on the severity of their responses, it was easy to identify one latent class as diseased and the other as healthy. Although our LCA results do not “prove” there are two subpopulations, the results suggest that this may be a useful description for the observed data. Based on the LCA model, TBUT and Schirmer’s test were superior to the other seven tests. Had TBUT and Schirmer’s test results not been censored, it is likely their superiority would have been even greater. Although there appear to be two distinct classes, the degree of separation is small and the amount of overlap is large. Although for approximately 30 patients, the posterior probability for a given latent class was 1 (and 0 for the opposite class), for most patients the posterior probability ranged between 1 and 0. Compared to other commonly used statistical techniques, LCA methodology is complex and requires substantial mathematical and statistical knowledge to construct a model and interpret the results.

ICA is a method typically used in signal separation and deconvolution. The description by Sullivan et al.¹² appears to be the first and only published use of ICA to create composite scores from diagnostic tests. As such, ICA has never been validated for this application. The main strength of ICA is in the simplicity of implementation. The method occurs in two steps: the transformation of measurements onto (0,1) and the computation of weights via ICA. Once weights have been calculated for a population, the calculation of composites is accomplished through basic arithmetic, a feature that cannot be said for LCA. However, using simulated data and multiple transformations of our DES data, we consistently found that measurements with higher variance produced greater weights (data not shown). This likely arises from a simplification necessary to make the model solvable: fixing S to have a variance of 1. Because $A \times S = X$, and weights are calculated from $1/A$, the weights are affected by the variance of the transformed measurements, X . As such, a weakness of the ICA is the potential for analytic manipulation through alterations in the method of transformation.

Besides creating composites, the ICA and LCA methods produce measures of diagnostic test performance. Unfortunately, the best-performing test by LCA (TBUT) is the worst-performing test by ICA methodology. Likewise, the best-performing test by ICA (osmolarity) is among the worst performing by LCA. Because test performance has not been well characterized for these tests, it is not clear which method is producing the most reliable results. It is important to note that there are other tear film parameters, such as tear meniscus volume and lipid layer appearance and degree of spreading,

which were not included in our model. Although these are not routinely measured metrics, their inclusion may have improved the performance of our model.

To conclude, this study demonstrates that it is possible to generate composite scores for DES using two different methodologies. LCA may be more robust than ICA, given its prior record in determining presence of disease in the absence of a gold standard and its ability to handle ordinal and continuous variables; however, it is more complex to implement. Both methodologies, however, require mathematical manipulation beyond that which is routinely performed by dry eye researchers. Our data suggest that the interpretation of ICA results requires scrutiny of how the data transformation was performed. Based on this study, however, no definitive conclusions can be made regarding which method is superior. Nonetheless, this article highlights that different methodologies generate different composite scores. This needs to be considered before a single method is adopted as the “gold standard” representation of disease severity.

Acknowledgments

This study was supported by a grant from the Veterans Affairs Medical Center (AG) and institutional grants from the National Institutes of Health Center Core Grant P30EY01480, Research to Prevent Blindness Unrestricted Grant, and Department of Defense Grant W81XWH-09-1-0675.

Disclosure: C. See, None; R.A. Bilonick, None; W. Feuer, None; A. Galor, None

References

1. The epidemiology of dry eye disease: report of the Epidemiology Subcommittee of the International Dry Eye WorkShop. *The Ocular Surface*. 2007;5:93-107.
2. Mertzani P, Abetz L, Rajagopalan K, et al. The relative burden of dry eye in patients' lives: comparisons to a U.S. normative sample. *Invest Ophthalmol Vis Sci*. 2005;46:46-50.
3. Miljanovic B, Dana R, Sullivan DA, Schaumberg DA. Impact of dry eye syndrome on vision-related quality of life. *Am J Ophthalmol*. 2007;143:409-415.
4. Schiffman RM, Walt JG, Jacobsen G, Doyle JJ, Lebovics G, Sumner W. Utility assessment among patients with dry eye disease. *Ophthalmology*. 2003;110:1412-1419.
5. Rajagopalan K, Abetz L, Mertzani P, et al. Comparing the discriminative validity of two generic and one disease-specific health-related quality of life measures in a sample of patients with dry eye. *Value Health*. 2005;8:168-174.
6. CBDM.T, The Market and Business Intelligence Company Reviews The Dry Eye Disease Treatment Market. Available at: <http://www.prweb.com/releases/cbdrm/dryeyedisease/prweb3924374.htm>. Accessed July 14, 2013.
7. Petris R. Data dry eye market forecast to 2017. Available at: <http://www.dryeyezone.com/talk/entry.php?291-GlobalData-dry-eye-market-forecast-to-2017>. Accessed April 6, 2012.
8. Yu J, Asche CV, Fairchild CJ. The economic burden of dry eye disease in the United States: a decision tree analysis. *Cornea*. 2011;30:379-387.
9. Schein OD, Tielsch JM, Munoz B, Bandeen-Roche K, West S. Relation between signs and symptoms of dry eye in the elderly. A population-based perspective. *Ophthalmology*. 1997;104:1395-1401.
10. Galor A, Feuer W, Lee DJ, Florez H, Venincasa VD, Perez VL. Ocular surface parameters in older male veterans. *Invest Ophthalmol Vis Sci*. 2013;54:1426-1433.
11. Pflugfelder SC, Tseng SC, Sanabria O, et al. Evaluation of subjective assessments and objective diagnostic tests for

- diagnosing tear-film disorders known to cause ocular irritation. *Cornea*. 1998;17:38–56.
12. Sullivan BD, Whitmer D, Nichols KK, et al. An objective approach to dry eye disease severity. *Invest Ophthalmol Vis Sci*. 2010;51:6125–6130.
 13. Hyvarinen A, Oja E. Independent component analysis: algorithms and applications. *Neural Netw*. 2000;13:411–430.
 14. See CW, Alemayehu W, Melese M, et al. How reliable are tests for trachoma? A latent class approach. *Invest Ophthalmol Vis Sci*. 2011;52:6133–6137.
 15. Chalmers RL, Begley CG, Caffery B. Validation of the 5-Item Dry Eye Questionnaire (DEQ-5): Discrimination across self-assessed severity and aqueous tear deficient dry eye diagnoses. *Cont Lens Anterior Eye* 2010;33:55–60.
 16. Bron AJ, Evans VE, Smith JA. Grading of corneal and conjunctival staining in the context of other dry eye tests. *Cornea*. 2003;22:640–650.
 17. Foulks GN, Bron AJ. Meibomian gland dysfunction: a clinical scheme for description, diagnosis, classification, and grading. *Ocul Surf*. 2003;1:107–126.
 18. Tomlinson A, Bron AJ, Korb DR, et al. The international workshop on meibomian gland dysfunction: report of the diagnosis subcommittee. *Invest Ophthalmol Vis Sci*. 2011;52:2006–2049.
 19. Amari SI, Cichocki A, Yang HH. A new learning algorithm for blind signal separation. *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press; 1996;757–763.
 20. Boker SM, Neale MC, Maes HH, et al. *OpenMx 1.3 User Guide*. Charlottesville, VA: The OpenMx Project; 2012.
 21. Boker SM, Neale MC, Maes HH, et al. OpenMX: an open source extended structural equation modeling framework. *Psychometrika*. 2011;76:306–317.
 22. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>. Accessed November 20, 2012.
 23. Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford, UK: Oxford University; 2003.