Innovative Techniques and Technology

Gary N. Foulks, MD, editor

Statistical Approaches to Assessing Single and Multiple Outcome Measures in Dry Eye Therapy and Diagnosis

Alan Tomlinson, PhD, DSc,[1] Mario Hair, M.Phil,[2] and Angus McFadyen, PhD,[2]

ABSTRACT Dry eye is a multifactorial disease which would require a broad spectrum of test measures in the monitoring of its treatment and diagnosis. However, studies have typically reported improvements in individual measures with treatment. Alternative approaches involve multiple, combined outcomes being assessed by different statistical analyses. In order to assess the effect of various statistical approaches to the use of single and combined test measures in dry eye, this review reanalyzed measures from two previous studies (osmolarity, evaporation, tear turnover rate, and lipid film quality). These analyses assessed the measures as single variables within groups, pre- and post-intervention with a lubricant supplement, by creating combinations of these variables and by validating these combinations with the combined sample of data from all groups of dry eye subjects. The effectiveness of single measures and combinations in diagnosis of dry eye was also considered.

KEY WORDS dry eye, outcome measures, statistical analysis, tear physiology

Accepted for publication May 2013.

From [1]Vision Sciences, Glasgow Caledonian University, Glasgow, United Kingdom and [2]Independent Statistics Consultants, Glasgow.

This work was supported by a research grant from Pfizer Inc (IIR WS887596).

The authors have no commercial or proprietary interest in any concept or product discussed in this article.

Single-copy reprint requests to: Alan Tomlinson, PhD, DSc (Address below).

Corresponding author: Professor Alan Tomlinson, PhD, DSc, Department of Vision Sciences, Glasgow Caledonian University, Cowcaddens Road, Glasgow G4 0BA, UK. Tel: 0141-331-3380. Fax: 0141-331-3387. E-mail address: A.Tomlinson@gcu.ac.uk

I. INTRODUCTION

Studies attempting to define dry eye have varied widely in the test measures employed and in the effectiveness of the diagnosis. Criteria used for definition and classification have been based on symptomatology, clinical test results (Schirmer wetting, tear film breakup time, ocular surface staining) and a range of tear physiology tests that measure distribution, stability, evaporation, etc. The number and variety of tests used in various studies have made it difficult to assess prevalence levels of dry eye and the efficacy of treatment regimens. Descriptions of the assessment techniques and their advantages and shortcomings have been well described in recent literature.[1-7]

The multifactorial etiology of dry eye disease makes it unlikely that a single test can provide a complete assessment of the condition; for example, if the only study outcome is a change in the Schirmer test, this may be inadequate for the assessment of subtypes such as evaporative dry eye or meibomian gland disease. The effectiveness of definition and diagnosis can be improved if multiple, combined outcomes are assessed by different statistical approaches. Previously, we have shown the benefit of discriminant function analysis in the diagnosis of dry eye,[8,9] In recent international workshops re-evaluating dry eye disease, multiple tests have been advocated for diagnosis and evaluation of therapeutic effects.[3,6]

There are significant problems in comparing the effectiveness of single and combined test measures in the published dry eye studies because of differences in study design, test technique, criteria (cut-off value) for diagnosis, and the ever-present problem of selection and spectrum bias. A comprehensive survey by Khanal et al of published studies lists diagnostic sensitivities and specificities varying from 10-99% and 49-100%, respectively.[8] In the present review, data from two previous studies of the tear physiology of dry eye subjects (osmolarity [**OSM**], evaporation [**EVAP**], tear turnover rate [**TTR**], and lipid film quality [**TFI**]) were reanalyzed to *compare directly* single versus multiple outcome measures in the effectiveness of diagnosis and the

OUTLINE

Abbreviations

ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
EVAP	Evaporation rate
HPMC	Hydroxypropyl methylcellulose
KMO	Kaiser-Meyer-Olkin (measure of sampling adequacy)
OSM	Osmolarity
TFI	Lipid film quality
TTR	Tear turnover rate

evaluation of therapeutic benefits of treatment. This approach avoids many of the problems inherent in comparing results from different studies.[5-7]

In this review, we re-evaluated data from two studies of dry eye conducted in our laboratory. In the first study,[10] two groups of subjects (n_1=27, n_2=26) were compared, and in the second study,[11] three groups (n_3=24, n_4=24, n_5=25) were evaluated. In both previous studies, the same four aspects of tear physiology were measured as single variables within groups pre- and post-intervention with a lubricant supplement. The techniques of measurement of these variables, evaporimetry,[12] interferometry,[13] TTR,[14] and tear OSM,[10] have been well described in the literature.

The purpose of the current investigation was to evaluate the effectiveness of single and multiple test measures in determining the effectiveness of diagnosis and treatment of dry eye disease. This was carried out by applying various statistical techniques to existing sets of data (a range of laboratory measures of tear physiology) collected in our laboratory. By this means it was possible to avoid the biases and inconsistencies inherent in comparisons of the results from studies from different clinical sites. The statistical methods used ranged from Analysis of Variance (**ANOVA**) and Analysis of Covariance (**ANCOVA**) models for single measure analyses, to multivariate statistical techniques such as factor analysis and discriminant analysis for multiple measures.

The difference between treatment groups on a single outcome measure is analyzed over time in Section III.A by the commonly employed repeated measures ANOVA models. However, when the groups being compared have significantly different baseline levels, it is more appropriate to use ANCOVA with the baseline measures as the covariate, as illustrated in this section.

When considering multiple outcome measures statistically as in Sections III.B and C, it is common practice to attempt to reduce the dimensionality, ie, number of variables, while still retaining a large proportion of the variation of the data. A commonly used technique, factor analysis, is used in Sections III.B.1 and 2 and III.C.2 and 3 to create factors (linear combinations of the original variables which represent the original variables). These factors or indices are usually independent of one another (orthogonal). Once a number of indices have been developed, the difference between treatment groups on each index over time is assessed in Sections III.B.4 and III.C.5, using ANOVA and ANCOVA.

It is also possible to assess the effectiveness of these indices for the initial diagnosis of dry eye. Multivariate discriminant analysis and logistic regression are commonly used in such clinical scenarios when the predicted outcome is dichotomous (dry eye or not dry eye). However, when a single independent index as an outcome variable is analyzed, the clinical interpretation is simpler with use of discriminant analysis. This is illustrated in Section III.C.6.

II. STATISTICAL METHODOLOGY

The present study reassessed the effects of artificial tears on aspects of tear physiology (EVAP, TTR, OSM and TFI) by adopting a range of approaches to the statistical analyses of the effect of treatment for dry eye. The physiology measures were assessed first as single variables, then by creating combinations of these variables, and finally by repeating with the combined sample of data from five groups of dry