Field-wide meta-analyses of observational associations can map selective availability of risk factors and the impact of model specifications

Stylianos Serghiou a, Chirag J. Patel b, Yan Yu Tan a, Peter Koay c,d, John P.A. Ioannidis e,f,g,h,*

a College of Medicine and Veterinary Medicine, The University of Edinburgh, 47 Little France Crescent, Edinburgh EH16 4TJ, Edinburgh, UK
b Department of Biomedical Informatics, Harvard Medical School, 10 Shattuck Street, 4th Floor, Boston, MA 02115, USA
c Ophthalmology Department, St John’s Hospital, Howden South Road, Livingston, West Lothian, EH54 6PP, UK
d The Princess Alexandra Eye Pavilion, Chalmers Street, Edinburgh EH3 9HA, UK
e Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, 1265 Welch Rd, MSOB X306, Stanford, CA 94305, USA
f Department of Health Research and Policy, Stanford University School of Medicine, 150 Governor’s Lane, Stanford, CA 94305, USA
g Department of Statistics, Stanford University School of Humanities and Sciences, 390 Serra Mall, Stanford, CA 94305, USA
h Meta-Research Innovation Center at Stanford (METRICS), Stanford School of Medicine, 1070 Arastradero Road, Palo Alto, CA 94304, USA

Accepted 7 September 2015; Published online 28 September 2015

Abstract

Objectives: Instead of evaluating one risk factor at a time, we illustrate the utility of “field-wide meta-analyses” in considering all available data on all putative risk factors of a disease simultaneously.

Study Design and Setting: We identified studies on putative risk factors of pterygium (surfer’s eye) in PubMed, EMBASE, and Web of Science. We mapped which factors were considered, reported, and adjusted for in each study. For each putative risk factor, four meta-analyses were done using univariate only, multivariate only, preferentially univariate, or preferentially multivariate estimates.

Results: A total of 2052 records were screened to identify 60 eligible studies reporting on 65 putative risk factors. Only 4 of 60 studies reported both multivariate and univariate regression analyses. None of the 32 studies using multivariate analysis adjusted for the same set of risk factors. Effect sizes from different types of regression analyses led to significantly different summary effect sizes (P-value < 0.001). Observed heterogeneity was very high for both multivariate (median I², 76.1%) and univariate (median I², 85.8%) estimates. No single study investigated all 11 risk factors that were statistically significant in at least one of our meta-analyses.

Conclusion: Field-wide meta-analyses can map availability of risk factors and trends in modeling, adjustments and reporting, as well as the impact of differences in model specification. © 2016 Elsevier Inc. All rights reserved.

Keywords: Meta-analysis; Exposome-wide association study; Observational study; Risk factor epidemiology; Statistical modeling; Big data

1. Introduction

Meta-analyses of observational studies on putative risk factors [1–3] are popular, and thousands of such meta-analyses have been published. Although useful in combining large numbers of relevant studies into a summary effect size, these quantitative syntheses suffer from significant limitations [4–6].

Most meta-analysis publications focus on studying the association of one or at most a few putative risk factors to a particular outcome. However, given the massive volume of published studies on risk factor epidemiology, there is typically a large number of risk factors that have been studied for any given outcome. Different investigators may have evaluated different risk factors and may have only reported subsets of them in their published articles. Moreover, there is often lack of consensus about whether risk factors should be analyzed and reported in univariate or multivariate models or both [7], and sometimes this can have a substantial effect on the results [5,8]. When multivariable models are used, there is a large diversity in terms of what other variables should be included for standard adjustments [9–11]. The pool of such variables can be very large, and it depends on what has been

Funding: This work was supported by a NIEHS K99/R00 (K99ES023504) and a PhRMA Foundation grant to C.J.P. The work of Dr. Ioannidis is also supported by an unrestricted gift from Sue and Bob O’Donnell.

Conflict of interest: None.

* Corresponding author. Tel.: +1 (650) 725-5465.
E-mail address: jioannid@stanford.edu (J.P.A. Ioannidis).

http://dx.doi.org/10.1016/j.jclinepi.2015.09.004
0895-4356/© 2016 Elsevier Inc. All rights reserved.
measured and what the investigators of each study consider essential for adjustments.

Finally, the reporting of an observational study may select only a few of the many models that have been assessed, and the choice may be based not only on a priori considerations about which model is the best, but also on the results [12]. Selective reporting bias can be a major threat to the validity of the available published information [13]. Meta-analyses that try to synthesize the selectively reported results, despite the advent of post hoc bias correction [14], may simply reinforce these biases by averaging the impact of diverse biases on the strength of the assessed associations [7,15].

Here, we propose that instead of merely assessing one putative risk factor at a time and synthesizing a limited set of potentially biased results, a meta-analytic approach to such observational data should assess the entire field of putative risk factors—we call this approach a field-wide meta-analysis.

What is new?

Key findings

- Field-wide meta-analyses can map the selective availability of risk factors as well as the patterns of modeling, adjustments and reporting, of risk factors across studies.

What this adds to what was known?

- Across all 60 studies on the risk factor epidemiology of pterygium, 32 used multivariate analyses, but no two studies used identical variables in their multivariate models.
- Differences in model specification can confer differences in summary effect sizes.

What is the implication and what should change now?

- We propose that instead of assessing one putative risk factor at a time, a meta-analysis of observational data should assess the entire field of putative risk factors—what we call this approach a field-wide meta-analysis.

2. Materials and methods

2.1. Systematic review

For a detailed overview of Methods please refer to Appendix A at www.jclinepi.com. Briefly, all studies quantitatively exploring the association of any putative risk factor to the presence of pterygium and providing (or allowing the calculation of) the effect size and 95% confidence interval (CI) were eligible. We systematically reviewed EMBASE Classic and EMBASE (1947 to October 15, 2013), ISI Web of Science (1900 to October 15, 2013), PubMed (1950 to October 15, 2013) and Cochrane Central Register of Controlled Trials (up to October 15, 2013). We customized our search strategy to every database and selected for studies where the main outcome measure was the presence of pterygium. Data were collected by two reviewers (S.S. and Y.Y.T.), who also extracted the following data items: type of modeling (multivariate, univariate or both), adjustment variables (in multivariate models), high-level risk factor domain (e.g., sunlight exposure), more granular risk factor specification (e.g., sunlight exposure for > 5 hours/day vs. not), metric type (e.g., odds ratio [OR] and relative risk [RR]), metric value (e.g., OR = 0.78), and 95% CI or standard error (whichever was available). Whenever the same study population was used more than once to study the same risk factors, we only extracted data from the latest article reporting on that study population. In studies where two multivariate models were used (e.g., a small model adjusting for age and gender and a large model adjusting for all statistically significant factors), we extracted the multivariate data from the largest model for which these were available. Metrics specific to continuous variables (e.g., Cohen’s d) were transformed into ORs using an online tool [18] and standard formulae [19]. Only one study reported RR [20]; given that the incidence of the outcome of interest (pterygium) in that study was less than 10%, the RRs were considered to adequately approximate ORs [21]. On the basis of what is known regarding the pathogenesis of pterygium [22,23], we denoted a priori interest in the following 10 risk factors: sunlight exposure, occupation type, educational attainment, income, area of residence, use of sunglasses, use of spectacles, use of hat, dry eyes, and latitude. Our complete database can be found here: https://goo.gl/9Ti1oG.

2.2. Mapping of risk factors, type of analysis, and adjustments

R [24] was used to format and analyze our database. The figures illustrating the results of this analysis were color
coded using the “Conditional Formatting” function of Microsoft Excel (Redmond, Washington; USA). As noted above, all recorded risk factors were categorized into high-level domains and more granular risk factor specifications (thus, each domain includes several risk factor specifications). For convenience, whenever we use the term “risk factors,” we refer to risk factor domains.

2.3. Meta-analyses for four modeling scenarios

Standard errors were calculated using: ln(upper limit of CI/lower limit of CI)/(2 × 1.96). We meta-analyzed only domains/specifications for which three or more estimates were available. All estimates were synthesized using a random-effects model due to evident between-study heterogeneity. The summary OR was estimated using the random-effects model due to evident between-study heterogeneity. The summary OR was estimated using the restricted maximum-likelihood ratio method. Four meta-analyses were done, one for each of: (1) multivariate data, (2) univariate data, (3) multivariate data combined with univariate data, whenever multivariate data were unavailable (henceforth called preferentially multivariate), and (4) univariate data combined with multivariate data, whenever univariate data were available (henceforth called preferentially univariate). We decided a priori to use \( \alpha = 0.01 \) for claiming nominal statistical significance. Given the multiple meta-analyses performed, \( \alpha = 0.05 \) would be too lenient, whereas a Bonferroni correction (\( \alpha = 0.05/76 = 0.0006 \)) would be too stringent, given that many factors are correlated.

The nonparametric Friedman test was used to identify whether any of the four matched pairs of modeling differs from at least one of the rest. Post hoc multiple comparison after Nemenyi was then used to delineate which specific groups were significantly different from each other, as per the R package “PMCMR” [25]. Where both domains and risk factor specifications were eligible, only domains were included to avoid bias. Meta-analyses were done using the “metafor” 1.9-1 package [26].

2.4. Heterogeneity metrics

Between-study heterogeneity was assessed using the Q statistic and the \( I^2 \) metric, which we calculated using the “metafor” package.

2.5. Excess significance testing

Publication bias, outcome bias, and selective analyses may lead to an excess of statistically significant results, for which we tested by applying the excess significance test to factors reported by at least three studies [16]. Briefly, for every meta-analyzed risk factor, we compare the amount of observed significant results (O) at \( \alpha = 0.05 \) to the amount of expected significant results (E), where \( E = \sum \text{power of each study within a specific meta-analysis}. \) Power was calculated using simulations on R, taking as plausible effect for the risk factor the effect seen in the most precise study (lowest standard error). The difference between O and E was assessed using the binomial test, with \( \alpha = 0.1 \), as previously suggested [17].

3. Results

3.1. Literature search and study description

We initially identified 2,052 records, from which 1,992 were excluded (Fig. 1) leaving 60 eligible studies (Table 1; Appendix B at www.jclinepi.com). No new studies were imported through reference checking. In two of the eligible articles, not all the mentioned risk factors were analyzable because not enough information was available to obtain ORs [27,28].

There were 32 cross-sectional, 22 case-control, 3 prospective cohort, and 3 retrospective cohort studies. Twenty-one of 32 cross-sectional, 8 of 22 case-control, and 3 of 3 prospective cohort studies used multivariate analyses. The following population types were used, with the proportion of those reporting multivariate analyses in parentheses: general (18 of 24), hospital based (3 of 12), work specific (0 of 10), rural (6 of 7), and ethnically specified (5 of 7); the differences reflect the tendency of earlier studies (before 1990) to use case-control studies to investigate specific occupations using univariate analysis. Most studies emanated from Asia (34 of 60), followed by America (10 of 60), Oceania (9 of 60), Africa (5 of 60), and Europe (3 of 60)—one study took place both in America and Asia [29]. The most commonly studied age groups were those aged 40 years old or older (23 of 60). Eligible studies involved 362,548 participants, including 20,094 with at least unilateral pterygium (excluding a study not providing adequate sample size information [30]).

3.2. Mapping of risk factors

A total of 65 risk factors (Fig. 2A) have been studied within 183 different risk factor specifications (Fig. 2B; Supplementary Fig. 1 at www.jclinepi.com). Each risk factor was only studied in a median of one article (interquartile range [IQR], 1–4), and the 10 risk factors of a priori-specified interest were studied in a median of eight articles (IQR, 6–13). Although each of our a priori-determined risk factors have been reported at least five times, only 6 of 60 studies investigated at least half of these 10 factors and the maximum number of these risk factors investigated was 6 of 10.

A total of 25 risk factors were studied at least 3 times, and 17 were studied at least 5 times. Age, occupation type, level of education, sunlight exposure, dry eye, and ethnicity were studied in at least five different risk factor specifications (range, 5–34).
3.3. Mapping of univariate or multivariate modeling and their adjustments

A total of 32 of 60 studies reported a multivariate model to adjust for potential confounders; the rest (28 of 60) only used univariate analysis. Of the studies using multivariate modeling, 5 studies only reported multivariate data; 17 studies reported only multivariate data for some risk factors, whereas for others they either reported both multivariate and univariate data, only univariate data, or raw data from which univariate data could be extracted; 5 studies reported only univariate data for some risk factors but both univariate and multivariate data for the rest; and 1 only reported univariate data, merely indicating that multivariate data were not significant. Seven of 32 studies did not report the multivariate data for risk factors that were found to be significant in univariate but not significant in multivariate analysis. Only 4 of 60 studies reported both univariate and multivariate data for all risk factors that were reported. The 32 studies using multivariate analysis adjusted for a total of 38 risk factors. Among these studies, we did not find a single pair considering exactly the same variables (Fig. 3). However, almost all these studies adjusted for gender (30 of 32) and age (29 of 32). Ten of 32 studies (31.3%) did not adjust for outdoor occupation or sunlight exposure.

3.4. Meta-analyses for four modeling scenarios

A meta-analysis was done for all four types of modeling for 19 risk factors (Fig. 4; Supplementary Table 1 at www.jclinepi.com) and 10 risk factor specifications. Ten factors were significantly associated with pterygium in multivariate meta-analysis (old age, male gender, high sunlight exposure, outdoor occupation, rural area of residence, low level of education, low income, latitude of residence > 30°, use
<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Journal</th>
<th>Design</th>
<th>Nn</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li and Cui</td>
<td>2013</td>
<td>Cornea</td>
<td>Prospective cohort</td>
<td>5,057</td>
<td>323</td>
</tr>
<tr>
<td>Liu et al.</td>
<td>2013</td>
<td>Intern Eye Sci</td>
<td>Cross-sectional</td>
<td>3,067</td>
<td>280</td>
</tr>
<tr>
<td>Marmamula et al.</td>
<td>2013</td>
<td>Invest Ophthalm Vis Sci</td>
<td>Cross-sectional</td>
<td>5,586</td>
<td>655</td>
</tr>
<tr>
<td>Rim et al.</td>
<td>2013</td>
<td>Cornea</td>
<td>Cross-sectional</td>
<td>14,920</td>
<td>4,307</td>
</tr>
<tr>
<td>Sherwin et al.</td>
<td>2013</td>
<td>Acta Ophthalmol</td>
<td>Cross-sectional</td>
<td>641</td>
<td>70</td>
</tr>
<tr>
<td>Sun et al.</td>
<td>2013</td>
<td>Ophthalmic Epidemiol</td>
<td>Cross-sectional</td>
<td>6,685</td>
<td>401</td>
</tr>
<tr>
<td>Tano et al.</td>
<td>2013</td>
<td>Acta Ophthalmol</td>
<td>Cross-sectional</td>
<td>2,312</td>
<td>101</td>
</tr>
<tr>
<td>Zhao et al.</td>
<td>2013</td>
<td>Invest Ophthalm Vis Sci</td>
<td>Prospective cohort</td>
<td>2,628</td>
<td>129</td>
</tr>
<tr>
<td>Ang et al.</td>
<td>2012</td>
<td>Ophthalmology</td>
<td>Cross-sectional</td>
<td>8,906</td>
<td>900</td>
</tr>
<tr>
<td>Anguria et al.</td>
<td>2012</td>
<td>S Afr Med J</td>
<td>Case-control</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>Asokan et al.</td>
<td>2012</td>
<td>Ophthalmic Physiol Opt</td>
<td>Cross-sectional</td>
<td>3,924</td>
<td>740</td>
</tr>
<tr>
<td>Rezvan et al.</td>
<td>2012</td>
<td>Acta Med Iran</td>
<td>Cross-sectional</td>
<td>5,190</td>
<td>489</td>
</tr>
<tr>
<td>Sudhal kar</td>
<td>2012</td>
<td>Eye (lond)</td>
<td>Case-control</td>
<td>107</td>
<td>310</td>
</tr>
<tr>
<td>Zhong et al.</td>
<td>2012</td>
<td>Invest Ophthalm Vis Sci</td>
<td>Cross-sectional</td>
<td>2,133</td>
<td>832</td>
</tr>
<tr>
<td>Vito et al.</td>
<td>2012</td>
<td>Eye (lond)</td>
<td>Cross-sectional</td>
<td>619</td>
<td>42</td>
</tr>
<tr>
<td>Peiretti et al.</td>
<td>2010</td>
<td>Invest Ophthalm Vis Sci</td>
<td>Case-control</td>
<td>112</td>
<td>123</td>
</tr>
<tr>
<td>Fotouhi et al.</td>
<td>2009</td>
<td>Eye (lond)</td>
<td>Cross-sectional</td>
<td>4,565</td>
<td>59</td>
</tr>
<tr>
<td>Shirima et al.</td>
<td>2009</td>
<td>Am J Ophthalmal</td>
<td>Cross-sectional</td>
<td>3,747</td>
<td>1,154</td>
</tr>
<tr>
<td>West and Munoz</td>
<td>2009</td>
<td>Br J Ophthalmol</td>
<td>Cross-sectional</td>
<td>4,774</td>
<td>722</td>
</tr>
<tr>
<td>Demirok et al.</td>
<td>2008</td>
<td>Ann Ophthalmal (skokie)</td>
<td>Case-control</td>
<td>60</td>
<td>102</td>
</tr>
<tr>
<td>Lin et al.</td>
<td>2008</td>
<td>Environ Health Perspect</td>
<td>Case-control</td>
<td>383</td>
<td>134</td>
</tr>
<tr>
<td>Nemesure et al.</td>
<td>2008</td>
<td>Ophthalmology</td>
<td>Prospective cohort</td>
<td>1,888</td>
<td>219</td>
</tr>
<tr>
<td>Omoti et al.</td>
<td>2008</td>
<td>Eye (lond)</td>
<td>Case-control</td>
<td>154</td>
<td>174</td>
</tr>
<tr>
<td>Davies et al.</td>
<td>2007</td>
<td>Niger J Physiol Sci</td>
<td>Case-control</td>
<td>110</td>
<td>85</td>
</tr>
<tr>
<td>Lu et al.</td>
<td>2007</td>
<td>Clin Exp Ophthalmal</td>
<td>Cross-sectional</td>
<td>2,229</td>
<td>323</td>
</tr>
<tr>
<td>Uptonmwan et al.</td>
<td>2007</td>
<td>East Afr Med J</td>
<td>Case-control</td>
<td>123</td>
<td>162</td>
</tr>
<tr>
<td>Paula et al.</td>
<td>2006</td>
<td>Eye (lond)</td>
<td>Case-control</td>
<td>624</td>
<td>115</td>
</tr>
<tr>
<td>Tan et al.</td>
<td>2006</td>
<td>Eye (lond)</td>
<td>Cross-sectional</td>
<td>477</td>
<td>81</td>
</tr>
<tr>
<td>Taylor et al.</td>
<td>2006</td>
<td>Arch Environ Occup Health</td>
<td>Cross-sectional</td>
<td>304</td>
<td>68</td>
</tr>
<tr>
<td>Yan et al.</td>
<td>2006</td>
<td>Chinese Med J-peking</td>
<td>Case-control</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Balachandra</td>
<td>2005</td>
<td>Pac Health Dialogue</td>
<td>Cross-sectional</td>
<td>753</td>
<td>109</td>
</tr>
<tr>
<td>Mathur et al.</td>
<td>2005</td>
<td>Int Ophthalmal</td>
<td>Cross-sectional</td>
<td>1,169</td>
<td>148</td>
</tr>
<tr>
<td>Hayashida and Kitaoka</td>
<td>2003</td>
<td>Acta Med Nagasaki</td>
<td>Cross-sectional</td>
<td>1,558</td>
<td>164</td>
</tr>
<tr>
<td>Gazzard et al.</td>
<td>2002</td>
<td>Br J Ophthalmal</td>
<td>Cross-sectional</td>
<td>1,210</td>
<td>113</td>
</tr>
<tr>
<td>Twelker and Bailey</td>
<td>2002</td>
<td>Invest Ophthalm Vis Sci</td>
<td>Case-control</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Wong et al.</td>
<td>2001</td>
<td>Am J Ophthalmal</td>
<td>Cross-sectional</td>
<td>1,232</td>
<td>120</td>
</tr>
<tr>
<td>McCarty et al.</td>
<td>2000</td>
<td>Br J Ophthalmol</td>
<td>Cross-sectional</td>
<td>5,147</td>
<td>141</td>
</tr>
<tr>
<td>Saw et al.</td>
<td>2000</td>
<td>Acta Ophthalmol Scand</td>
<td>Case-control</td>
<td>125</td>
<td>61</td>
</tr>
<tr>
<td>Threlfall and English</td>
<td>1999</td>
<td>Am J Ophthalmal</td>
<td>Case-control</td>
<td>135</td>
<td>150</td>
</tr>
<tr>
<td>Kadayificilar et al.</td>
<td>1998</td>
<td>Acta Ophthalmol Scand</td>
<td>Case-control</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Khoo et al.</td>
<td>1998</td>
<td>Int Ophthalmal</td>
<td>Case-control</td>
<td>125</td>
<td>61</td>
</tr>
<tr>
<td>Panchapakesan et al.</td>
<td>1998</td>
<td>Aust N Z J Ophthalmal</td>
<td>Cross-sectional</td>
<td>3,629</td>
<td>266</td>
</tr>
<tr>
<td>Corzo-Alvarez and</td>
<td>1997</td>
<td>Invest Clin</td>
<td>Case-control</td>
<td>143</td>
<td>143</td>
</tr>
<tr>
<td>Urribarri-Delgado</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mackenzie et al.</td>
<td>1992</td>
<td>Ophthalmology</td>
<td>Case-control</td>
<td>278</td>
<td>278</td>
</tr>
<tr>
<td>Taylor et al.</td>
<td>1992</td>
<td>Arch Ophthalmal</td>
<td>Cross-sectional</td>
<td>838</td>
<td>140</td>
</tr>
<tr>
<td>Ibechukwu</td>
<td>1990</td>
<td>East Afr Med J</td>
<td>Case-control</td>
<td>1</td>
<td>98</td>
</tr>
<tr>
<td>Taylor et al.</td>
<td>1989</td>
<td>Arch Ophthalmal</td>
<td>Cross-sectional</td>
<td>838</td>
<td>140</td>
</tr>
<tr>
<td>Spierer et al.</td>
<td>1985</td>
<td>Metab Pediatr Syst Ophthalmal</td>
<td>Retrospective cohort</td>
<td>600</td>
<td>93</td>
</tr>
<tr>
<td>Karai and Horiguchi</td>
<td>1984</td>
<td>Br J Ophthalmal</td>
<td>Case-control</td>
<td>191</td>
<td>214</td>
</tr>
<tr>
<td>Moran and Hollows</td>
<td>1984</td>
<td>Br J Ophthalmal</td>
<td>Cross-sectional</td>
<td>64,314</td>
<td>2,194</td>
</tr>
<tr>
<td>Taylor</td>
<td>1980</td>
<td>Br J Ophthalmal</td>
<td>Cross-sectional</td>
<td>350</td>
<td>67</td>
</tr>
<tr>
<td>Biedner et al.</td>
<td>1979</td>
<td>Ann Ophthalmal</td>
<td>Case-control</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>Deleis and Dhir</td>
<td>1967</td>
<td>Arch Ophthalmal</td>
<td>Case-control</td>
<td>1,252</td>
<td>151</td>
</tr>
<tr>
<td>Darrell and Bachrach</td>
<td>1963</td>
<td>Arch Ophthalmal</td>
<td>Case-control</td>
<td>1,904</td>
<td>463</td>
</tr>
<tr>
<td>Kerkenzov</td>
<td>1956</td>
<td>Trans Ophthalmal Soc Aust</td>
<td>Cross-sectional</td>
<td>3,000</td>
<td>287</td>
</tr>
<tr>
<td>Ten Doeschate and Fischer</td>
<td>1949</td>
<td>Ophthalmologica</td>
<td>Retrospective cohort</td>
<td>180,000</td>
<td>147</td>
</tr>
</tbody>
</table>
Fig. 2. (A) A “data microarray” illustrating what studies investigated which putative risk factors, for all risk factors studied three or more times. Risk factors are ordered based on how many times they were studied. Risk factors in light gray are the a priori factors of interest. Each row represents a separate article in the order these were presented in Table 1. (B) A “data microarray” illustrating the risk factor specifications for the “sunlight” risk factor domain.
of spectacles, and use of sunglasses). Use of sunglasses lost
its significance in all other modes of analysis, male gender
lost significance in univariate analysis, and dry eyes gained
significance in univariate analysis. No single study investig-
ated all 11 risk factors found to be significant in at least
one meta-analysis.

Considering risk factor domains, the summary estimate
for different types of analysis was statistically significantly
different ($F_{(27.1; P-value < 0.0001)}$ because of differ-
ences between multivariate vs. univariate and preferentially
multivariate vs. univariate analyses ($P < 0.0005$ for both).
Univariate analyses displayed a tendency toward higher ef-
ect sizes; specifically, in 13 of 19 cases, the univariate anal-
ysis yielded larger effect sizes than the multivariate analysis.
(When taking risk factor specifications into account, this
ratio became 19 of 28—for one of the risk factors, the OR
for multivariate was equal to that for univariate analysis).

3.5. Heterogeneity metrics and excess significance

Heterogeneity between all summary effect sizes was
consistently very high, with median heterogeneity
$I^2 = 76.1\%$ (IQR, 59.9–88.0%) for multivariate/preferentially
multivariate meta-analyses, against $I^2 = 81.2\%$ (IQR, 80.8–
94.3%) for univariate/preferentially univariate meta-
analyses; excluding risk factor domains led to a small nonsignificant
decrease. In 18 of 29 risk factors, heterogeneity was larger
for univariate than multivariate models (binomial
$P$-value = 0.18). The O exceeded the E number of studies with
statistically significant results at $\alpha = 0.1$ in 21 of 116 meta-
analyses.

4. Discussion

We have presented and implemented the concept of field-
wide meta-analyses of risk factors using data from pterygium
risk factor epidemiology. Our example illustrates what is
likely to be common in many fields where efforts are made
to identify risk factors for a given outcome of interest: a large
number of studies evaluate a large number of risk factors,
most of which are reported in the minority of studies. It is un-
clear if this means that these factors were not studied or stud-
ied but not reported. Moreover, there is a large diversity in
modeling and adjustments—in our case study of pterygium,
not a single multivariate model specification was identical
between at least two studies. Different combinations of data
emanating from either univariate or multivariate analyses
yield different meta-analytical effect size estimates. How-
ever, differences in combinations did not seem to affect

---

**Fig. 3.** A “data microarray” illustrating what putative risk factors were
adjusted for by each article. Only a priori-determined risk factors of
interest are shown for convenience. Articles are ordered in descend-
ing order based on how many putative risk factors they adjusted for
(this order only holds true for the risk factors shown). No two
articles adjusted for the same putative risk factors. Refer to
Supplementary Fig. 2 at www.jclinepi.com for a data microarray
illustrating all risk factor domains for which adjustments were made.
Black, adjusted for; gray, not adjusted for.

It shows that although many domains have been investigated extensively, they have been specified very differently across studies. Refer to Supplementary Fig. 1 at www.jclinepi.com for the data microarray illustrating all specifications for all risk factor domains. Light gray risk factor, a priori factor of interest; light gray, not analyzed; medium gray, univariate analysis; dark gray, multivariate analysis; black, both multivariate and univariate analysis; AMD, age-
related macular degeneration; UVA, ultraviolet A; UVB, ultraviolet B.
whether our summary estimates reach statistical significance or not. Large heterogeneity, small-study effects, and excess significance signals are suggestive of reporting biases, although they probably have modest specificity in that regard. Such examples of between-study differences probably are the norm rather than the exception and have been previously shown to affect meta-analyses [7,31,32].

The observed variability in model specification in observational epidemiology may be related to a combination of different reasons and motives. Researchers perhaps do not actively try to align their research to previously published models, and editors/referees do not actively promote this practice. Perhaps, they do not see a need in doing so or even explicitly try to do things differently, using a different model specification and consideration of different risk factors as a means to claim innovation [33]. Furthermore, there are often many valid ways to define a putative risk factor, unless there is a generally accepted consensus. Even when there is consensus, reporting of models may still be sketchy and selective, despite efforts for standardization of reporting [34,35]. Eventually, there is substantial room for flexible analyses and for bias because of conflict of interest or prior beliefs that may result in selective reporting and partial interpretation of the results.

The phenomenon whereby effect size estimates change with alternative data approaches has been called “vibration of effects” [36,37]. Vibration can be generated from alternative choices, not only in adjustments for covariates in model specification [37, 38], but also from different definitions of variables, eligibility criteria, and statistical methods of analysis. Sometimes, analytical choices [36,37] may be aimed at facilitating the identification of as strong an effect as possible [39], leading to potentially inflated or widely different estimates [40].

As we show, multivariate models tend to give weaker effects than univariate ones, but this is not always the case. Nevertheless, the optimal choice of model may be difficult to identify. For example, adjusting for confounders is appropriate, whereas adjusting for variables that may be in the path that explains the effect of a risk factor may inappropriately diminish the adjusted effect size. Often it is difficult to understand which variables are confounders and which are not.

Despite our comprehensive and systematic approach to gathering and evaluating all the literature related to pterygium risk factor epidemiology, limitations exist. First, although we went at great length to ensure that all relevant articles were identified by creating a sensitive literature...
search strategy, translating articles published in languages other than English, and having two researchers going through and validating our database, we cannot exclude the possibility that some articles were overlooked. Second, we focused on what is reported, but it is quite likely that additional unreported risk factors and models were measured and fit. However, if something is not reported, it does not inform the research literature. Third, it is possible that in some studies, some risk factors may not have been relevant to model, for example, if all participants had the same value (e.g., if a study only enrolled women, gender would not be relevant to adjust for). Fourth, outcomes of interest other than pterygium may exhibit different field-wide profiles with more (or even less) consistency. Routine application of this approach across diverse fields of risk factor epidemiology will inform the extent to which these analytical choices make a difference. Fifth, although our meta-analyses have identified summary estimates for all risk factors ever studied in the context of pterygium and these have been duly reported, the quality of the data available cannot be necessarily trusted to permit firm conclusions based on those results.

Approaches similar to our field-wide meta-analysis have been used in genetic epidemiology to generate field synopses, where databases are dedicated to collecting all genetic association studies on specific outcomes (phenotypes) [41–43]. However, in the case of genetics, adjustments for covariates are not an issue and human-genome epidemiology has made tremendous progress with the advent of platforms that allow covering the entire genome. Exposure-wide assessments, such as “Exposome/Environment-wide Association Studies” (EWAS), are similarly being proposed in nongenetic variables [44–49], to not only facilitate discovery, but also standardize variable adjustments and variables. However, until EWAS-like approaches become more commonplace and widely accepted, single studies are likely to continue covering very different aspects of the putative risk factors’ space. Field-wide evaluations provide a way of recognizing the quality of current data and selective availability of evidence and aid to characterize how susceptible to potential bias entire fields of risk factor epidemiology may be. They also provide visual tabulation of the gaps in a field, thus guiding further research in the field. To enhance field-wide analyses, we recommend executing EWAS to standardize, use of reproducible methods to identify covariates [11], estimating and reporting required power before commencing data collection and calibrating P-values with a database of gold standard negative controls [8,50].

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2015.09.004.

References
