Research Article

Statistics in Medicine

Received 6 October 2010, Accepted 11 April 2011 Published online 5 July 2011 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4284

Mixed treatment comparison of repeated measurements of a continuous endpoint: an example using topical treatments for primary open-angle glaucoma and ocular hypertension

Helen A. Dakin,[a*†] Nicky J. Welton,[b] A. E. Ades,[b] Sarah Collins,[c] Michelle Orme[c] and Steven Kelly[d]

Mixed treatment comparison (MTC) meta-analyses estimate relative treatment effects from networks of evidence while preserving randomisation. We extend the MTC framework to allow for repeated measurements of a continuous endpoint that varies over time. We used, as a case study, a systematic review and meta-analysis of intraocular pressure (IOP) measurements from randomised controlled trials evaluating topical ocular hypotensives in primary open-angle glaucoma or ocular hypertension because IOP varies over the day and over the treatment course, and repeated measurements are frequently reported. We adopted models for conducting MTC in WinBUGS (The BUGS Project, Cambridge, UK) to allow for repeated IOP measurements and to impute missing standard deviations of the raw data using the predictive distribution from observations with standard deviations. A flexible model with an unconstrained baseline for IOP variations over time and time-invariant random treatment effects fitted the data well. We also adopted repeated measures models to allow for class effects; assuming treatment effects to be exchangeable within classes slightly improved model fit but could bias estimated treatment effects if exchangeability assumptions were not valid. We enabled all timepoints to be included in the analysis, allowing for repeated measures to increase precision around treatment effects and avoid bias associated with selecting timepoints for meta-analysis.The methods we developed for modelling repeated measures and allowing for missing data may be adapted for use in other MTC meta-analyses. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: mixed treatment comparison meta-analysis; network meta-analysis; meta-regression; panel data; primary open-angle glaucoma; ocular hypertension

1. Introduction

Whereas standard meta-analytical techniques estimate differences between two treatments based on head-to-head trials, mixed treatment comparison (MTC) or network meta-analysis estimates the relative efficacy of multiple treatments by taking account of the entire evidence network (including both direct and indirect evidence) [1–7]. These methods have been applied to continuous [3–6, 8, 9] and binary [3, 6, 7, 10] endpoints in both Bayesian [3–5, 8–10] and frequentist [6, 7] frameworks, although Bayesian methods have been used more widely to date because of their intuitive interpretation and the flexibility to use alternative model specifications and estimate different output parameters.

Trials commonly report results at multiple timepoints, which raises additional challenges for meta-analysis: particularly for endpoints that systematically vary over the day and/or over the course of

[a] *Health Economics Research Centre, University of Oxford, UK*
[b] *School of Social and Community Medicine, University of Bristol, UK*
[c] *Abacus International, Oxfordshire, UK*
[d] *Pfizer Ltd, Surrey, UK*
**Correspondence to: Helen A. Dakin, Health Economics Research Centre, Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford, OX3 7LF, UK.*
[†] *E-mail: helen.dakin@dph.ox.ac.uk*

treatment, such as intraocular pressure (IOP), which affects the risk of primary open-angle glaucoma (POAG) and blindness [11]. Although methods to allow for repeated measures in pairwise meta-analyses have been reported [12, 13], most meta-analyses analyse data at each time separately or pool outcomes at a single timepoint, which requires judgements about which timepoints are analysed that make the resulting inference vulnerable to bias. Among 13 published meta-analyses evaluating the IOP-lowering effects of three or more drugs [11, 14–25][‡], four did not explicitly allow for temporal IOP variations [11, 14, 15, 21], whereas two analysed data at end of trial or at 3 months and investigated the effect of treatment duration using meta-regression [25, 26]. Seven studies pooled data separately at different times of day [16, 20] or at peak and trough [18, 19, 22–24], which produces multiple estimates of treatment effect but does not assess whether any differences in relative effects between timepoints represent genuine trends or simply random fluctuation. Although two meta-analyses used MTC [17, 18], neither allowed for repeated measures or temporal trends within MTC.

Two studies have presented methods allowing for repeated measures in MTC, although these cannot be directly applied to outcomes such as IOP. Lu *et al*. used survival analysis methods to model repeated measures of a binary endpoint (the cumulative proportion of patients healed) within MTC [27]. However, this method cannot be generalised to continuous outcomes that may increase or decrease over time. Wandel *et al*. present an MTC model that allows for repeated measures of a continuous outcome by assuming that observations at each timepoint are exchangeable around study-arm-specific means [28]. However, this model implicitly assumes that variability between repeated measures is due to chance and that the outcome measure of interest does not show any structural patterns over time, which is frequently not the case. In particular, this would not be an appropriate model of IOP, which is generally highest in the early morning and falls during the day [29–31], although diurnal patterns vary between patients [30, 32, 33]. We are unaware of any previous studies describing MTC models that allow for trends over time in continuous outcomes.

We set out to adapt the MTC framework to estimate relative effects while allowing for repeated measurements of a continuous endpoint that shows random and systematic variation. A secondary objective was to assess the impact of controlling for key covariates and modelling class effects and develop methods for imputing missing data. Methodological developments are illustrated using a meta-analysis of IOP measurements from 50 trials evaluating multiple classes of treatments for POAG and ocular hypertension (OH); this large dataset provides repeated IOP measurements taken at multiple times of day and several follow-up visits, although some trials did not report standard errors. The methods developed can be applied to similar datasets in other settings.

2. Methods

We conducted a comprehensive systematic review of randomised controlled trials evaluating topical therapies for POAG or OH to identify studies that could be used to develop MTC models allowing for repeated measures. The methods and results of the systematic review are available at Appendix A. Fifty trials evaluating 40 treatments met the inclusion criteria for the meta-analysis.

In order to take account of the network of evidence on ocular hypotensives (Figure 1) and allow both direct and indirect evidence to inform estimates of efficacy, we conducted the meta-analysis using MTC in a Bayesian framework.

2.1. Mixed treatment comparison models for repeated observations of a continuous outcome

Our dataset includes up to 18 repeated IOP measurements for each study arm that are taken at different times of day and/or different follow-up visits (details available at Appendix A). Baseline observations (taken at week 0) are excluded and the number and timing of repeated measurements varies between trials. We adapted the standard model for MTC on a single continuous outcome to develop six new models that allow for repeated measures. Code for all models is available at Appendix B.

All models assume a normal likelihood (based on central limit theorem), which was adapted from earlier models [2, 3, 8, 9, 34] to allow for repeated measurements of each study arm:

$$y_{s,k,i} \sim N\left(\nu_{s,k,i}, \mathrm{var}(y_{s,k,i})\right) \tag{1}$$

[‡]*Meta-analyses were identified through a MEDLINE search conducted on 24 August 2010 using PubMed.*

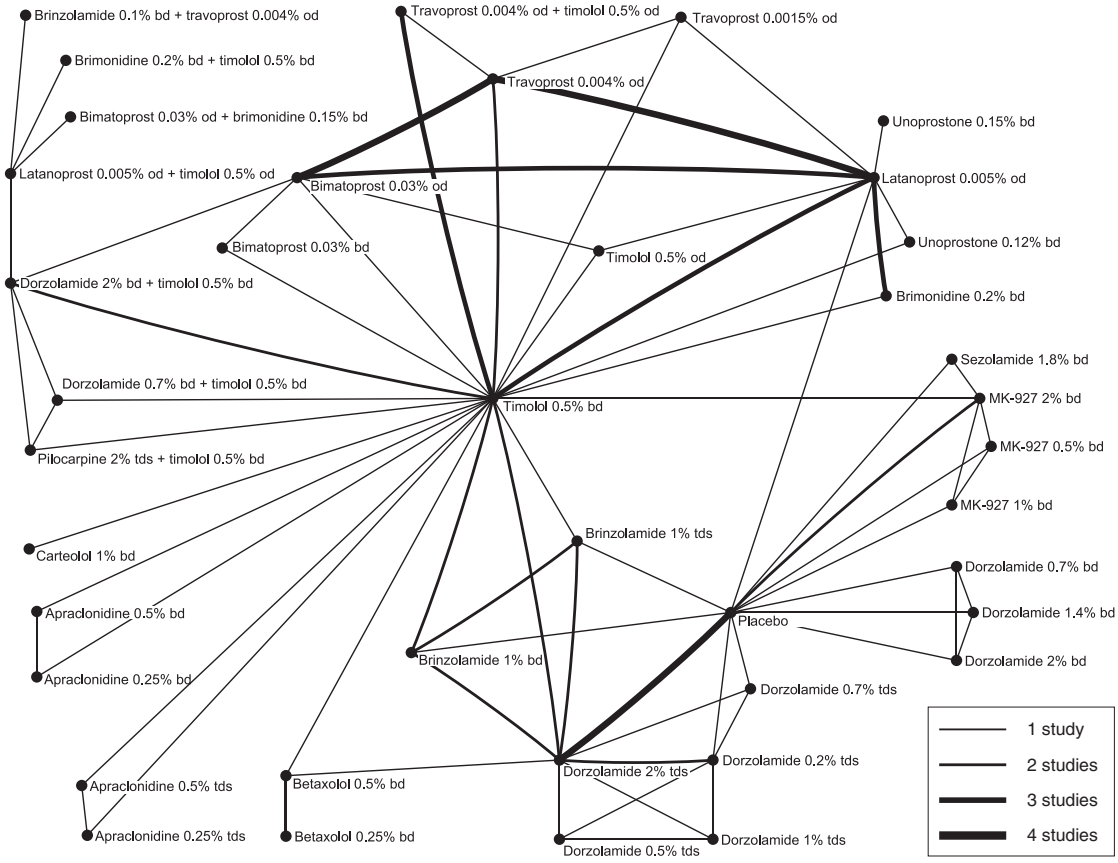

Figure 1. Network diagram showing the network of evidence across all 50 studies included in the MTC meta-analysis. Each line represents one or more randomised controlled trial; line thickness is proportional to the number of trials. bd, twice daily; od, once daily; tds, three times daily.

where: $y_{s,k,i}$ is the observed mean IOP at timepoint i in arm k of trial s, which has standard error (SE, $\sqrt{\text{var}(y_{s,k,i})}$) equal to that observed in trial s; and $v_{s,k,i}$ is the predicted (mean) outcome at timepoint i in arm k of trial s.

Model 1: Mean intraocular pressure and random relative effects constant over time

Model 1 makes the simplest assumption that allows for repeated measures: namely that IOP and relative treatment effects are constant over the day and over the course of treatment and that any variations in IOP between repeated measurements of arm k, study s are due to sampling error. In this model, $v_{s,k,i}$ is therefore the same for all timepoints (i) of any given arm (k) of study s:

$$v_{s,k,i} = \mu_s + \delta_{s,k} \tag{2}$$

where μ_s is the estimated mean outcome in the baseline arm of trial s (for which $k = 1$) and $\delta_{s,k}$ is the difference in mean outcome between arm k and arm 1 for study s (with $\delta_{s,k} = 0$ for $k = 1$).

In all models, the pooled treatment effects (d_t) of treatment t relative to reference treatment A are the *basic* parameters [35]. All other pooled relative effects are obtained from these basic parameters via the consistency equations [36], which follow from exchangeability assumptions [37]; for example the difference in pooled mean IOP between treatments B and C equals $d_B - d_C$. Timolol 0.5% twice daily (bd) is taken to be the baseline treatment because it lies at the centre of the evidence network (Figure 1), having head-to-head comparisons against 22 other treatments.

For random treatment effect models, such as model 1, the treatment effects in arm k of study s ($\delta_{s,k}$) are assumed to be drawn from a common distribution with mean $d_{t_{s,k}} - d_{t_{s,1}}$ and (between-studies) standard deviation (SD) τ:

$$\delta_{s,k} \sim N\left(d_{t_{s,k}} - d_{t_{s,1}}, \tau^2\right) \tag{3}$$

where $t_{s,k}$ indicates the treatment used in the k^{th} arm of study s. We assume that τ is constant across treatment comparisons (homogenous variance). For multi-arm trials, Equation (3) is replaced by a multivariate normal random-effects distribution, in which the correlations between relative treatment effects are set equal to 0.5, based on the assumption of heterogeneous variances across treatment comparisons [3, 34]. μ_s and d_t were assigned separate non-informative Gaussian priors ($\mu_s \sim N(0, 10\,000)$, $d_t \sim N(0, 10\,000)$), whereas the between-studies SD (τ) was given a wide uniform prior ($\tau \sim U(0, 20)$).

Model 2: Unconstrained baseline, random relative effects constant over time
Exploratory data analysis (available at Appendix A) suggested that IOP varies over the day and over the treatment course. Because assuming constant IOP (as in model 1) is unlikely to be appropriate, model 2 estimates mean IOP in arm 1 ($\mu_{s,i}$) independently for each timepoint of each study, allowing estimated IOP levels in arm 1 to fluctuate in line with observed IOP:

$$\nu_{s,k,i} = \mu_{s,i} + \delta_{s,k} \tag{4}$$

We assigned separate, uninformative Gaussian priors for each timepoint of each study ($\mu_{s,i} \sim N(0, 10000)$); all other model details are identical to model 1.

Whereas model 1 assumes that IOP is constant over time, model 2 makes no assumptions about the relationship between IOP and time and is therefore valid for outcome measures showing systematic temporal trends, even if the relationship between time and outcome is non-linear or varies between studies. However, as for model 1, we assumed relative treatment effects ($\delta_{s,k}$) to be constant over time.

Model 3: Unconstrained baseline, fixed relative effects constant over time
Model 3 comprises a variant of model 2 with fixed treatment effects, whereby we assumed relative effects to be identical for all studies, such that $\delta_{s,k} = d_{t_{s,k}} - d_{t_{s,1}}$ and all studies estimate a common mean treatment effect.

Model 4: Unconstrained baseline, random relative effects vary over the day
In model 4, we estimated treatment effects separately for four time bins (0700–0800 h, 0900–1000 h, 1100–1500 h and 1600–2100 h). This allows relative treatment effects (i.e. the difference in IOP between treatment arms) to vary over time without making strong assumptions about the relationship between time and treatment effect, assuming only that the treatment effects are piecewise constant within each of these time intervals. We estimated mean treatment effects for each time bin b ($d_{t,b}$) independently and assumed them to be constant across the bin.

$$\nu_{s,k,i} = \mu_{s,i} + \delta_{s,k,i} \quad \delta_{s,k,i} \sim N\left(d_{t_{s,k},b_{s,i}} - d_{.t_{s,1},b_{s,i}}, \tau^2\right) \quad d_{t,b} \sim N(0, 10000) \tag{5}$$

Because mean post-baseline IOP for each study arm at any given time of day generally varied by less than 1 mmHg (see Appendix A), we assumed that treatment effects were constant over the course of treatment.

Model 5: Unconstrained baseline, random relative effects constant within treatment class and over time
We also evaluated two models allowing for class effects (models 5 and 6). Model 5 assumes that treatment effects are constant within each treatment class (e.g. assuming that all prostaglandins reduce IOP by the same amount). This model is identical to model 2, except that d_t equals D_c, the pooled mean treatment effect for treatment class c relative to class 1 (beta-blockers).

$$d_t = D_{c_t} \quad D_c \sim N(0, 10000) \tag{6}$$

We defined treatment classes by chemical class (details available at Appendix B); we considered unoprostone to comprise a separate class (docosanoid prostaglandins) because its chemical structure, mode of action and efficacy differ from licensed prostaglandins [38, 39].

Model 6: Unconstrained baseline, random relative effects exchangeable within treatment class and constant over time
Model 6 assumes that treatment effects are exchangeable within classes. In other words, we assumed the treatment effects for treatments in the same class to be similar, and we drew from a normal distribution with class mean D_c and variance T^2, which indicates the variability between treatments in the same class. This model was identical to model 2, except that

$$\begin{array}{ll} d_t \sim N\left(D_{c_t}, T^2\right) & \text{for classes including >1 treatment} \\ d_t = D_{c_t} & \text{for classes including only 1 treatment} \\ D_c \sim N(0, 10\,000) \qquad T \sim U(0, 20) & \end{array} \tag{7}$$

For simplicity, we assumed the within-class SD (T) to be common to all treatment classes, although this assumption was relaxed in sensitivity analysis. We estimated all treatment effects relative to timolol 0.5% bd, which is considered separately from other beta-blockers.

2.2. *Methods for imputation of missing SD*

Across the 278 datapoints, 61 (22%) did not report SDs, SEs or 95% confidence intervals (CIs) around mean IOP. In order to include these datapoints within the analysis and allow for variations in sample size, we predicted missing SDs based on studies reporting SDs and propagated the uncertainty around predicted values into the meta-analysis.

We conducted a separate meta-analysis in WinBUGS (code available at Appendix B) to estimate the distribution of SDs for all observations meeting inclusion criteria that reported SDs or for which SDs could be calculated from the SE or 95% CI. A gamma distribution with estimated parameters alpha=11.54 and beta=3.37 was found to fit the SD data best. We randomly sampled predicted values for missing SDs from this gamma distribution and used them in MTC meta-analyses to impute missing SD values and allow for uncertainty around predictions.

2.3. *Modelling covariates*

We evaluated the effect of controlling for one of four arm-level covariates within the best-fitting model to assess whether differences between arms affected relative treatment effects: baseline IOP; whether treatment comprised an unfixed dose combination (rather than monotherapy or fixed-dose combination); whether treatment was a gel-forming timolol solution (not standard eye-drops); and whether treatment comprised a prostaglandin administered before 1800 h (contrary to dosing recommendations [40–42]). We evaluate each covariate on an incremental basis, whereby the meta-regression coefficient (β) is multiplied by the difference in the covariate value (X) between arm k and arm 1:

$$\delta_{s,k} \sim N\left(d_{t_{s,k}} - d_{t_{s,1}} + \beta(X_{s,k} - X_{s,1}), \tau^2\right) \quad \beta \sim N(0, 10000) \tag{8}$$

$X_{s,k}$and $X_{s,1}$ cancel out when baseline IOP or dosing is identical in both arms; covariates therefore only affect relative treatment effects when $X_{s,k}$ differs between arms k and 1. For dosing covariates, β indicates the extent to which the predicted treatment effect for $t_{s,k}$ vs $t_{s,1}$ increases when arm k is unfixed/gel-forming solution/non-evening prostaglandin dosing and arm 1 is not. When baseline IOP is considered as an arm-level covariate, β indicates the extent to which predicted treatment effects increase (or decrease) for every 1 mmHg increase in baseline IOP imbalance.

We also evaluated two study-level covariates (year of publication and mean baseline IOP across all arms):

$$\delta_{s,k} \sim N\left(d_{t_{s,k}} - d_{t_{s,1}} + \left(\beta_{t_{s,k}} - \beta_{t_{s,1}}\right)\left(X_s - \bar{X}\right), \tau^2\right) \tag{9}$$

where $\beta_{t_{s,k}}$ equals zero if arm k evaluated timolol 0.5% bd and equals β otherwise and where $\beta \sim N(0, 10000)$.

We added each covariate to the best fitting model individually to assess its effect on model fit (Section 2.4). We assumed the effect of all six covariates to be common across all treatments and studies.

2.4. *Implementation*

Models were fitted using WinBUGS version 1.4.3 (The BUGS Project,Cambridge, UK [43]); code for all models is available at Appendix B. We run two chains with different initial values for each model. 10 000–50 000 sampled updates of each chain followed a burn-in of 50 000–200 000 updates; we assessed convergence using graphical measures within WinBUGS (e.g. BGR [44] and history) and the ratio of Monte Carlo error divided by SD.

When choosing between the models described in Section 2.1, we considered models with lower residual deviance [10,45] to fit the data better; we considered models to fit well if residual deviance was lower than the number of datapoints [46]. We also calculated deviance information criterion (DIC) [43,47] by

adding residual deviance to the effective number of parameters (pD, estimated using the DIC tool in WinBUGS) [43, 47], which takes account of model complexity as well as model fit. We added each covariate to the best-fitting model separately to evaluate whether it improved model fit; we included any covariates or covariate interactions that reduced DIC by at least three in the final model, because differences of 3–5 or more are generally considered important [43, 47]. We considered meta-regression coefficients to be statistically significant if the 95% credible interval (CrI) around β did not include zero.

3. Results

3.1. Model selection and results

We evaluated six models allowing for repeated IOP measurements. Model 1 assumed that IOP levels were constant over the day and over the course of treatment, and that relative treatment effects (i.e. the difference in IOP between treatment arms) were also constant over time. As expected, this model fitted the data very poorly, with a residual deviance of 2632 (Table I), indicating that predicted values differ from observed IOP by 3.55 SEs; by contrast, residual deviance for a well-fitting model of 742 datapoints should be around 742 [46].

Model 2 estimated mean IOP in the baseline arm separately for each timepoint of each study, thereby allowing IOP to vary over time without making any assumptions about this relationship other than that the true *difference* in IOP between study arms is constant over time. This model fitted the data very well, having a residual deviance of 704 (substantially lower than the number of datapoints). The finding that allowing IOP levels to vary between timepoints in the same study substantially improves model fit strongly suggests that it is inappropriate to ignore diurnal and longer-term fluctuations in IOP within meta-analyses of ocular hypotensives.

Model 3, a variant of model 2 with fixed-effects, fitted the data less well (residual deviance 968 compared with the 742 data points), indicating that there is substantial heterogeneity in treatment effects across studies, justifying use of a random-effects model. Use of random-effects is also supported by the relatively high between-studies SD (τ) estimated in model 2 (0.679 mmHg, 95% CrI: 0.502, 0.909).

We relaxed the assumption that treatment effects are constant over time in model 4. In this model, we estimated treatment effects independently for four parts of the day (0700–0800 h, 0900–1000 h, 1100–1500 h and 1600–2100 h), which were chosen to maximise the number of treatments for which treatment effects could be estimated at each time. This model fitted the data less well than model 2 (residual deviance: 748), suggesting that it is reasonable to assume that treatment effects are constant over time. However, model 4 had a substantially lower between-studies SD than model 2 (0.218 versus 0.679), which suggests that some of the heterogeneity in treatment effect within model 2 is explained by differences in the time of day at which IOP measurements are taken. The timepoint-specific treatment effects estimated by model 4 also differed significantly from the overall effects estimated by model 2 for

Table I. Goodness of fit measures for the six models investigated.

Model	pD	Residual deviance*	DIC
Models estimating effects for individual treatments separately			
1. IOP and random treatment effects constant over time	291	2632	2923
2. Unconstrained baseline but constant random treatment effects (allows for IOP variation over time)	381	704	1084
3. As model 2 but fixed treatment effects	369	968	1337
4. As model 2 but variable random treatment effects that are estimated separately for 4 different parts of the day	466	748	1214
Models allowing for class effects			
5. As model 2 but random treatment effects assumed to be identical for treatments in the same class	380	709	1089
6. As model 2 but random treatment effects assumed to be exchangeable among treatments in the same class	378	703	1080

Note: * The data set includes 742 data points; well-fitting models would therefore be expected to have a residual deviance of 742. DIC, deviance information criterion (DIC = pD + residual deviance); IOP, intraocular pressure; pD, effective number of parameters (estimated using DIC tool in WinBUGS).

nine out of the 40 treatments ($p < 0.05$). The differing results may suggest that treatment effects vary over the day in a way that is not fully captured by model 4 or that cannot be estimated reliably from this dataset. Alternatively, model fit may be worsened by estimating treatment effects independently for each bin based only on the subset of trials that reported IOP at that time of day.

We also evaluated two models of class effects. Model 5 made the very strong assumption that treatment effects were constant within classes (e.g. that all prostaglandins have identical effects on IOP) and had worse fit than model 2 (residual deviance: 709 versus 704 for model 2; DIC: 1089 versus 1084 for model 2).

Model 6 assumed that treatment effects are exchangeable within treatment classes: for example that the treatment effect for prostaglandin t is drawn from the same distribution as other prostaglandins. This model had similar fit to model 2 (residual deviance: 703, versus 704 for model 2) and improved DIC because of the lower number of parameters (1080 versus 1084 for model 2), but produced different estimates of relative effects for individual treatments (Figure 2). Model 6 estimated the within-class SD (T) to be 0.32 (95% CrI: 0.02, 0.69) mmHg, suggesting that the relative effects for treatments within the same class are very similar, although this model had a larger between-study SD than model 2 (0.75 versus 0.68 mmHg). Because model 6 makes additional assumptions regarding exchangeability of treatment effects within classes, its results would be prone to bias if this exchangeability assumption were not valid for all treatment classes. Model 6 also makes the strong assumption that between-treatment within-class SD (T) is common to all classes. Although there is no clinical basis for this assumption, a sensitivity analysis allowing T to vary had marginally worse model fit (residual deviance: 704) and produced within-class SDs that were sensitive to priors, suggesting that our dataset is insufficiently large to reliably estimate T separately for each class.

Because models 2 and 6 had similar model fit but produced different results, results for both models are presented (Figure 2) in line with best practice [43, 47]. We explored the effect of adding covariates and changing model specification within model 2 because this model avoids assuming exchangeability between treatments in each class.

We added six covariates into model 2 one at a time to assess the effect on model fit (Section 2.3), which were all non-significant ($p > 0.05$), worsened or had negligible effect on model fit (results available at Appendix C) and increased between-studies SD (demonstrating that they do not account for between-studies heterogeneity). We therefore did not evaluate interactions between covariates, and no covariates were included in the final model.

The results of model 2 suggest that the prostaglandins bimatoprost, latanoprost and travoprost comprise the most effective licensed monotherapy agents considered in the meta-analysis (Figure 2, Appendix C). Model 2 showed marked variation between drugs in each class and between different doses, which was substantially reduced when exchangeable class effects were assumed. Some such variations (e.g. the finding that betaxolol is less effective than timolol) are likely to reflect genuine differences or dose-response relationships that have been observed previously [18, 19], although other differences may be artifactual: particularly the difference between twice-daily and three-times-daily dorzolamide dosing, which may result from a lack of head-to-head trials comparing dosing frequencies (Figure 1).

3.2. Sensitivity analyses

We compared the results of repeated measures models with those of the standard MTC model for single observations within a sensitivity analysis that included only one mean IOP measurement from each study.[§] SEs around treatment effects were, on average, 21% higher in this analysis than in model 2 and treatment effects differed by an average of 0.447 mmHg, although the difference between the models was statistically significant for only one treatment (betaxolol 0.25% bd). This confirms that allowing for repeated measures increases precision in addition to avoiding arbitrary decisions about which timepoint to use.

We evaluated a logarithmic model to assess whether treatment has a multiplicative effect on IOP because previous research [30, 48] and contact with clinicians working in the field suggested that the treatments that produce greatest reductions in mean IOP may also reduce the magnitude of diurnal fluctuations; this model fitted very poorly, suggesting that additive treatment effects were appropriate.

[§]*The IOP measurement included in this analysis was arbitrarily chosen to be the measurement closest to 0901 h 28.1 days after start of treatment. Studies not meeting I1 and I2 inclusion criteria (Appendix A) and those that reported data only in the afternoon or more than 13 weeks after start of treatment were excluded. This sensitivity analysis was conducted using model 2 but including only one observation per study.*

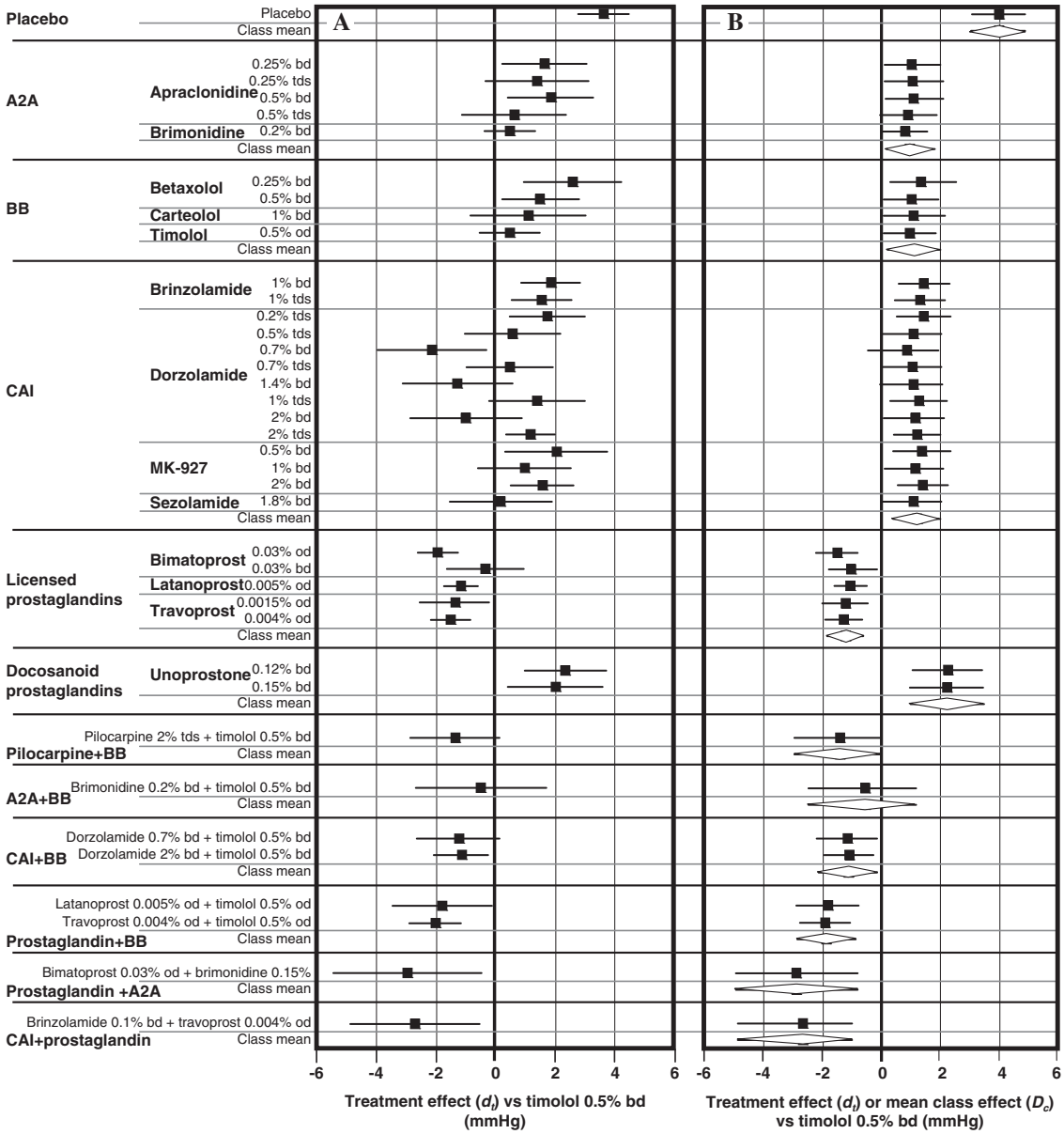

Figure 2. Results of the best fitting models. A. Model 2 (relative effects estimated independently for each treatment); B. Model 6 (relative effects assumed to be exchangeable within classes). Error bars indicate 95% credible intervals around relative effects for individual treatments. Diamonds indicate class means (D_c) and their 95% credible intervals. bd, twice daily; od, once daily; tds, three times daily; A2A, A2-adrenergics/sympathomimetics; BB, beta-blockers; CAI, carbonic anhydrase inhibitors.

Further sensitivity analyses demonstrated that widening or narrowing the priors for τ, d_t and $\mu_{s,i}$ had no effect on the results of model 2. Omitting studies that did not report SD or SE slightly improved model fit, but had no significant effect on mean relative effects, increased SEs around all treatment effects and prevented estimation of treatment effects for three of the 40 treatments.

4. Discussion

In this study, we have developed models to conduct MTC on datasets with repeated measurements of a continuous endpoint that varies over time. This allowed us to use all datapoints to estimate average treatment effects, thereby increasing precision and minimising the potential for bias that could arise from analysts selecting a single timepoint at which to conduct their analysis. However, reporting bias could still affect our analysis if trial investigators were selective in the timepoints at which IOP measurements

were taken and/or results were presented: particularly if trials on one or more treatments preferentially reported results at times when between-treatment differences in IOP were greatest.

Wandel *et al*. previously presented a model of repeated measures in MTC which assumes that the repeated measurements are randomly drawn from a common normal distribution [28]. Our models relax this assumption by allowing outcome estimates for the baseline arm to follow observed data even if outcomes vary systematically over the day or over the course of treatment, but without making any assumptions about the shape of these trends. The same methods and WinBUGS code (available at Appendix B) can be applied to meta-analyses of continuous outcomes in other disease areas and are valid for use on outcomes that show systematic temporal trends as well as those that show random fluctuations. By allowing for repeated measures, our study included many more studies and treatments than previous MTC meta-analyses on ocular hypertensives [17, 18]; the results of our analysis generally confirmed the findings of these previous studies [17, 18], but highlighted some differences in treatment rankings that may be due to differences in modelling methodology or study characteristics.

The results suggest that models allowing for variations in IOP over time fit this dataset much better than models that assume IOP is constant over time. The same is likely to be true for other outcome measures, such as blood pressure, that show diurnal variations or change over a longer period. This highlights an additional danger of 'naïve' comparisons in which outcomes from separate arms of different studies are pooled and compared without respect for randomisation because absolute IOP measurements at 0800 h from one study are not comparable with measurements taken at 2000 h in a different study, even if the same treatment is given to a similar population. Temporal variations will increase the risk of bias in naïve comparisons over and above the problems of selection bias and overestimating precision that have been described previously [49].

Although models 2–6 allow for variations in IOP over time, they do not make any assumptions about the nature of the relationship between IOP and time except for assuming constant additive treatment effects, thereby avoiding the bias that could be introduced by using an inappropriate functional form. However, our models cannot predict mean on-treatment IOP at any given time of day or provide insights into how IOP varies over the day or over the course of treatment or how these trends vary between treatments. We also investigated parametric models that assumed that IOP had a kinked linear relationship with the number of weeks since start of treatment and used various functional forms to model the relationship between IOP and time of day, including several sinusoidal functions, spliced linear functions and polynomial, exponential and logarithmic functions (details available on request). However, all of the parametric forms that we explored fitted the data poorly, with the best fitting parametric model having a mean residual deviance of 1.54 per datapoint. Attempts to capture the relationship between time and IOP in a single model may have been hindered by differences between treatments or trials in the relationship between IOP and time and/or by differences in the timing and frequency of drug administration. However, given adequate data on time since last dose and a sufficiently large dataset, it may be possible to identify alternative parametric models of how IOP changes over time and how such trends are affected by treatment. Parametric meta-regression models of the relationship between outcomes and time may also be developed for other widely-researched outcome measures: particularly if the appropriate functional form is known and if patient-level data are available to inform model development.

The finding that model 2 (which assumes constant treatment effects) fits the data better than model 4 (which allows treatment effects to vary over the day) suggests that assuming constant treatment effects is a reasonable simplification, although some previous studies suggest that diurnal IOP variations are affected by treatment [30, 33, 48, 50]. However, it is possible that treatment effects do vary over time in a way not captured by model 4: particularly as allowing treatment effects to vary over the day substantially reduced between-studies heterogeneity. In particular, previous research [30, 48] and the results of our parametric models suggest that the treatments producing greatest reductions in IOP (prostaglandins, beta-blockers and combination therapies) may also reduce the magnitude of diurnal IOP fluctuation. This therefore remains an area in which further research is needed: particularly because patients with larger diurnal IOP variations are at greater risk of developing glaucoma or experiencing disease progression [51].

We imputed SDs for studies that did not report measures of variability based on the distribution of SDs within studies that did report them. At least seven previous meta-analyses in glaucoma have imputed missing SDs, although four relied on SDs around baseline IOP being available for each study [18–20, 22] and three applied the overall mean SD [17, 21] or SE [52] to all studies that did not report them, which makes no allowance for uncertainty. By contrast, the methods we used allow for varying sample size and

propagate the uncertainty around imputed SDs into the main meta-analysis; to our knowledge, this has not been done previously within MTC.

Additional analyses evaluated class effects, which have been investigated previously [53]. Although assuming treatments in the same class to have identical efficacy worsened model fit, model 6 (which assumed treatment effects to be exchangeable within classes) fitted the data slightly better than model 2 but estimated different relative effects for individual therapies because of shrinkage towards class means. The differences in estimated treatment effects between models 2 and 6 are due to the assumption of exchangeability within treatment class; decisions around which model is most appropriate therefore depend on whether this exchangeability assumption is considered credible. Although in some situations, exchangeability may be a reasonable simplification that increases precision around treatment effects, it is debatable whether relative effects of specific treatments are genuinely exchangeable. Dose-response relationships will also violate the exchangeability assumption when doses of the same drug are modelled as different treatments in the same class. It is therefore essential to consider in each specific application whether class effects are clinically realistic and how treatments are best grouped. Class-effect models may be particularly useful for clinical guidelines concerning use of drug classes, because they directly estimate class means and the variability between treatments in the same class.

We excluded around 15% of studies from our analysis as they reported absolute or percentage change in IOP from baseline or the proportion of patients meeting targets rather than absolute on-treatment IOP. Future studies could adapt our models to include these alternative measures. Drug doses could also be modelled in a more sophisticated way by allowing for dose as a covariate, which could reduce model complexity and improve consistency between treatment effects at different doses without assuming doses to be exchangeable, but would require additional assumptions about dose relationships.

Although the low residual deviance indicates that the evidence network is consistent, there was a large amount of heterogeneity that was not explained by publication date, baseline IOP or dosing regimen. Whereas differences in the time of IOP measurement may account for some heterogeneity, future analyses could explore heterogeneity further by controlling for additional covariates, such as time since last dose, previous treatment or study quality, although these variables are not reported consistently in all trials.

Although we would expect measurements of the same participants at different timepoints to be correlated, we assumed that the likelihood for each repeated measurement of any given study arm is independent because none of the 264 studies reviewed reported the correlations between repeated observations. One approach to this problem would be to perform sensitivity analyses using different assumed values for the autocorrelation over time. This is analogous to standard meta-analyses on differences between repeated measures based on studies that report only summary statistics 'before' and 'after'; such analyses often assume an arbitrary value for the correlation (e.g. +0.5 [54]). However, with more than two repeated measures, this becomes more complex, as we might expect correlations to weaken over time. We hypothesise that correlations in the likelihood are unlikely to affect mean treatment effects, although the uncertainty around treatment effect estimates may be sensitive to such assumptions; this hypothesis could be tested in future work where individual patient data are available.

The models we developed to allow for repeated measures within meta-analyses of continuous outcome measures that vary over time may be adapted for use in other meta-analyses. However, such analyses require primary trials to report full results (including patient numbers and variability) at all timepoints investigated.

Appendix A: Systematic review of randomised controlled trials evaluating topical therapies for primary open-angle glaucoma or ocular hypertension

Methods of the systematic review

We searched Medline, EMBASE and the Cochrane Library up to 30 May 2008 to identify relevant studies that could be used to explore the effect of controlling for repeated measures in an MTC of POAG and OH trials. The search terms included: 'glaucoma,' 'ocular hypertension,' 'randomized controlled trial,' 'clinical trial,' and generic/brand names for commonly-used topical medications for OH and/or POAG.

Two reviewers independently assessed identified studies to ascertain whether they met a set of predefined inclusion/exclusion criteria (I1, Table AI). They extracted study characteristics and the timing and results of intraocular pressure (IOP) measurements for all studies meeting I1 inclusion criteria. They rounded the timing of IOP measurements to the nearest whole week and nearest hour. A second reviewer checked the extracted data, and they resolved any discrepancies through discussion; a third

Table AI. Inclusion and exclusion criteria.

Criterion	Included	Excluded
I1: Inclusion/exclusion criteria pre-defined prior to literature searches. Data were extracted on all studies meeting these criteria.		
Population	• Qualifying disease: ocular hypertension and/or POAG • Age: ≥ 18 years • Race: any • Treatment history: treatment-naïve patients or previously treated patients who had discontinued any previous IOP-lowering medication ≥ 2 weeks before measurements baseline	• Studies on patients who had secondary glaucoma, angle closure, pseudoexfoliation glaucoma or surgery prior to enrolment were excluded as effect of ocular hypotensives is likely to be different in these conditions than in POAG.
Study perspective	• Prospective	• Retrospective
Type of study	• RCT	• Any studies with non-randomised allocation to treatment • Studies with a run-in period unless they included a washout of ≥ 2 weeks before the baseline measurement • Crossover trials without a ≥ 2 week washout period between treatments
Language	• English language	-
Study duration	• At least 1 week of treatment and follow-up	-
Sample size	• Any	-
Interventions/ treatments	• Evaluated at least one topical pharmacological treatment licensed in the UK for POAG/ocular hypertension that was administered onto the eye as drops or gel. No restrictions were placed on drug dose or on whether treatment was given as monotherapy or fixed/unfixed combination therapy. Licensed treatments included beta-blockers (betaxolol hydrochloride, carteolol hydrochloride, levobunolol hydrochloride, metipranolol and timolol maleate), prostaglandin analogues (bimatoprost, latanoprost and travoprost), A2-adrenergics/sympathomimetics (brimonidine tartrate and dipivefrine hydrochloride), carbonic anhydrase inhibitors (brinzolamide and dorzolamide) and miotics (pilocarpine).	• Studies on systemic therapies • Studies that did not include at least one treatment licensed in the UK for use in POAG or ocular hypertension
Study outcomes	• Studies were included in the systematic review if they reported any data on IOP at baseline *and* following treatment, whether absolute IOP or change from baseline	
I2: Inclusion/exclusion criteria defined after data collection had commenced but prior to analysis. Studies meeting these criteria were included in the meta-analysis.		
Analysis	• Studies analysed at level of patients	• Studies for which the unit of analysis was the eye rather than the patient
Study population	-	• Studies that exclusively recruited patients who have previously responded to one of the treatments used in that trial (enriched-enrollment design) • Studies that exclusively recruited patients who had previously failed to respond to one of the treatments used in that trial
Criteria for inclusion of datapoints in the analysis	• Studies were included in the meta-analysis if mean absolute IOP and the number of patients for whom IOP was measured were reported separately for each treatment arm. • In order to model how IOP varies over the weeks since start of treatment, for crossover studies, we included only the IOP measurements taken during the	IOP measurements were excluded from the analysis if: • The number of patients with IOP measurements at that timepoint was not reported. • The study did not report both the number of weeks since start of

Table AI. *Continued*

Criterion	Included	Excluded
	first period of treatment within the analysis (i.e. from baseline to the point at which each patient ceased their first therapy). Crossover studies that combined data from patients who received any given treatment in the first period with data from patients who received this treatment second and did not report data on the two periods separately were excluded from the meta-analysis. • For any parallel-group studies where an additional treatment was added in or the dose was changed X weeks after the start of the trial, only IOP measurements taken during the first X weeks of the study were included in the meta-analysis. • For studies reporting data by subgroup as well as for the overall population, we used data for the overall population in the analysis. However, if the study provided data only by subgroup (or if one subgroup did not meet the inclusion criteria for the systematic review), we used the data for each subgroup in the analysis and treated each subgroup as a separate RCT.	treatment and the time of day at which the IOP measurement was taken or if the the IOP value was the average across multiple days, the average across several different times of day (diurnal IOP) or the mean of patients' peak/trough measurements. The analysis was restricted to mean IOP measurements taken at a single known timepoint to allow temporal trends to be explored and as mean peak/trough measurements and standard errors around IOP values averaged over several timepoints may not be comparable with measurements taken at a single time of day. • The IOP value comprised the average across two or more arms of the study. • The IOP measurement was taken during an extension phase of an RCT in which patients in one or more of the study arms did not receive the treatment to which they were initially randomised. • The IOP measurement was taken <1 week or >52 weeks after start of treatment. Measurements taken within a week of starting therapy were excluded as trends and treatment effects observed before all treatments had taken full effect are likely to differ from later measurements. Measurements taken more than a year after start of treatment were excluded as only one study provided data beyond a year, making any predictions in this period unreliable. • IOP measurements taken at baseline and studies that provided mean IOP only at the baseline visit and provided no usable on-treatment data were excluded.
Connectivity		• Studies that were not connected to the evidence network after all other inclusion/exclusion criteria had been applied were excluded.

Note: IOP, increased ocular pressure; MTC, mixed treatment comparison; POAG, primary open-angle glaucoma; RCT, randomised controlled trial; SD, standard deviation.

reviewer checked outliers and studies with unusual designs again. Reviewers applied additional inclusion/exclusion criteria (I2, Table AI) to the studies meeting the systematic review criteria in order to exclude any studies that did not report sufficient data for inclusion in meta-analyses of absolute IOP or that were not comparable with the majority of studies.

Results of the systematic review

Electronic database searches identified 5330 papers, of which 264 studies met the inclusion criteria for the systematic review (I1, Figure A1); a list of these trials is available from the authors on request. Of these 264 trials, 50 studies met the narrower (I2) inclusion criteria and were included in the meta-analysis (Table AII).

Of the 50 studies included in the meta-analysis, 32 had two treatment arms, 10 had three, seven had four and one had five. On average, studies reported mean on-treatment IOP at 5.6 (range: 1–18) timepoints taken on an average of 2.4 (range: 1–7) study visits, giving a total of 278 study-observations, each with IOP data for 2–5 study arms. The time of day when post-baseline IOP measurements were

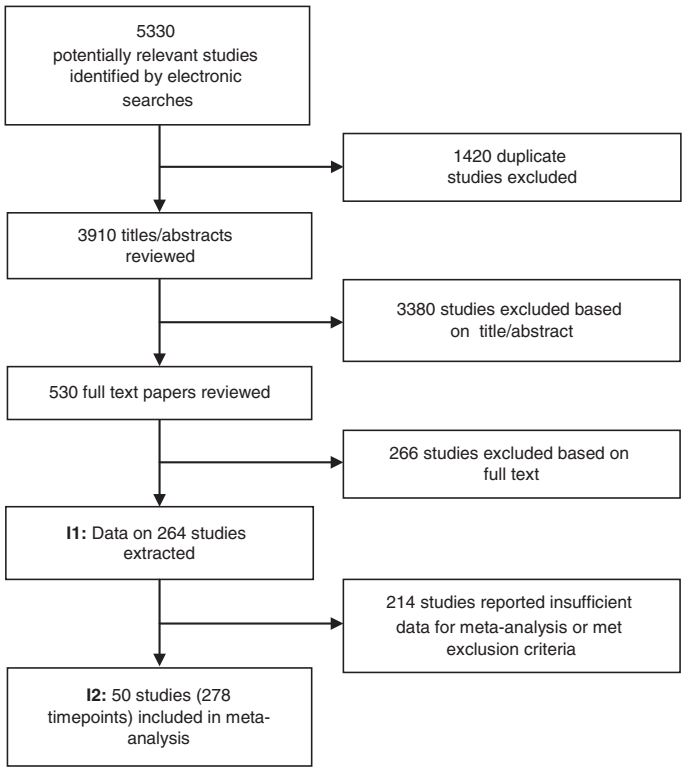

Figure A1. Flow diagram of the results of the search strategy and identified studies.

taken ranged from 0700 h to 2100 h, with 65% (181/278) of measurements being taken in the morning (0700 h–1200 h), 28% (77/278) in the afternoon (1300 h–1700 h) and 7% (20/278) in the evening (1800 h-2100 h, Figure A2). The number of weeks since start of treatment varied from one to 52, although 62% (174/278) of datapoints were in the first 8 weeks of treatment and only 4% (11/278) were taken more than 25 weeks after start of treatment. The times of day at which IOP measurements were most commonly reported were 0800 h (23% of datapoints), 1000 h (19%) and 1600 h (16%), whereas the follow-up points most commonly reported were 0 weeks (31% of datapoints), 2 weeks (17%), 4 weeks (11%) and 2 months (9%) after start of treatment. However, each specific combination of time of day and weeks since start of treatment was reported relatively rarely: for example, although the most common post-baseline timepoint was 0800 h, 2 weeks after start of treatment, only 15 out of 50 trials reported mean IOP at this time (Figure A2).

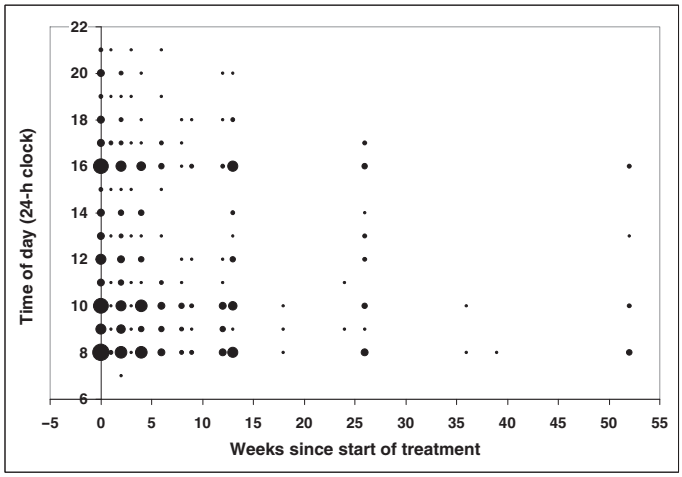

Figure A2. Bubble-plot illustrating the number of trials reporting data at each timepoint. The area of each bubble is proportional to the number of trials included in the meta-analysis that reported data at that timepoint.

Table AII. Summary of studies included in the meta-analysis.

Study name	Sample size	Study duration (weeks)	Treatment Arm 1	Arm 2	Arm 3	Arm 4	Arm 5	Total no. timepoints	No. post-baseline timepoints
Alagoz 2008 [59]	90	26	Bimatoprost 0.03% od	Travoprost 0.04% od	-	-	-	4	3
Arcieri 2008 [60]	34	20	Bimatoprost 0.03% od	Latanoprost 0.005% od	Travoprost 0.004% od	-	-	2	1
Barnebey 2005 [61]	258	13	Timolol 0.5% bd	Travoprost 0.004% od	Travoprost 0.004% od + timolol 0.5% od	-	-	12	9
Bartlett 1999 [62]	100	12	Carteolol 1% bd	Timolol 0.5% bd	-	-	-	2	1
Bron 1991 [63]	48	2	MK-927 2% bd	Placebo	Sezolamide 1.8% bd	-	-	14	11
Camras 2005 [64]	303	26	Brimonidine 0.2% bd	Latanoprost 0.005% od	-	-	-	8	4
Chen 2003 [65]	40	4	Brimonidine 0.2% bd	Timolol 0.5% bd	-	-	-	5	5
Coleman 2003 [66]	177	13	Bimatoprost 0.03% od	Dorzolamide 2% bd + timolol 0.5% bd	-	-	-	12	8
Denis 2006 [67]	83	6	Travoprost 0.004% od + timolol 0.5% od	Travoprost 0.004% od + timolol 0.5% od	-	-	-	9	6
DuBiner 2001 [68]	127	12	Brimonidine 0.2% bd	Latanoprost 0.005% od	-	-	-	3	2
EGPS Group 2005 [69]	1081	260	Dorzolamide 2% tds	Placebo	-	-	-	3	10
Garcia-Sanchez 2004 [70]	344	26	Brimonidine 0.2% bd + timolol 0.5% bd	Latanoprost 0.005% od + timolol 0.5% od	-	-	-	6	3
Gross 2008 [71]	109	2	Travoprost 0.004% od	Travoprost 0.004% od	-	-	-	4	2
Hughes 2005 [72]	316	12	Travoprost 0.004% od + timolol 0.5% od	Travoprost 0.004% od + timolol 0.5% od	-	-	-	12	9
Jampel 2002 [73]	165	8	Latanoprost 0.005% od	Unoprostone 0.15% bd	-	-	-	6	3
Kampik 2002 [74]	379	26	Brimonidine 0.2% bd	Latanoprost 0.005% od	-	-	-	2	2
Kitazawa 1994 [75]	113	2	Dorzolamide 2% tds	Dorzolamide 0.2% tds	Dorzolamide 0.5% tds	Dorzolamide 1% tds	-	5	5
Lippa 1991 [76]	76	2	MK-927 2% bd	MK-927 0.5% bd	MK-927 1% bd	Placebo	-	16	8
Lippa 1992 [77]	73	2	Dorzolamide 0.7% bd	Dorzolamide 1.4% bd	Dorzolamide 2% bd	Placebo	-	7	7
Liu 1999 [78]	26	2	Latanoprost 0.005% od	Placebo	-	-	-	3	3
Martinez-de-la-casa 2004 [79]	44	13	Latanoprost 0.005% od + timolol 0.5% od	Brinzolamide 0.1% bd + travoprost 0.004% od	-	-	-	15	12
Maul 2007 [80]	302	6	Latanoprost 0.005% od	Travoprost 0.004% od	-	-	-	5	4
Mishima 1996 [81]	184	12	Latanoprost 0.005% od	Timolol 0.5% bd	-	-	-	5	4
Nagasubramanian 1993 [82]	69	13	Apraclonidine 0.25% tds	Apraclonidine 0.5% tds	Timolol 0.5% bd	-	-	7	5
Netland 2001 [83]	801	52	Latanoprost 0.005% od	Timolol 0.5% bd	Travoprost 0.0015% od	Travoprost 0.004% od	-	21	18

Table AII. *Continued*

Study name	Sample size	Study duration (weeks)	Treatment Arm 1	Arm 2	Arm 3	Arm 4	Arm 5	Total no. timepoints	No. post-baseline timepoints
Netland 2003 [84]	28	12	Bimatoprost 0.03% od + brimonidine 0.15% bd	Latanoprost 0.005% od + timolol 0.5% od	-	-	-	4	3
Noecker 2006 [85]	94	12	Bimatoprost 0.03% od	Travoprost 0.004% od	-	-	-	4	3
Parrish 2003 [86]	411	12	Bimatoprost 0.03% od	Latanoprost 0.005% od	Travoprost 0.004% od	-	-	8	4
Sacca 1998 [87]	32	13	Betaxolol 0.25% bd	Betaxolol 0.5% bd	-	-	-	21	14
Sall 2000 [88]	463	12	Brinzolamide 1% bd	Brinzolamide 1% tds	Dorzolamide 2% tds	Placebo	-	11	8
Schenker 2000 [89]	241	52	Timolol 0.5% od	Timolol 0.5% od	-	-	-	6	5
Schuman 2005 [90]	403	13	Timolol 0.5% bd	Travoprost 0.004% od + timolol 0.5% od	Travoprost 0.004% od + timolol 0.5% od	-	-	12	9
Serle 1992 [91]	36	6	MK-927 2% bd	Timolol 0.5% bd	-	-	-	24	16
Shedden 2001 [92]	286	24	Timolol 0.5% bd	Timolol 0.5% od	-	-	-	12	10
Shin 2004 [93]	253	13	Dorzolamide 2% bd + timolol 0.5% bd	Latanoprost 0.005% od + timolol 0.5% od	-	-	-	6	3
Silver 1998 [94]	574	13	Brinzolamide 1% bd	Brinzolamide 1% tds	Dorzolamide 2% tds	Timolol 0.5% bd	-	11	8
Stewart 1996 [95]	230	13	Apraclonidine 0.25% bd	Apraclonidine 0.5% bd	Timolol 0.5% bd	-	-	7	5
Stewart 1998a [96]	43	12	Timolol 0.5% od	Timolol 0.5% od	-	-	-	2	1
Stewart 1998b [97]	36	4	Timolol 0.5% bd	Unoprostone 0.12% bd	-	-	-	2	3
Strahlman 1995 [98]	523	52	Betaxolol 0.5% bd	Dorzolamide 2% tds	Timolol 0.5% bd	-	-	15	12
Strahlman 1996 [99]	333	6	Dorzolamide 0.2% tds	Dorzolamide 0.7% tds	Dorzolamide 2% tds	Placebo	-	8	6
Strahlman Pilocarpine Comparison Study [100]	261	28	Dorzolamide 0.7% bd + timolol 0.5% bd	Dorzolamide 2% bd + timolol 0.5% bd	Pilocarpine 2% tds + timolol 0.5% bd	Timolol 0.5% bd	Timolol 0.5% bd	6	6
Strahlman pilot study [100]	32	1	Dorzolamide 2% bd + timolol 0.5% bd	Timolol 0.5% bd	-	-	-	14	7
Susanna 2001 [101]	108	8	Latanoprost 0.005% od	Unoprostone 0.12% bd	-	-	-	6	3
Walters 2004 [102]	115	4	Bimatoprost 0.03% od	Latanoprost 0.005% od	Timolol 0.5% od	-	-	5	3
Wang 2004 [103]	50	6	Brinzolamide 1% bd	Timolol 0.5% bd	-	-	-	1	1
Watson 1996 [104]	294	26	Latanoprost 0.005% od	Timolol 0.5% bd	-	-	-	10	7
Weinreb 1990 [105]	352	13	Betaxolol 0.25% bd	Betaxolol 0.5% bd	-	-	-	5	4
Whitcup 2003 [106]	602	12	Bimatoprost 0.03% bd	Bimatoprost 0.03% od	Timolol 0.5% bd	-	-	6	2
Wilkerson 1993 [107]	28	4	Dorzolamide 2% tds	Placebo	-	-	-	12	7

Note: bd, twice daily; BL, baseline; od, once daily; tds, three times daily

The studies meeting I2 criteria evaluated 13 different monotherapies (apaclonidine, betaxol, bimatoprost, brimonidine, brinzoloamide, carteolol, dorzolomide, latanoprost, MK-927, sezolomide, timolol, travoprost, unoprostone), placebo and seven combination therapies in which one of the aforementioned drugs was used alongside timolol, pilocarpine, brimonidine or brinzoloamide (Figure 1). We used the treatments at a range of concentrations, dosing frequencies and formulations. In order to maximise the power of the analysis without 'lumping' together treatments with potentially different efficacy [55], we treated each combination of concentration and dosing frequency as a separate treatment.

Initial analyses considered timolol eye-drops and timolol gel-forming drops to be equivalent and made no distinction between fixed dose combinations (in which both active molecules are included in the same eye drop solution) and unfixed dose combination therapy (in which each active molecule is given in a separate solution). However, we evaluated the effect of controlling for use of gel-forming drops and unfixed dosing as covariates (Section 2.3).

The studies meeting I2 criteria formed a connected evidence network (Figure 1) including 40 treatments; timolol 0.5% bd lay at the centre of the evidence network, having head-to-head comparisons against 22 other treatments. Eighteen studies (128 datapoints) evaluated timolol 0.5% bd and eight studies (60 datapoints) evaluated placebo.

The relationship between time and IOP varied substantially between studies, although some consistent patterns were evident over the course of treatment. In all but two study arms receiving active treatment, mean IOP fell substantially between baseline and week 1, before stabilising around a lower value, with mean IOP measurements taken at the same time of day varying by less than 1 mmHg over all post-baseline measurements in 87% of study arms. Although diurnal variations were generally larger, with the highest and lowest mean IOP levels for each follow-up visit differing by more than 1 mmHg in 66% of study arms, the strength and shape of the relationship between time and IOP varied substantially. IOP levels tended to be highest in the early morning (as reported previously [56–58]), with IOP falling between 0800 h and 1000 h in most studies, although trends after 1000 h differed between studies and study arms.

APPENDIX B: WinBUGS code used in the analysis

Code used to estimate prior for missing standard deviations:

```
model{
for (i in 1:1112){
            sd[i]~dgamma(alpha,beta)     }
mu~dgamma(.001,.001)
beta~dgamma(.001,.001)

alpha<- mu*beta

sd.new[1]~dgamma(alpha,beta)     }
```

Model 2: Unconstrained baseline, random relative effects constant over time

```
model{
for(i in 1:NData){
                for (k in 1:na[study[i]]) {
            sdev1[i,k]<- sdev[i,k]*(1-equals(sdev[i,k],0)) + sdev.imp.cut*equals(sdev[i,k],0)
            se[i,k]<-sdev1[i,k]/sqrt(n[i,k])              #Calculate SE from sd and n
                    py[i,k]<- pow(se[i,k],-2)             #Define precision from se
              y[i,k] ~ dnorm(my[i,k],py[i,k])             #normal likelihood

my[i,k] <- mu[i] + delta[study[i],k]                      # model

                                        dev[i,k]<- (y[i,k]-my[i,k])*(y[i,k]-my[i,k])*py[i,k]
                                        residual[i,k] <- y[i,k]-my[i,k]
                    }
```

```
        mu[i] ~ dnorm(0,.0001)                    # flat priors for baseline
        dev.i[i]<-sum(dev[i,1:na[study[i]]])
}
resdev<- sum(dev.i[])
sdev.imp.cut<-cut(sdev.imp)
sdev.imp~dgamma(11.54,3.37)    #Alpha and beta parameters estimated in the previously mentioned
                               #meta-analysis of SD data

for ( j in 1:NStudy){

    w[ j,1] <-0
        delta[ j,1]<-0
            for (k in 2:na[ j]) {
        delta[ j,k] ~ dnorm(md[ j,k],taud[ j,k])        # trial-specific LOR distributions
        md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k]
        taud[ j,k] <- tau *2*(k-1)/k                    #precision of LOR distributions
        w[ j,k] <- (delta[ j,k] - d[t[ j,k]] + d[t[ j,1]])  #adjustment, multi-arm RCTs
        sw[ j,k] <-sum(w[ j,1:k-1])/(k-1)               # cumulative adjustment for multi-arm trials
    }
}

d[1]<-0
for (k in 2:NTrt){
        d[k] ~ dnorm(0,.0001)            # vague priors for basic parameters
}

sd~dunif(0,20)                           # vague prior for random-effects standard deviation
tau<-1/pow(sd,2)

}
```

Variations to the code for Model 2 used in other models

Model 1: Mean intraocular pressure and random relative effects constant over time

The line

```
my[i,k] <- mu[i] + delta[study[i],k]
```

was replaced with the line

```
my[i,k]<- mu[study[i]] + delta[study[i],k]
```

And the line mu[i] dnorm(0,.0001) was moved into the loop "for (j in 1:NStudy)"

Model 3: Unconstrained baseline, fixed relative effects constant over time

The line

```
delta[ j,k]  dnorm(md[ j,k],taud[ j,k])
```

was replaced with

```
delta[ j,k] <- d[t[ j,k]] - d[t[ j,1]]
```

and the following lines were removed

```
sd dunif(0,20)
tau<-1/pow(sd,2)
```

Model 4: Unconstrained baseline, random relative effects vary over day, being estimated separately for four time bins:

The line

```
my[i,k] <- mu[i] + delta[study[i],k]
```

was replaced with

```
my[i,k] <- mu[i] + delta[i,k]
```

The line

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k]
```

was replaced with

```
md[i,k] <- equals(bin[i],1)*(d1[t[study[i],k]] - d1[t[study[i],1]]) +equals(bin[i],2)*(d2[t[study[i],k]] -
d2[t[study[i],1]]) + equals(bin[i],3)*(d3[t[study[i],k]] - d3[t[study[i],1]]) +
equals(bin[i],4)*(d4[t[study[i],k]] - d4[t[study[i],1]]) + sw[i,k]
```

The line

```
w[ j,k] <- (delta[ j,k] - d[t[ j,k]] + d[t[ j,1]])
```

was replaced with

```
w[i,k] <- (delta[i,k] - (equals(bin[i],1)*(d1[t[study[i],k]] - d1[t[study[i],1]])
+equals(bin[i],2)*(d2[t[study[i],k]] - d2[t[study[i],1]]) + equals(bin[i],3)*(d3[t[study[i],k]] -
d3[t[study[i],1]]) + equals(bin[i],4)*(d4[t[study[i],k]] - d4[t[study[i],1]])))
```

The lines

```
d[1]<-0
for (k in 2:NTrt){
      d[k]   dnorm(0,.0001)     }
```

were replaced with

```
d1[1]<-0
d2[1]<-0
d3[1]<-0
d4[1]<-0
for (k in 2:NTrt){
      d1[k]   dnorm(0,.0001)
      d2[k]   dnorm(0,.0001)
      d3[k]   dnorm(0,.0001)
      d4[k]   dnorm(0,.0001)     }
```

Model 5: Unconstrained baseline, random relative effects constant within treatment class and over time

Treatments were grouped into the 12 classes shown in Table AIII. The lines

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k]
w[ j,k] <- (delta[ j,k] - d[t[ j,k]] + d[t[ j,1]])
```

were replaced with

```
md[ j,k] <- d[class[ j,k]] - d[class[ j,1]] + sw[ j,k]
w[ j,k] <- (delta[ j,k] - d[class[ j,k]] + d[class[ j,1]])
```

And the line

```
for (k in 2:Trt){
```

was replaced with

```
for (k in 2:NClass){
```

Model 6: Unconstrained baseline, random relative effects exchangeable within treatment class and constant over time

The lines

```
d[1]<-0
for (k in 2:NTrt){
      d[k]   dnorm(0,.0001)                    # vague priors for basic parameters
}
```

were replaced with

```
d[1]<-0
for (k in 2:NTrt){
      d[k]   dnorm(MeanClassd[Class[k]],tauClassd[Class[k]])     }
for (z in 1:7){
      MeanClassd[z]   dnorm(0,.0001)           # vague priors for basic parameters
      SDClassd[z]<-SDallClass
```

Table AIII. Treatments were categorised into the following classes:

Treatment	Class
Betaxolol Carteolol Timolol	Beta-blockers (BB)
Apraclonidine Brimonidine	A2-adrenergics/sympathomimetics (A2A)
Brinzolamide Dorzolamide MK-927 Sezolamide	Carbonic anhydrase inhibitors (CAI)
Bimatoprost Latanoprost Travoprost	Licensed prostaglandins
Unoprostone	Docosanoid prostaglandins
Bimatoprost + timolol Latanoprost + timolol Travoprost + timolol	Prostaglandin + BB
Dorzolamide + timolol	CAI + BB
Brimonidine + timolol	A2A + BB
Placebo	Placebo
Pilocarpine + timolol	Pilocarpine + BB
Bimatoprost + brimonidine	Prostaglandin + A2A
Brinzolamide + travoprost	CAI + prostaglandin

```
        tauClassd[z]<-1/pow(SDClassd[z],2)     }
  for (z in 8:12){
        MeanClassd[z]   dnorm(0,.0001)                  # vague priors for basic parameters
        tauClassd[z]<-1000 }
  SDallClass dunif(0,10)                                  # vague prior for within-class SD
```

We inserted additional code to calculate weighted mean differences between classes:

```
  for (c in 1:(NClass-1))        {
        for (k in (c+1):NClass) {
        wmClassd[c,k] <- MeanClassd[k] - MeanClassd[c]              }              }
```

We categorised treatments into 12 classes using the same classification as the constant class effect model, with the exception of the baseline treatment (timolol 0.5% bd), which we excluded from the beta-blocker class and assigned a fixed treatment effect of zero against which all other treatment effects were calculated. We ordered classes such that classes 1–7 included two or more treatments or doses, whereas classes 8–12 each included only one treatment at a single dose.

Models including arm-level covariates:

The line

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k]
```

was replaced with one of the following:

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + betaUnfixed * (unfixed[ j,k] - unfixed[ j,1])
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + betaDaytime * (daytime[ j,k] - daytime[ j,1])
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + betaGeldrop * (geldrop[ j,k] - geldrop[ j,1])
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + betaDiffBLIOP * (BLIOParm[ j,k]-BLIOParm[ j,1])
```

and one of the following priors was added at the end of the code:

```
betaUnfixed dnorm(0,.0001)
```

```
betaGeldrop dnorm(0,.0001)
betaDaytime dnorm(0,.0001)
betaDiffBLIOP dnorm(0,.0001)
```

Models including arm-level covariates:
The line

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k]
```

was replaced with one of the following:

```
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + (betaBLIOPk[t[ j,k]] - betaBLIOPk[t[ j,1]]) *
    (BLIOP[ j]-24.72705)
md[ j,k] <- d[t[ j,k]] - d[t[ j,1]] + sw[ j,k] + (betaPubDatek[t[ j,k]] - betaPubDatek[t[ j,1]]) *
    (PubDate[ j]-2000.16)
```

And the lines

```
d[1]<-0
for (k in 2:NTrt){
    d[k]   dnorm(0,.0001)                    # vague priors for basic parameters
}
```

were replaced with

```
d[1]<-0
betaBLIOPk[1] <-0                  OR          betaPubDatek[1]<-0
for (k in 2:NTrt){
    d[k]   dnorm(0,.0001)                    # vague priors for basic parameters
    betaBLIOPk[k]<-betaBLIOP       OR          betaPubDatek[k]<-betaPubDate
    }
betaPubDate dnorm(0,.0001)         OR          betaBLIOP dnorm(0,.0001)
```

Sensitivity analysis with multiplicative treatment effects:
The line `my[i,k] <- mu[i] + delta[study[i],k]`
was replaced with `log(my[i,k]) <- mu[i] + delta[study[i],k]`
Additional lines were also added to calculate exp(delta[study[i],k]) and exp(d[k]).

Appendix C: Additional results

Table AIV. Goodness of fit measures for variants of model 2 that included covariates.

Model	pD	Residual deviance*	DIC	Meta-regression coefficient (95% CrI)
2a: Controlling for unfixed dosing of combination therapy (dummy equalling 1 for unfixed combinations and 0 for monotherapy or fixed-dose combinations)	380	703	1083	−0.48 (−1.42, 0.47)†
2b: Controlling for gel-forming solution (dummy equalling 1 for gel-forming solution and 0 for standard eye drops)	382	704	1086	0.11 (−2.17, 2.37)†
2c: Controlling for non-evening dosing of prostaglandins (dummy equalling 1 for prostaglandins administered before 6pm and 0 otherwise)	380	703	1084	−0.49 (−1.6, 0.61)†
2d: Controlling for imbalance in baseline IOP between arms (baseline IOP in each arm, mmHg)	381	704	1085	−0.22 (−0.69, 0.26)†
2e: Controlling for mean baseline IOP across all arms (mmHg)	381	704	1085	0.02 (−0.19, 0.23)‡
2f: Controlling for publication year	381	704	1085	−0.01 (−0.15, 0.12‡

Note: CrI, credible interval; DIC, deviance information criterion (DIC=pD + residual deviance); IOP, intraocular pressure; pD, effective number of parameters (estimated using DIC tool in WinBUGS).
* The data set includes 742 datapoints; well-fitting models would therefore be expected to have a residual deviance of 742.
† Meta-regression coefficient indicates the extent to which the mean difference between two study arms increases when the covariate takes the value 1 rather than 0 or for every 1 mmHg difference in baseline IOP between arms.
‡ Meta-regression coefficient indicates the extent to which the mean difference between the treatment used in arm k and treatment 1 (timolol 0.5% bd) increases for every one unit increase in the study level covariate; covariates were centred at the mean value to improve convergence.

Table AV. Results of model 2 (without covariates).

Treatment	Treatment effect (d_t) versus timolol 0.5% bd	Number of trials (no. pts)
Timolol 0.5% bd	N/A	18 (4653)
Placebo	3.592 (2.745, 4.454)*	8 (2080)
Apraclonidine 0.25% bd†	1.628 (0.196, 3.037)*	1 (230)
Apraclonidine 0.25% tds†	1.376 (−0.349, 3.106)	1 (69)
Apraclonidine 0.5% bd†	1.831 (0.399, 3.252)*	1 (230)
Apraclonidine 0.5% tds	0.619 (−1.135, 2.345)	1 (69)
Betaxolol 0.25% bd	2.557 (0.926, 4.196)*	2 (382)
Betaxolol 0.5% bd	1.479 (0.217, 2.773)*	3 (899)
Bimatoprost 0.03% od	−1.959 (−2.642, −1.275)*	7 (1583)
Bimatoprost 0.03% od + brimonidine 0.15% bd†	−2.978 (−5.460, −0.483)*	1 (28)
Bimatoprost 0.03% bd†	−0.374 (−1.657, 0.923)	1 (602)
Brimonidine 0.2% bd	0.469 (−0.384, 1.321)	4 (841)
Brimonidine 0.2% bd + timolol 0.5% bd	−0.513 (−2.691, 1.679)	1 (325)
Brinzolamide 1% tds†	1.519 (0.504, 2.550)*	2 (921)
Brinzolamide 1% bd	1.842 (0.836, 2.832)*	3 (969)
Brinzolamide 0.1% bd + travoprost 0.004% od†	−2.718 (−4.888, −0.562)*	1 (44)
Carteolol 1% bd†	1.088 (−0.851, 3.005)	1 (100)
Dorzolamide 1% tds†	1.371 (−0.227, 2.983)	1 (113)
Dorzolamide 2% bd†	−1.008 (−2.878, 0.856)	1 (72)
Dorzolamide 2% bd + timolol 0.5% bd	−1.162 (−2.096, −0.256)*	4 (719)
Dorzolamide 2% tds	1.151 (0.340, 1.964)*	7 (3,000)
Dorzolamide 0.2% tds†	1.702 (0.466, 2.968)*	2 (443)
Dorzolamide 0.7% bd†	−2.151 (−3.985, −0.334)*	1 (72)
Dorzolamide 0.7% bd + timolol 0.5% bd†	−1.257 (−2.647, 0.103)	1 (257)
Dorzolamide 0.5% tds†	0.565 (−1.061, 2.170)	1 (113)
Dorzolamide 0.7% tds†	0.448 (−1.001, 1.900)	1 (330)
Dorzolamide 1.4% bd†	−1.321 (−3.140, 0.548)	1 (72)
Latanoprost 0.005% od	−1.170 (−1.738, −0.616)*	13 (3270)
Latanoprost 0.005% od + timolol 0.5% od	−1.803 (−3.472, −0.129)*	4 (650)
Pilocarpine 2% tds + timolol 0.5% bd†	−1.377 (−2.880, 0.109)	1 (257)
Sezolamide 1.8% bd†	0.155 (−1.570, 1.863)	1 (48)
Timolol 0.5% od	0.456 (−0.551, 1.457)	4 (686)
Travoprost 0.0015% od†	−1.384 (−2.563, −0.221)*	1 (787)
Travoprost 0.004% od	−1.522 (−2.192, −0.878)*	8 (2,141)
Travoprost 0.004% od + timolol 0.5% od	−2.041 (−2.917, −1.175)*	4 (1021)
Unoprostone 0.15% bd†	1.997 (0.384, 3.571)*	1 (164)
Unoprostone 0.12% bd†	2.318 (0.953, 3.693)*	2 (140)
MK-927 2% bd†	1.565 (0.483, 2.614)*	3 (160)
MK-927 1% bd†	0.945 (−0.614, 2.519)	1 (76)
MK-927 0.5% bd†	2.027 (0.305, 3.739)*	1 (76)

Note: * $p<0.05$

† Regimen that has been evaluated in one or more trials, but which is not licensed or is unavailable in the UK.

Acknowledgements

The authors would like to thank all those involved in the extraction and validation of this large dataset: particularly Stephen Mitchell, Carmelina Guarnieri and Baris Deniz. We would also like to thank Andrew McNaught and Dina Christopoulou for advice on clinical interpretation and appropriateness of the modelling assumptions and Maria Alva for her involvement in the discussions on modelling approach.

This study was funded by Pfizer Ltd.

References

1. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; **331**:897–900. DOI: 331/7521/897 [pii]10.1136/bmj.331.7521.897.

2. Ades AE, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, Lu G. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *PharmacoEconomics* 2006; **24**:1–19. DOI: 10.2165/00019053-200624010-00001.
3. Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine* 1996; **15**:2733–2749. DOI: 10.1002/(SICI)1097-0258(19961230)15:24<2733::AID-SIM562>3.0.CO;2-0.
4. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value in Health* 2008; **11**:956–964. DOI: 10.1111/j.1524-4733.2008.00347.x.
5. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004; **23**:3105–3124. DOI: 10.1002/sim.1875.
6. Hasselblad V. Meta-analysis of multitreatment studies. *Medical Decision Making* 1998; **18**:37–43.
7. Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 2002; **21**:2313–2324. DOI: 10.1002/sim.1201.
8. Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *American Journal of Epidemiology* 2009; **169**:1158–1165. DOI: kwp014 [pii] 10.1093/aje/kwp014.
9. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *Journal of Clinical Epidemiology* 2009; **62**:857–864. DOI: S0895-4356(08)00276-X [pii] 10.1016/j.jclinepi.2008.10.001.
10. Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006; **166**:1269–1275. DOI: 166/12/1269 [pii] 10.1001/archinte.166.12.1269.
11. National Collaborating Centre for Acute Care. Glaucoma: Diagnosis and management of chronic open angle glaucoma and ocular hypertension, April 2009. http://www.nice.org.uk/nicemedia/pdf/CG85FullGuideline.pdf, Accessed: [28 September 2010].
12. Peters JL, Mengersen KL. Meta-analysis of repeated measures study designs. *Journal of Evaluation in Clinical Practice* 2008; **14**:941–950. DOI: JEP1010 [pii] 10.1111/j.1365-2753.2008.01010.x.
13. Ishak KJ, Platt RW, Joseph L, Hanley JA, Caro JJ. Meta-analysis of longitudinal studies. *Clinical Trials* 2007; **4**:525–539. DOI: 4/5/525 [pii] 10.1177/1740774507083567.
14. Li N, Chen XM, Zhou Y, Wei ML, Yao X. Travoprost compared with other prostaglandin analogues or timolol in patients with open-angle glaucoma or ocular hypertension: meta-analysis of randomized controlled trials. *Clinical & Experimental Ophthalmology* 2006; **34**:755–764. DOI: CEO1237 [pii] 10.1111/j.1442-9071.2006.01237.x.
15. Stewart WC, Pfeiffer N, Mathis HM. Meta-analysis of articles evaluating routine intraocular pressure control in monotherapy in the United States and Germany. *European Journal of Ophthalmology* 2009; **19**:769–775. DOI: 6D1BCFE8-5BA3-4C79-B6A2-F53F6E7BC2C1[pii].
16. Stewart WC, Konstas AG, Nelson LA, Kruft B. Meta-analysis of 24-hour intraocular pressure studies evaluating the efficacy of glaucoma medicines. *Ophthalmology* 2008; **115**:1117–1122.e1. DOI: S0161-6420(07)01076-7 [pii] 10.1016/j.ophtha.2007.10.004.
17. Orme M, Collins S, Dakin H, Kelly S, Loftus J. Mixed treatment comparison and meta-regression of the efficacy and safety of prostaglandin analogues and comparators for primary open-angle glaucoma and ocular hypertension. *Current Medical Research and Opinion* 2010; **26**:511–528. DOI: 10.1185/03007990903498786.
18. van der Valk R, Webers C, Lumley T, Hendrikse F, Prins MH, Schouten J. A network meta-analysis combined direct and indirect comparisons between glaucoma drugs to rank effectiveness in lowering intraocular pressure. *Journal of Clinical Epidemiology* 2009; **62**:1279–1283. DOI: S0895-4356(09)00176-0 [pii] 10.1016/j.jclinepi.2008.04.012.
19. van der Valk R, Webers CA, Schouten JS, Zeegers MP, Hendrikse F, Prins MH. Intraocular pressure-lowering effects of all commonly used glaucoma drugs: a meta-analysis of randomized clinical trials. *Ophthalmology* 2005; **112**:1177–1185. DOI: S0161-6420(05)00286-1 [pii] 10.1016/j.ophtha.2005.01.042.
20. Aptel F, Cucherat M, Denis P. Efficacy and tolerability of prostaglandin analogs: a meta-analysis of randomized controlled clinical trials. *Journal of Glaucoma* 2008; **17**:667–673. DOI: 10.1097/IJG.0b013e318166655700061198-200812000-00011 [pii].
21. Denis P, Lafuma A, Khoshnood B, Mimaud V, Berdeaux G. A meta-analysis of topical prostaglandin analogues intra-ocular pressure lowering in glaucoma therapy. *Current Medical Research and Opinion* 2007; **23**:601–608. DOI: 10.1185/030079907X178720.
22. Cheng JW, Cai JP, Li Y, Wei RL. A meta-analysis of topical prostaglandin analogs in the treatment of chronic angle-closure glaucoma. *Journal of Glaucoma* 2009; **18**:652–657. DOI: 10.1097/IJG.0b013e31819c49d400061198-200912000-00003 [pii].
23. Webers CA, Beckers HJ, Zeegers MP, Nuijts RM, Hendrikse F, Schouten JS. The Intraocular Pressure-Lowering Effect of Prostaglandin Analogs Combined with Topical beta-Blocker Therapy A Systematic Review and Meta-analysis. *Ophthalmology* 2010; **117**:2067–2074. DOI: S0161-6420(10)00271-X [pii] 10.1016/j.ophtha.2010.03.024.
24. Webers CA, van der Valk R, Schouten JSAG, Zeegers MP, Prins MH, Hendrikse F. Intraocular pressure-lowering effect of adding dorzolamide or latanoprost to timolol: a meta-analysis of randomized clinical trials. *Ophthalmology* 2007; **114**:40–46. DOI: S0161-6420(06)00999-7 [pii] 10.1016/j.ophtha.2006.07.018.
25. Eyawo O, Nachega J, Lefebvre P, Meyer D, Rachlis B, Lee CW, Kelly S, Mills E. Efficacy and safety of prostaglandin analogues in patients with predominantly primary open-angle glaucoma or ocular hypertension: a meta-analysis. *Clinical Ophthalmology* 2009; **3**:447–456.
26. Honrubia F, Garcia-Sanchez J, Polo V, de la Casa JM, Soto J. Conjunctival hyperaemia with the use of latanoprost versus other prostaglandin analogues in patients with ocular hypertension or glaucoma: a meta-analysis of randomised clinical trials. *The British Journal of Ophthalmology* 2009; **93**:316–321. DOI: bjo.2007.135111 [pii] 10.1136/bjo.2007.135111.
27. Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Stat Med* 2007; **26**:3681–3699. DOI: 10.1002/sim.2831.

28. Wandel S, Jüni P, Tendal B, Nüesch E, Villiger PM, Welton NJ, Reichenbach S, Trelle S. Effects of glucosamine, chondroitin,or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. *BMJ* 2010; **341**:c4675.
29. Read SA, Collins MJ, Iskander DR. Diurnal variation of axial length, intraocular pressure, and anterior eye biometrics. *Investigative Ophthalmology & Visual Science* 2008; **49**:2911–2918. DOI: iovs.08-1833 [pii] 10.1167/iovs.08-1833.
30. David R, Zangwill L, Briscoe D, Dagan M, Yagev R, Yassur Y. Diurnal intraocular pressure variations: an analysis of 690 diurnal curves. *British Journal of Ophthalmology* 1992; **76**:280–283.
31. Liu J, Zhang X, Kripke DF, Weinreb RN. Twenty-four-hour intraocular pressure pattern associated with early glaucomatous changes. *Investigative Ophthalmology & Visual Science* 2003; **44**:1586–1590.
32. Noel C, Kabo AM, Romanet JP, Montmayeur A, Buguet A. Twenty-four-hour time course of intraocular pressure in healthy and glaucomatous Africans: relation to sleep patterns. *Ophthalmology* 2001; **108**:139–144. DOI: S0161-6420(00)00411-5 [pii].
33. Wilensky JT. Diurnal variations in intraocular pressure. *Transactions of the American Ophthalmological Society* 1991; **89**:757–790.
34. Ades AE, Welton N, Lu G. Introduction to Mixed Treatment Comparisons 2007. https://www.bris.ac.uk/cobm/docs/intro%20to%20mtc.doc, Accessed: [28 September 2010].
35. Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: the confidence profile method. *Medical Decision Making* 1990; **10**:15–23.
36. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 2006; **101**:447–459.
37. Lu G, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009; **10**:792–805. DOI: kxp032 [pii] 10.1093/biostatistics/kxp032.
38. Bhattacherjee P, Paterson CA, Percicot C. Studies on receptor binding and signal transduction pathways of unoprostone isopropyl. *Journal of Ocular Pharmacology and Therapeutics* 2001; **17**:433–441. DOI: 10.1089/108076801753266811.
39. Susanna R, Giampani J, Borges AS, Vessani RM, Jordao ML. A double-masked, randomized clinical trial comparing latanoprost with unoprostone in patients with open-angle glaucoma or ocular hypertension. *Ophthalmology* 2001; **108**:259–263. DOI: S0161-6420(00)00495-4 [pii].
40. Alcon Laboratories (UK) Ltd. Summary of product characteristics for Travatan, 2007. http://www.medicines.org.uk/EMC/medicine/8031/SPC/TRAVATAN/, Accessed:[13th July 2010].
41. Allergan Ltd. Summary of product characteristics for Lumigan 0.1 mg/ml, 2010. http://www.medicines.org.uk/EMC/medicine/22807/SPC/Lumigan+0.1mg+ml/, Accessed: [13th July 2010].
42. Pfizer Ltd. Summary of product characteristics for Xalatan 0.005% eye drops solution, 2010. http://www.medicines.org.uk/EMC/medicine/9043/SPC/Xalatan+0.005%25+eye+drops+solution/, Accessed: [13th July 2010].
43. Bayesian inference using Gibbs Sampling (BUGS). WinBUGS with DoodleBUGS version 1.4.7. 2003-7. http://www.mrc-bsu.cam.ac.uk/bugs/, Accessed: [20th August 2010].
44. Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* 1998; **7**:455.
45. McCullloch P, Nelder JA. *Generalized Linear Models*, (2nd edn). Chapman and Hall: London, 1989.
46. Dempster AP. The direct use of likelihood for significance testing. *Statistics and Computing* 1997; **7**:247–252.
47. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B* 2002; **64**:583–639.
48. Orzalesi N, Rossetti L, Invernizzi T, Bottoli A, Autelitano A. Effect of timolol, latanoprost, and dorzolamide on circadian IOP in glaucoma or ocular hypertension. *Investigative Ophthalmology & Visual Science* 2000; **41**:2566–2573.
49. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, Bradburn M, Eastwood AJ. Indirect comparisons of competing interventions. *Health Technology Assessment* 2005; **9**(26). DOI: 98-51-99 [pii]. 148 pp.
50. Stewart WC, Konstas AG, Kruft B, Mathis HM, Stewart JA. Meta-analysis of 24-h intraocular pressure fluctuation studies and the efficacy of glaucoma medicines. *Journal of Ocular Pharmacology and Therapeutics* 2010; **26**:175–180. DOI: 10.1089/jop.2009.0124.
51. Ishida K, Yamamoto T, Kitazawa Y. Clinical factors associated with progression of normal-tension glaucoma. *Journal of Glaucoma* 1998; **7**:372–377.
52. Ma J, Liu W, Hunter A, Zhang W. Performing meta-analysis with incomplete statistical information in clinical trials. *BMC Medical Research Methodology* 2008; **8**:56. DOI: 1471-2288-8-56 [pii] 10.1186/1471-2288-8-56.
53. Dominici F, Parmigiani G, Wolpert RL, Hasselblad V. Meta-analysis of migraine headache treatments: combining information from heterogeneous designs. *Journal of the American Statistical Association* 1999; **94**:16–28.
54. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions, September 2009. http://www.cochrane-handbook.org/, Accessed: [23rd August 2010]. Version 5.0.2.
55. Caldwell DM, Gibb DM, Ades AE. Validity of indirect comparisons in meta-analysis. *Lancet* 2007; **369**:270. author reply 271.
56. Read SA, Collins MJ, Iskander DR. Diurnal variation of axial length, intraocular pressure, and anterior eye biometrics. *Investigative Ophthalmology & Visual Science* 2008; **49**:2911–2918.
57. David R, Zangwill L, Briscoe D, Dagan M, Yagev R, Yassur Y. Diurnal intraocular pressure variations: an analysis of 690 diurnal curves. *The British Journal of Ophthalmology* 1992; **76**:280–283.
58. Liu JH, Zhang X, Kripke DF, Weinreb RN. Twenty-four-hour intraocular pressure pattern associated with early glaucomatous changes. *Investigative Ophthalmology & Visual Science* 2003; **44**:1586–1590.
59. Alagoz G, Gurel K, Bayer A, Serin D, Celebi S, Kukner S. A comparative study of bimatoprost and travoprost: effect on intraocular pressure and ocular circulation in newly diagnosed glaucoma patients. *Ophthalmologica* 2008; **222**:88–95.

60. Arcieri ES, Pierre Filho PT, Wakamatsu TH, Costa VP. The effects of prostaglandin analogues on the blood aqueous barrier and corneal thickness of phakic patients with primary open-angle glaucoma and ocular hypertension. *Eye* 2008; **22**:179–183.
61. Barnebey HS, Orengo Nania S, Flowers BE, Samples J, Mallick S, Landry TA, Bergamini MV. The safety and efficacy of travoprost 0.004%/timolol 0.5% fixed combination ophthalmic solution. *American Journal of Ophthalmology* 2005; **140**:1–7.
62. Bartlett JD, Olivier M, Richardson T, Whitaker R, Pensyl D, Wilson MR. Central nervous system and plasma lipid profiles associated with carteolol and timolol in postmenopausal black women. *Journal of Glaucoma* 1999; **8**:388–395.
63. Bron A, Lippa EA, Gunning F, Bénichou C, Lesure P, Sirbat D, Royer J, Flament J, Clineschmidt C, Panebianco D. Multiple-dose efficacy comparison of the two topical carbonic anhydrase inhibitors sezolamide and MK-927. *Archives of Ophthalmology* 1991; **109**:50–53.
64. Camras CB, Sheu WP, United States Latanoprost Brimonidine Study Group. Latanoprost or brimonidine as treatment for elevated intraocular pressure: multicenter trial in the United States. *Journal of Glaucoma* 2005; **14**:161–167.
65. Chen MJ, Chou JC, Hsu WM, Liu JH. The efficacy and safety of brimonidine 0.2% compared with timolol 0.5% in glaucoma: a randomized clinical trial on Taiwanese patients. *Journal of the Chinese Medical Association : JCMA* 2003; **66**:276–281.
66. Coleman AL, Lerner F, Bernstein P, Whitcup SM. A 3-month randomized controlled trial of bimatoprost (LUMIGAN) versus combined timolol and dorzolamide (Cosopt) in patients with glaucoma or ocular hypertension. *Ophthalmology* 2003; **110**:2362–2368.
67. Denis P, Andrew R, Wells D, Friren B. A comparison of morning and evening instillation of a combination travoprost 0.004%/timolol 0.5% ophthalmic solution. *European Journal of Ophthalmology* 2006; **16**:407–415.
68. DuBiner HB, Mroz M, Shapiro AM, Dirks MS, Brimonidine versus Latanoprost Study Group. A comparison of the efficacy and tolerability of brimonidine and latanoprost in adults with open-angle glaucoma or ocular hypertension: a three-month, multicenter, randomized, double-masked, parallel-group trial. *Clinical Therapeutics* 2001; **23**: 1969–1983.
69. Miglior S, Zeyen T, Pfeiffer N, Cunha Vaz J, Torri V, Adamsons I, European Glaucoma Prevention Study Group. Results of the European Glaucoma Prevention Study. *Ophthalmology* 2005; **112**:366–375.
70. Garcia-Sanchez J, Rouland JF, Spiegel D, Pajic B, Cunliffe I, Traverso C, Landry J. A comparison of the fixed combination of latanoprost and timolol with the unfixed combination of brimonidine and timolol in patients with elevated intraocular pressure. A six month, evaluator masked, multicentre study in Europe. *British Journal of Ophthalmology* 2004; **88**:877–883.
71. Gross RL, Peace JH, Smith SE, Walters TR, Dubiner HB, Weiss MJ, Ochsner KI. Duration of IOP reduction with travoprost BAK-free solution. *Journal of Glaucoma* 2008; **17**:217–222.
72. Hughes BA, Bacharach J, Craven ER, Kaback MB, Mallick S, Landry TA, Bergamini MV. A three-month, multicenter, double-masked study of the safety and efficacy of travoprost 0.004%/timolol 0.5% ophthalmic solution compared with travoprost 0.004% ophthalmic solution and timolol 0.5% dosed concomitantly in subjects with open angle glaucoma or ocular hypertension. *Journal of Glaucoma* 2005; **14**:392–399.
73. Jampel HD, Bacharach J, Sheu W-P, Wohl LG, Solish AM, Christie W, Latanoprost/Unoprostone Study Group. Randomized clinical trial of latanoprost and unoprostone in patients with elevated intraocular pressure. *American Journal of Ophthalmology* 2002; **134**:863–871.
74. Kampik A, Arias Puente A, O'Brart DP, Vuori ML, European Latanoprost Study Group. Intraocular pressure-lowering effects of latanoprost and brimonidine therapy in patients with open-angle glaucoma or ocular hypertension: a randomized observer-masked multicenter study. *Journal of Glaucoma* 2002; **11**:90–96.
75. Kitazawa Y, Azuma I, Iwata K, Tsukahara S, Shiose Y, Araie M, Shirato S, Mizogami K, Mishima H, Futa R, Komemushi S. Dorzolamide, a topical carbonic anhydrase inhibitor: A two-week dose- response study in patients with glaucoma or ocular hypertension. *Journal of Glaucoma* 1994; **3**:275–279.
76. Lippa EA, Aasved H, Airaksinen PJ, Alm A, Bertelsen T, Calissendorff B, Dithmer O, Eriksson LO, Gustad L, Hovding G, et al. Multiple-dose, dose-response relationship for the topical carbonic anhydrase inhibitor MK-927. *Archives of Ophthalmology* 1991; **109**:46–49.
77. Lippa EA, Carlson LE, Ehinger B, Eriksson LO, Finnström K, Holmin C, Nilsson SE, Nyman K, Raitta C, Ringvold A. Dose response and duration of action of dorzolamide, a topical carbonic anhydrase inhibitor. *Archives of Ophthalmology* 1992; **110**:495–499.
78. Liu CL, Chou JC, Hsu WM, Liu JH. Effects of latanoprost 50 mug/ml on Chinese patients with primary open-angle glaucoma and ocular hypertension. *Chinese Medical Journal (Taipei)* 1999; **62**:703–709.
79. Martinez de la Casa JM, Castillo A, Garcia Feijoo J, Mendez Hernandez C, Fernandez Vidal A, Garcia Sanchez J. Concomitant administration of travoprost and brinzolamide versus fixed latanoprost/timolol combined therapy: three-month comparison of efficacy and safety. *Current Medical Research and Opinion* 2004; **20**:1333–1339.
80. Maul E, Carrasco FG, Costa VP, Casiraghi JF, Vargas E, Sarmina JS, Mayol R. A 6-week, multicenter, randomized, double-masked, parallel-group study comparing travoprost 0.004% to latanoprost 0.005% followed by 6-week, open-label treatment with travoprost 0.004%. *Clinical Therapeutics* 2007; **29**:1915–1923.
81. Mishima HK, Masuda K, Kitazawa Y, Azuma I, Araie M. A comparison of latanoprost and timolol in primary open-angle glaucoma and ocular hypertension. A 12-week study. *Archives of Ophthalmology* 1996; **114**:929–932.
82. Nagasubramanian S, Hitchings RA, Demailly P, Chuniaud M, Pannarale MR, Pecori Giraldi J, Stodtmeister R, Parsons DG. Comparison of apraclonidine and timolol in chronic open-angle glaucoma. A three-month study. *Ophthalmology* 1993; **100**:1318–1323.
83. Netland PA, Landry T, Sullivan EK, Andrew R, Silver L, Weiner A, Mallick S, Dickerson J, Bergamini MV, Robertson SM, Davis AA, Travoprost Study Group. Travoprost compared with latanoprost and timolol in patients with open-angle glaucoma or ocular hypertension. *American Journal of Ophthalmology* 2001; **132**:472–484.

84. Netland PA, Michael M, Rosner SA, Katzman B, Macy JI. Brimonidine purite and bimatoprost compared with timolol and latanoprost in patients with glaucoma and ocular hypertension. *Advances in Therapy* 2003; **20**:20–30.
85. Noecker RJ, Earl ML, Mundorf TK, Silverstein SM, Phillips MP. Comparing bimatoprost and travoprost in black Americans. *Current Medical Research and Opinion* 2006; **22**:2175–2180.
86. Parrish RK, Palmberg P, Sheu WP, XLT Study Group. A comparison of latanoprost, bimatoprost, and travoprost in patients with elevated intraocular pressure: a 12-week, randomized, masked-evaluator multicenter study. *American Journal of Ophthalmology* 2003; **135**:688–703.
87. Sacca SC, Macri A, Rolando M, Ciurlo G. Effect of betaxolol on primary open-angle glaucoma and normal-tension glaucoma patients. *Journal of Ocular Pharmacology & Therapeutics* 1998; **14**:191–201.
88. Sall K, Assil KK, Beehler CC, Cacioppo L, Caine R, DiGaetano M, Friedlaender M, Friedland B, Greenidge K, Gross RL, Higginbortham E, Horwitz B, Iwach A, Kantor RL, Keates EU, Lopatynsky M, Lowry GM, Morris JB, Samples J, Stevenson OD, Terry SA, Tubbs CB, Walters TR, Weiss MJ. The efficacy and safety of brinzolamide 1% ophthalmic suspension (Azopt((TM))) as a primary therapy in patients with open-angle glaucoma or ocular hypertension. Brinzolamide Primary Therapy Study Group. *Survey of Ophthalmology* 2000; **44**:155–162.
89. Schenker HI, Silver LH. Long-term intraocular pressure-lowering efficacy and safety of timolol maleate gel-forming solution 0.5% compared with Timoptic XE 0.5% in a 12-month study. *American Journal of Ophthalmology* 2000; **130**:145–150.
90. Schuman JS, Katz GJ, Lewis RA, Henry JC, Mallick S, Wells DT, Sullivan EK, Landry TA, Bergamini MV, Robertson SM. Efficacy and safety of a fixed combination of travoprost 0.004%/timolol 0.5% ophthalmic solution once daily for open-angle glaucoma or ocular hypertension. *American Journal of Ophthalmology* 2005; **140**:242–250.
91. Serle JB, Lustgarten J, Lippa EA, Camras CB, Framm L, Payne JE, Deasy D, Podos SM. Six week safety study of 2% MK-927 administered twice daily to ocular hypertensive volunteers. *Journal of Ocular Pharmacology* 1992; **8**:1–9.
92. Shedden A, Laurence J, Tipping R, Timoptic XE 0.5% Study Group. Efficacy and tolerability of timolol maleate ophthalmic gel-forming solution versus timolol ophthalmic solution in adults with open-angle glaucoma or ocular hypertension: a six-month, double-masked, multicenter study. *Clinical Therapeutics* 2001; **23**:440–450.
93. Shin DH, Feldman RM, Sheu WP, Fixed Combination Latanoprost/Timolol Study Group. Efficacy and safety of the fixed combinations latanoprost/timolol versus dorzolamide/timolol in patients with elevated intraocular pressure. *Ophthalmology* 2004; **111**:276–282.
94. Silver LH, Allen R, Camras C, Cohen R, Cook D, Dobler A, Belger HJ, Bron A, Cohn H, Collignon J, Cunha Vaz J, Dascotte JC, Demailly P, Detry M, Friedman R, Haik B, Karp D, Laibovitz R, Levy N, De Waard P, George JL, Gruber, D, Hamard P, Pericoi M, Kestelyn P, Lesieur G, Malecaze F, Matt A, Mundorf T, Nelson E, Shields S, Skov P, Stewart R, Wasserstrom J, Mester U, Nordmann JP, Renard JP, Sangers D, Tassignon MJ, Zeyen T, Ulrich WD, Verdonck C. Clinical efficacy and safety of brinzolamide (azopt(TM)), a new topical carbonic anhydrase inhibitor for ‘ and ocular hypertension. *American Journal of Ophthalmology* 1998; **126**:400–408.
95. Stewart WC, Laibovitz R, Horwitz B, Stewart RH, Ritch R, Kottler M. A 90-day study of the efficacy and side effects of 0.25% and 0.5% apraclonidine versus 0.5% timolol. Apraclonidine primary therapy study group. *Archives of Ophthalmology* 1996; **114**:938–942.
96. Stewart WC, Leland TM, Cate EA, Stewart JA. Efficacy and safety of timolol solution once daily versus timolol gel in treating elevated intraocular pressure. *Journal of Glaucoma* 1998; **7**:402–407.
97. Stewart WC, Stewart JA, Kapik BM. The effects of unoprostone isopropyl 0.12% and timolol maleate 0.5% on diurnal intraocular pressure. *Journal of Glaucoma* 1998; **7**:388–394.
98. Strahlman E, Tipping R, Vogel R. A double-masked, randomized 1-year study comparing dorzolamide (Trusopt), timolol, and betaxolol. International dorzolamide study group. *Archives of Ophthalmology* 1995; **113**:1009–1016.
99. Strahlman E, Tipping R, Vogel R. A six-week dose-response study of the ocular hypotensive effect of dorzolamide with a one-year extension. Dorzolamide Dose-Response Study Group. *American Journal of Ophthalmology* 1996; **122**:183–194.
100. Strahlman ER, Vogel R, Tipping R, Clineschmidt CM. The use of dorzolamide and pilocarpine as adjunctive therapy to timolol in patients with elevated intraocular pressure. The dorzolamide additivity study group. *Ophthalmology* 1996; **103**:1283–1293.
101. Susanna R, Giampani J, Borges AS, Vessani RM, Jordao ML. A double-masked, randomized clinical trial comparing latanoprost with unoprostone in patients with open-angle glaucoma or ocular hypertension. *Ophthalmology* 2001; **108**:259–263.
102. Walters TR, DuBiner HB, Carpenter SP, Khan B, VanDenburgh AM, Bimatoprost Circadian IOPSG. 24-Hour IOP control with once-daily bimatoprost, timolol gel-forming solution, or latanoprost: a 1-month, randomized, comparative clinical trial. *Survey of Ophthalmology* 2004; **49**(2):S26–S35. Suppl 1.
103. Wang TH, Huang JY, Hung PT, Shieh JW, Chen YF. Ocular hypotensive effect and safety of brinzolamide ophthalmic solution in open angle glaucoma patients. *Journal of the Formosan Medical Association = Taiwan yi zhi* 2004; **103**:369–373.
104. Watson P, Stjernschantz J. A six-month, randomized, double-masked study comparing latanoprost with timolol in open-angle glaucoma and ocular hypertension. The latanoprost study group. *Ophthalmology* 1996; **103**:126–137.
105. Weinreb RN, Caldwell DR, Goode SM, Horwitz BL, Laibovitz R, Shrader CE, Stewart RH, Williams AT. A double-masked three-month comparison between 0.25% betaxolol suspension and 0.5% betaxolol ophthalmic solution. *American Journal of Ophthalmology* 1990; **110**:189–192.
106. Whitcup SM, Cantor LB, VanDenburgh AM, Chen K. A randomised, double masked, multicentre clinical trial comparing bimatoprost and timolol for the treatment of glaucoma and ocular hypertension. *The British Journal of Ophthalmology* 2003; **87**:57–62.
107. Wilkerson M, Cyrlin M, Lippa EA, Esposito D, Deasy D, Panebianco D, Fazio R, Yablonski M, Shields MB. Four-week safety and efficacy study of dorzolamide, a novel, active topical carbonic anhydrase inhibitor. *Archives of Ophthalmology* 1993; **111**:1343–1350.