CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Prostaglandin Analogues for Ophthalmic Use: A Review of Comparative Clinical Effectiveness, Cost-Effectiveness, and Guidelines

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	February 18, 2020
Report Length:	46 Pages

CADTH

Authors: Saadul Islam, Carolyn Spry

Cite As: Prostaglandin Analogues for Ophthalmic Use: A Review of Comparative Clinical Effectiveness, Cost-Effectiveness, and Guidelines. Ottawa: CADTH; 2020 February. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

Abbreviations

AD	Area density classification
AE	Adverse event
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
BIM	Bimatoprost
BUT	Tear film break-up time
DBP	Diastolic blood pressure
DPP	Diastolic perfusion pressure
ICER	Incremental cost-effectiveness ratio
IOP	Intraocular pressure
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
HR	Heart rate
LAT	Latanoprost
NMA	Network meta-analysis
NTG	Normal tension glaucoma
OHT	Ocular hypertension
OAG	Open angle glaucoma
OPP	Ocular perfusion pressure
OSDI	Ocular Surface Disease Index
PGA	Prostaglandin analogues
POAG	Primary open angle glaucoma
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PXG	Pseudo-exfoliative glaucoma
RCT	Randomized controlled trial
SBP	Systolic blood pressure
SPP	Systolic perfusion pressure
TAF	Tafluprost
TIM	Timolol
TRA	Travoprost
QALY	Quality adjusted life years
QOL	Quality of life
USD	United States dollar
VBM	Value-Based Medicine

Context and Policy Issues

Glaucoma is an umbrella term that refers to eye diseases involving progressive degeneration of the optic nerve.[1] This may lead to gradual irreversible vision loss and potential blindness if not detected or treated early. Although characterized traditionally by an elevated intraocular pressure (IOP), it is now known that glaucoma involves a characteristic atrophy of the optic nerve head, which may or may not be accompanied by elevated IOP.[2] Nonetheless, an elevated IOP is the most important risk factor for glaucoma.[1] IOP is dependent on secretion of aqueous humour by the ciliary body as well as drainage of aqueous humour from the eye.[1] The normal IOP ranges between 10 mmHg and 20 mmHg with an average value of 15 mmHg.[3] Ocular hypertension (OHT) is characterized by a higher than normal IOP level, in the absence of optic nerve damage or visual field loss. Open angle glaucoma (OAG), sometimes referred to as primary open angle glaucoma (POAG), is the most common form of glaucoma, accounting for more than 70% of glaucoma cases.[4] This is caused by a higher concentration of fluid being produced within the eyes, or when the drainage system is not working properly, resulting in a buildup of pressure on the optic nerve, ultimately damaging the optic nerve. A less common type of glaucoma is closed-angle glaucoma or narrow angle glaucoma, which happens when the drainage angle in the eye (formed by the cornea and the iris) closes or becomes blocked.

This usually occurs in old age, when the lens in the eye becomes larger, pushing the iris forward and narrowing the space between the iris and the cornea, thereby blocking the aqueous fluid from exiting through the drainage system, resulting in a buildup of fluid and an increase in eye pressure.[4] Secondary glaucoma can occur as a result of other conditions (e.g., infection, inflammation, trauma, or pseudoexfoliation), medication usage (e.g., corticosteroids), or ocular surgery. Finally, in normal tension glaucoma (NTG), the optic nerve is damaged without a concomitant increase in eye pressure. The pathophysiology of NTG is still unclear.[4]

There is limited epidemiological data available for glaucoma in Canada. A recently published document by Health Quality Ontario[5] reported more than 400,000 Canadians are affected with glaucoma, with the direct costs of vision loss from glaucoma estimated at $300 million annually. Similar findings were reported from the 2008-2009 Canadian Community Health Survey on Health Aging which estimated that 456,533 Canadians had a diagnosis of glaucoma.[6]

Treatments for glaucoma primarily involve lowering IOP levels to a normal range. The target IOP should be modified based on the patient's age, quality of life (QoL) and risk factors for progression.[3] Treatment strategies for patients with glaucoma include topical or systemic medications, laser therapy, and surgery, although the latter two are less common. Pharmacologic therapy is the most common method of lowering IOP and there are several types of agents available: prostaglandin analogues (PGAs, alternatively defined as prostamides), beta-blockers, carbonic anhydrase inhibitors, alpha adrenergic agonists, and direct-acting cholinergic agonists.[3] Of these, the most common first-line therapy is with PGAs due to favourable effectiveness, once-daily administration, and tolerability compared with the other agents.[3,7,8] Currently available PGAs include latanoprostene bunod (0.024%), latanoprost (LAT, 0.005%), travoprost (TRA, 0.004%), bimatoprost (BIM, available in two doses, 0.03% and 0.01%), and tafluprost (TAF). Patients who do not meet their target IOP may receive an additional agent, often timolol (TIM, 0.5%), a beta-blocker.[3]

In 2015, CADTH prepared a Rapid Response Summary with Critical Appraisal on the clinical effectiveness and cost-effectiveness of BIM compared with other PGAs for ophthalmic use.[9] The objective of the current report is to evaluate the evidence published as of 2015 on the clinical effectiveness and cost-effectiveness of BIM versus other PGAs for ophthalmic use. Additionally, evidence-based guidelines regarding the use of BIM for elevated intraocular pressure will be reviewed.

Research Questions

1. What is the comparative clinical effectiveness of bimatoprost versus other prostaglandin analogues for ophthalmic use?

2. What is the cost-effectiveness of bimatoprost versus other prostaglandin analogues for ophthalmic use?

3. What are the evidence-based guidelines for the use of bimatoprost for elevated intraocular pressure?

Key Findings

A total of 13 relevant publications were included in this report: five systematic reviews (including four meta-analysis and one network meta-analysis), seven randomized controlled

trials, and one cost-effectiveness study. No evidence-based guidelines were identified from the literature within the last five years regarding the use of bimatoprost in lowering intraocular pressure.

The overall findings from the systematic reviews and clinical trials are seemingly contradictory. Results from the systematic reviews showed bimatoprost be clinically superior over other prostaglandin analogues (including latanoprost, travoprost, and tafluprost) with respect to intraocular pressure; whereas with the exception of a crossover trial, none of the randomized trials with a parallel group design demonstrated the clinical superiority of bimatoprost. Most studies concluded that the clinical profile of the four prostaglandin analogues is similar; with all four prostamides effective in reducing intraocular pressure in patients with open angle glaucoma or ocular hypertension, irrespective of prior treatment status or other risk factors. In terms of adverse events, bimatoprost was found to be relatively less tolerated compared to the others and may result in hyperemia and growth of eyelashes. However, these findings are not consistently reported in all trials, and the safety profile of the four prostamides may be similar. A single cost-effectiveness study from the United States showed bimatoprost to be the most cost-effective prostamide, using robust methodology and different perspectives (societal, third party insurance, and ophthalmic), although its applicability is debatable in the Canadian health care context.

The mixed and inconclusive findings across the studies may be a result of a number of factors: inclusion of studies that are heterogeneous in nature with respect to patient population, intervention (including dosage), assessment of study outcomes, and overall study design; as well as methodological limitations arising from small sample size, and inadequate reporting of data. Overall, bimatoprost appears to be at least as effective as latanoprost, travoprost, and tafluprost, either as monotherapy, or in combination with timolol. However, bimatoprost showed to be the most cost-effective prostamide, which may support its use over other alternatives.

Methods

Literature Search Methods

A limited literature search was conducted by an information specialist on key resources including MEDLINE, Embase, the Cochrane Library, the University of York Centre for Reviews and Dissemination (CRD) databases, the websites of Canadian and major international health technology agencies, as well as a focused Internet search. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concept was bimatoprost. No search filters were applied to limit retrieval by publication type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2015 and January 20, 2020.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Adults with elevated intraocular pressure (e.g., open-angle glaucoma, angle-closure glaucoma, ocular hypertension)
Intervention	Bimatoprost as monotherapy or in combination with timolol
Comparator	• Latanoprost as monotherapy or in combination with timolol • Latanoprostene • Travoprost as monotherapy or in combination with timolol • Tafluprost
Outcomes	Q1: Clinical effectiveness (e.g., change in ocular pressure, duration of effect) and harm (e.g., adverse events, morbidity, mortality) Q2: Cost-effectiveness (e.g., cost per quality adjusted life year, cost per patient adverse event avoided, cost per clinical outcome, cost minimization) Q3: Recommendations regarding its place in therapy and appropriate use
Study Designs	Health Technology Assessments, Systematic Reviews, Randomized Controlled Trials, Economic Evaluations, Evidence-based Guidelines.

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2015. Notably, one systematic review by Lou et al.[10] was excluded despite meeting the selection criteria since this was described in a previously published Rapid Response report on the same topic.[9] Finally, guidelines with unclear methodology were excluded.

Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised by one reviewer using A MeaSurement Tool to Assess systematic Reviews (AMSTAR) II,[11] the network meta-analysis was critically appraised using the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) checklist,[12] the clinical studies were critically appraised using the Downs and Black checklist,[13] and the economic study was critically appraised using the Drummond Checklist.[14] Summary scores were not calculated for the included studies; rather, the strengths and limitations of each included study were described narratively.

Summary of Evidence

Quantity of Research Available

A total of 207 citations were identified in the literature search. Following screening of titles and abstracts, 182 citations were excluded and 25 potentially relevant reports from the electronic search were retrieved for full-text review. In addition, four potentially relevant publications were retrieved from the grey literature search for full-text review. Of these 29 potentially relevant articles, 16 publications were excluded for various reasons, while 13 publications met the inclusion criteria and were included in this report. These comprised five systematic reviews,[15-19] seven randomized controlled trials (RCTs),[20-26] and one cost-effectiveness study.[27] Appendix 1 presents the PRISMA[28] flowchart of the study selection. Additional references of potential interest are provided in Appendix 5.

Summary of Study Characteristics

A total of five relevant systematic reviews (including three with meta-analyses and one with a network meta-analysis),[15-19] seven RCTs,[20-26] and one cost-effectiveness study[27] were identified for inclusion in this review. No relevant health technology assessments or evidence-based guidelines were identified. Detailed study characteristics are available in Appendix 2,Table 2, Table 3, and Table 4.

Study Design

The five included systematic reviews[15-19] had objectives and inclusion criteria that were wider in scope than the current report. The systematic reviews compared a variety of PGAs and topical medical therapies (including beta-blockers, alpha-2 adrenergic agonists and carbonic anhydrase inhibitors) for the treatment of glaucoma and glaucomatous conditions. This report is limited to studies that compared BIM with a relevant PGA. The studies by Tang et al. (2019),[15] Rennie et al. (2019),[16] and Li et al. (2016)[19] only included RCTs, whereas Diaconita et al. (2018)[17] and Takagi et al. (2018)[18] included all study designs. The systematic reviews by Tang et al.[15] and Rennie et al.[16] consisted of 17 and 10 RCTs published up to June and March of 2018, respectively, of which 15 and 1 RCT compared BIM to other PGAs, respectively. Diaconita et al.[17] investigated the washout duration of different PGAs in 56 published studies until October 2016, of which one RCT compared BIM with LAT. Takagi et al.[18] evaluated the IOP-lowering effect of different PGAs in Japanese patients, and included 27 published studies up to February 2018, of which one single arm study (not relevant to this report) and three multi-arm studies included BIM as a treatment. Li et al.[19] published a systematic review with NMA comparing first-line medications for POAG, which consisted of 114 RCTs published until March 2014, of which 13 compared BIM with no treatment/placebo or another topical medication.

The seven RCTs included in this report varied in study design, and all but three were open-label (OL) in nature. Guven Yilmaz et al.[23] and Stalmans et al.[24] were single-blinded trials where study personnel were blinded to patients' treatment assignment; whereas Rosetti et al.[26] had a double-blinded placebo-controlled design. Among other notable design features, Blondeau et al.[20] and Guven Yilmaz et al.[23] were switch trials (switching to BIM from another drug) and Stalmans et al.[24] had a crossover design. The remaining trials had a parallel group comparison design.

The single cost-effectiveness study evaluated patient preference-based comparative effectiveness and cost-utility analyses of PGAs (including BIM) for the treatment of OAG.[27] Briefly, the study used a proprietary method called Value-Based Medicine (VBM), incremental cost-utility analysis (comparing the prostamides among themselves and with TIM), and average cost-utility analysis (comparing each drug with no therapy). VBM is a method that uses standardized, cost-utility analysis inputs and outputs, thereby allowing comparability across all medical fields. With this method, the model utilized standardized interventional benefit in terms of gains in quality of life (QoL) and/or length of life. A 20-year time horizon was applied; this was based on the average life expectancy for a person aged 65 years, as in a glaucoma clinical trial. Three perspectives were chosen for the cost-utility analysis, societal, third party insurer, and direct ophthalmic expenditure cost perspectives. A 3% annual discount rate was applied for all patient value outcomes (QALYs) and costs (all in USD). The following inputs were derived from a number of published studies: time to end-stage glaucoma at different IOPs, direct nonophthalmic medical costs, caregiver costs, and visual field correlation with utilities. Time tradeoff utilities were obtained from the Center for Value-Based Medicine Utility Database consisting of more than 55,000 patient utilities.

Average, diurnal, and IOP data were derived from the systematic review and NMA by Li et al.[19] Adverse events (AEs) were based upon a double-masked RCT of BIM, two meta-analyses, and FDA randomized, double-masked clinical trial files (it was unclear in the publication how these differed from the previously listed double-masked RCT). Drug prices were taken from 2015 Average Wholesale Price data. Medical costs were taken from the 2015 national average Medicare Fee Schedule. Employment costs were taken from Bureau of Labor and Statistics data.

Country of Origin

Authors of systematic reviews were based in China,[15] Australia,[16] Canada,[17] Japan,[18] and USA.[19] The location of the primary studies was not provided in all systematic reviews.

Of the included RCTs, Blondeau et al.,[20] Nazir et al.,[21] Moussa et al.,[22] Guven Yilmaz et al.,[23] and Maruyama et al.[25] were single-centre studies or conducted in a single country, namely, in Canada, Pakistan, Lebanon, Turkey, and Japan, respectively. Stalmans et al.[24] and Rosetti et al.[26] were multi-centre studies conducted in 7 European centres in Central and Western Europe.

The cost-effectiveness study by Brown and Brown[27] was conducted in the US.

Patient Population

All included systematic reviews consisted of adult (≥ 18 years of age) patients with glaucoma, including POAG, OAG, NTG, pseudo-exfoliative glaucoma (PXG), or OHT. As study-level data were not provided in all systematic reviews, the sample size, major demographic characteristics (including age, sex distribution, and treatment history), and trial duration could not be summarized for the relevant primary studies (involving BIM and at least one relevant PGA as comparator).

The RCTs included in this report comprised of both treatment-naive and treatment-experienced patients who were at various stages of disease development. The RCTs by Blondeau et al.,[20] and Moussa et al.[22] were conducted in treatment-naive patients who were newly diagnosed with OAG or OHT. The remaining RCTs included patients with OAG or OHT who had previously received prostaglandin monotherapies or combinational therapies with TIM for a period of four to six weeks but still had high IOP levels (around or exceeding the 21 mmHg cutoff). In addition to POAG, Blondeau et al.,[27] Stalmans et al.,[24] and Rosetti et al.[26] included a number of patients with NTG and secondary OAG (including PXG, pigmentary glaucoma or PG). All trials recruited adult patients only, with mean age ranging from 30 – 32 years in Nazir et al.,[21] to 60 – 68 years in the remaining trials.

The economic evaluation by Brown and Brown[27] was conducted in patients with OAG, using data from studies with at least 3-month follow-up, in which the baseline IOP was 26mm Hg.

Interventions and Comparators

The five included systematic reviews investigated a variety of interventions and comparators regarding the topical management of glaucoma. The Tang et al. study[15] included three PGAs, namely LAT, BIM, and TRA. Takagi et al.[18] evaluated the PGAs available in Japan, i.e. unoprostone (UNO), LAT, TRA, TAF, and BIM. Diaconita et al.[17] investigated the washout duration of PGAs or prostamide or combination drug which includes PGAs or prostamide. Rennie et al.[16] and Li et al.[19] considered common first line topical treatments used in glaucoma, including beta-blockers, PGAs, alpha-2 adrenergic

agonists and carbonic anhydrase inhibitors. Notably, study-level detail on the dose of PGAs (including BIM) was only provided in Tang et al.,[15] therefore, findings for the systematic reviews could not be summarized by PGA dose.

The RCTs meeting the inclusion criteria of this report all compared two or multiple PGAs; however, the choice of intervention, randomization ratio, dosage, and duration varied across the trials. In Blondeau et al,[20] treatment-naive newly diagnosed patients were given LAT for one month, and those who failed to achieve a 20% reduction in IOP were selected to be randomized to continue LAT, or switch to BIM or TRA for a month without any washout. In Nazir et al.,[21] patients received one drop of LAT (0.005%) or BIM (0.01%) daily for 29 days. Patients in Moussa et al.[22] received a single drop of preserved (with benzalkonium chloride, a commonly used preservative in glaucoma preparations) BIM (0.01%), LAT (0.05%), TRA (0.004%), or preservative-free TAF (0.0015%) daily for six months. In Guven Yilmaz et al.,[23] patients on bilateral monotherapy of hypotensive lipids switched to receive fixed combinations of BIM/TIM, LAT/TIM, and TRA/TIM once-daily for an average of four to five weeks. In Stalmans et al.,[24] patients previously treated with a preserved prostaglandin monotherapy were randomized to receive BIM (0.03%) or LAT drops once-daily for three months, then underwent six weeks of washout, followed by three months of treatment with the alternative regimen. In Maruyama et al.,[25] patients previously treated with LAT monotherapy switched to either BIM or LAT/TIM for 12 weeks. Finally, in Rosetti et al.,[26] patients on monotherapy underwent six weeks of washout with unfixed combination of LAT or TIM before being allocated daily drops of BIM or LAT/TIM for 12 weeks. Of the studies with dose information, only Maruyama et al.[25] included BIM 0.03%, the remaining studies used 0.01% concentration of BIM. In most cases, a single eyedrop was administered per day during the evening or at bedtime. All RCTs with the exception of the following three used PGA monotherapy: Guven Yilmaz et al.[23] where all randomized patients received a combination therapy, and Maruyama et al.[25] and Rosetti et al.[26] where BIM monotherapy was compared with LAT/TIM.

The cost-effectiveness study by Brown and Brown[27] compared the four prostamide glaucoma medications available in the US, including topical BIM 0.01%, LAT 0.005%, TRA 0.004%, and TAF 0.0015%.

Outcomes

All five systematic reviews included IOP as the primary efficacy outcome or used IOP to compute the outcome of choice. In the Tang et al.[15] study, the mean IOP reduction from baseline to endpoint was determined. In the Rennie et al. study,[16] the primary outcome of interest was post-intervention mean ocular perfusion pressure (OPP), defined as the pressure forcing blood into the eye against the resistance of the ocular vasculature.[16] There is evidence that low values of diastolic perfusion pressure is a major risk factor for glaucoma, and a decreased perfusion pressure to be a strong predictor of glaucoma progression.[26] Mean OPP data was obtained directly from the included RCTs, or computed using systolic or diastolic OPP and IOP. The primary outcome in Diaconita et al.[17] was the mean pre and post-washout IOP, which was used to compute the mean (and percentage of) IOP reduction. In Takagi et al.,[18] the ocular hypotensive efficacy was based on postdosing IOP reduction extracted directly from eligible articles, or computed from pre-dosing and postdosing final evaluation points. Finally, Li et al.[19] used IOP (measured at 3 months) as the primary outcome and visual field (at any follow-up time point) as the secondary outcome.

Two systematic reviews, Tang et al.[15] and Rennie et al.[16] assessed adverse events (AEs); which was further specified in Tang et al.[15] as the following: conjunctiva hyperemia, discomfort (itching, eye irritation, foreign body sensation), and growth of lashes.

Similar to the systematic reviews, the RCTs compared the clinical benefits of PGAs primarily on the basis of IOP difference following treatment and/or between treatment groups. IOP was measured in a similar process across the trials, with the use of a (Goldmann) calibrated applanation tonometer. However, the frequency, timing, and the number and blinding of measurement personnel differed. With the exception of Blondeau et al.[20] and Nazir et al.,[21] all trials involved repeated measurements of IOP at pre-specified times through the day to obtain an average at each timepoint. Blondeau et al.[20] and Guven Yilmaz et al.[23] further accounted for diurnal/circadian variation in IOP measurement. Below is a list of efficacy and safety endpoints measured in the trials:

- **IOP:** IOP difference before/during/after treatment, or between treatment groups, and IOP response rate.
- **Blood pressure, perfusion pressure:** Maruyama et al.[25] and Rosetti et al.[26] also evaluated outcomes based on blood pressure, e.g., systolic and diastolic blood pressure (SBP, DBP), heart rate (HR), and perfusion pressure (calculated as SBP/DBP – IOP).
- **Safety and tolerability:** Moussa et al.[22] evaluated PGA tolerability using both objective (clinical exam) and subjective methods (using the Ocular Surface Disease Index (OSDI) score, a scale from 0 to 100 where the score given is directly proportional to the severity of the symptoms). The ocular surface injury is classified as normal (0–12), minimal (13–22), moderate (23–32), and severe (33–100). Other notable safety endpoints in Stalmans et al.[24] and Maruyama et al.[25] involved a complete bilateral and comparative ophthalmologic examination for adverse events (AEs) including conjunctival injection score (grade: 0–3), corneal epitheliopathy score (area density classification; AD score), tear film break-up time (BUT), macroscopic conjunctival hyperemia, superficial punctiform keratitis, follicular conjunctivitis, iris pigmentation using photographic iris color, eyelash growth, and herpes reactivation.

The economic evaluation reported the incremental cost-effectiveness ratio (ICER) using cost per quality-adjusted life-years (QALYs) gained.[27]

Summary of Critical Appraisal

Additional details regarding the strengths and limitations of included publications are provided in Appendix 3, Table 5, Table 6, and Table 7.

Systematic review and NMA

The four included systematic reviews with meta-analysis were critically appraised using AMSTAR;[11] the ISPOR questionnaire[12] was used to appraise the systematic review with NMA.

These publications were conducted and reported in accordance with established guidelines for systematic reviews (PRISMA statement and PRISMA extension statement for NMAs). All five reports had a clearly defined objective and rationale, inclusion and exclusion criteria, and included flow charts illustrating study selection and provided reasons for study

exclusion. Key search terms and search strategies were provided in all reviews, increasing their reproducibility. A comprehensive and thorough literature review was conducted, covering multiple databases, with little or no restrictions placed on publication date. Study selection, data extraction, and study quality assessment were well-documented and generally done in duplicate for all but one report; Takagi et al.,[18] which provided no information on whether one or multiple authors were involved and how disagreements were resolved. In all but one systematic review, the risk of bias and limitations of primary study design were assessed using established tools, including the Cochrane Risk of Bias tool for RCTs, and the Downs and Black checklist; Takagi et al.[18] did not appear to conduct a quality assessment for the included primary studies. Finally, the authors of all five systematic reviews disclosed their sources of funding; only Takagi et al.[18] declared a conflict of interest arising from multiple authors being paid employees of a pharmaceuticals company.

Description of important characteristics of the included studies was provided. However, as the objectives of all systematic reviews were broader than the topic of interest for this Rapid Response report, the study details that were chosen to be reported were often inadequate for comparing study characteristics. All systematic reviews with the exception of Takagi et al.[18] were generally conducted with high scientific rigour; however, the quality of included studies within each systematic review varied. Overall, the risk of bias of the included RCTs in each systematic review was mixed, as reported by systematic review authors, and most primary studies had multiple domains with low, moderate, high or unclear risk of bias. Nonetheless, the quality of the studies and evidence was adequately considered in interpreting results and formulating conclusions.

The statistical methods for pooling results for direct (pairwise meta-analyses) or indirect (NMA) comparisons were well-established and appropriately done. With the exception of Takagi et al.,[18] the remaining three systematic reviews with meta-analysis assessed statistical heterogeneity, using the Cochrane Q test or I^2 statistic and extracted results of studies were weighted. In all cases, a random-effects model was used when $I^2 \geq 50\%$ and the Cochrane Q test P value was < 0.1 indicating moderate to significant heterogeneity, otherwise a fixed-effects model was used. The NMA included a random-effects model for both pairwise direct and indirect comparisons. Assessments of publication bias using Funnel plot were planned in two systematic reviews, Rennie et al.[16] and Diaconita et al.[17] However, it could not be assessed in Rennie et al.[16] due to a small number of included studies (10); whereas Diaconita et al.[17] found no evidence of plot asymmetry.

Randomized controlled trials

Strengths common to all seven RCTs were that the study objective, inclusion and exclusion criteria, methods for patient recruitment, and interventions being compared were described clearly. In addition, the main outcomes, potential confounders, baseline patient characteristics, and main findings were also clearly described, increasing the strength of reporting. Three of the seven included trials were blinded in nature, two single-blinded and one double-blinded RCT. In all cases, masking treatment allocation from the investigators and/or patients was appropriately done. The remaining four trials were open-label; therefore, patients and investigators were not blinded. However, the assessment of study outcomes as well as the diagnosis of patients for eligibility appeared to have been done in an objective manner. Statistical tests conducted in all RCTs were appropriate; the effects of confounders were assessed using appropriate analytical methods or controlled with measurement techniques or procedures. Study participants, care providers or investigators, and health care settings appeared to be representative of the population and care settings

of interest, increasing the external validity of the studies. The trials were conducted in different countries across North America, Europe, and Asia, including one study in Canada. Further, the trials included patients of various demographic characteristics, as well as treatment and disease history; therefore, the findings are generalizable to a large group of glaucoma patients. Finally, with the exception of Stalmans et al.[24] and Maruyama et al.,[25] the authors of the remaining five studies declared no potential conflicts of interest.

As for methodological limitations, the RCTs excluded patients based on factors that may interfere with the results. This included patients with significant comorbidities (including ophthalmic conditions), contraindication or hypersensitivity to any of the study drugs, prior history of ophthalmic surgery, and those who are unlikely to be compliant with the study protocol. While it is a common practice in clinical trials to recruit patients who are otherwise healthy and are free of interfering comorbidities or therapies, the exclusion of patients based on the above criteria does limit the generalizability of the study findings in these patients. However, the generalizability does not appear to be seriously limited. Three of the seven included RCTs did not conduct a power or sample size calculation (Nazir et al.,[21] Moussa et al.,[22] and Guven Yilmaz et al.,[23,20] and Maruyama et al.[25]). Therefore, the adequacy of power in detecting the effect estimate is a concern in these trials when non-significant results were observed. Finally, of the following four RCTs did not provide sufficient or any detail of the randomization procedure for treatment allocation: e.g. Moussa et al.,[22] Guven Yilmaz et al.,[23] Stalmans et al.,[24] and Maruyama et al.,[25] therefore the appropriateness of the randomization procedure as well as its success could not be assessed.

Cost-effectiveness studies

The economic evaluation included in this report[27] was generally conducted well with methodologies suitable for the respective context and decision problem. Notable strengths included a clearly stated research question, perspective, and time horizon; appropriately chosen Value-Based Medicine method (as this method utilizes standardized, cost-utility analysis inputs and outputs allowing comparability across all medical fields); relevant and justified source data for clinical, utilities and cost inputs; valid method for data extrapolation; clearly described method for estimation of medical, non-medical, societal, and health system costs; and a varied range of sensitivity and scenario analyses to assess model assumptions.

The economic evaluation had two notable limitations. The study applied a 3% annual discount rate, which even though is generally acceptable, is higher than the CADTH recommended 1.5%. This indicates that the authors of the study placed less importance on future benefits of the treatment(s) which may reflect the values and social construct in the US.[29] Finally, even though the authors provided adequate justification for choosing Value-Based Medicine for the study, the impact of not using the more commonly used Markov modeling is uncertain.

Summary of Findings

The overall findings of the included studies are highlighted below. Detailed summaries of the main findings are available in Appendix 4, Table 8, Table 9, and Table 10.

Comparative clinical effectiveness of bimatoprost versus other prostaglandin analogues for ophthalmic use

Change in intraocular pressure

Three systematic reviews compared PGAs (including BIM) with respect to changes in IOP from baseline, namely Tang et al.,[15] Takagi et al.,[18] and Li et al.[19] Of these, Tang et al.[15] compared the IOP-lowering effect of PGAs at three different timepoints, 1, 3, and 6 months post-treatment. Based on pooled analyses, BIM (0.03%) showed a greater IOP control compared with LAT (0.005%) and TRA (0.004%); however, the efficacy was demonstrated with longer term use (3 to 6 months). Takagi et al.[18] reported a 15% to 20% reduction in IOP with the use of BIM, LAT, TRA, and TAF; of which BIM consistently showed the strongest hypotensive profile, followed by TRA and TAF, which showed similar IOP-lowering efficacy, and LAT was found weaker compared to all three; however, no statistical testing for the reduction in IOP or the comparisons between PGAs was reported. Li et al.[19] compared the 3-month IOP lowering activity of BIM and other PGAs using direct and indirect estimates. Overall, all PGAs showed similar clinical efficacy; BIM generally showed clinical superiority over other prostamides, although results were inconclusive for some comparisons. These findings were consistent with the direct and indirect estimates comparing the PGAs, where BIM, LAT, and TRA were found to be the most efficacious prostamides in reducing IOP levels. The three PGAs had similar clinical profiles, with small between-treatment differences which may not be clinically meaningful.

The included RCTs that compared changes in IOP level from baseline all reported a statistically significant improvement (reduction) in IOP following treatment with all PGAs assessed. The clinical benefits were extended to trials of all duration; Nazir et al.[21] evaluated 30-day IOP difference from baseline, whereas Moussa et al.[22] and Stalmans et al.[24] evaluated IOP difference from baseline at regular intervals (e.g. 1 to 6 months). However, with the exception of Stalmans et al.,[24] a comparison between different PGAs did not result in a statistically significant between-group difference in any trial, irrespective of monotherapies or combination therapies used, prior treatment history, and other patient characteristics. Stalmans et al.[24] compared a preservative-free formulation of BIM and LAT in a crossover trial, and reported an average 1.6 (±0.5) mmHg higher IOP level at month 6 in LAT-treated patients compared with BIM; this result was statistically significant and considered to be clinically relevant by the study authors. Notably, the difference between the two treatment groups at the end of month 3 was very similar; no statistically significant difference was found. An analysis of change in IOP at month 6 from baseline within the same patient undergoing the two treatments showed an average 0.9 (± 0.2) mmHg higher IOP when treated with LAT compared with BIM; this result was statistically significant. Finally, the effect of treatment sequence was analyzed, and no statistically significant difference was found among patients treated with BIM first versus LAT first. The authors also argued that the potential bias from carry-over effect was minimized by accounting for the centre effect in the analysis, and the use of a 6-week washout cycle as well as a 3-month treatment cycle. Based on findings from Guven Yilmaz et al,[23] there was evidence of diurnal variation of IOP; however, the treatment groups showed no statistically significant difference in diurnal or nocturnal IOP.

In terms of clinical effectiveness, the authors of the respective trials concluded that the clinical profile of the PGAs (including BIM, monotherapy or in combination with TIM) were similar.

Ocular Perfusion Pressure

Rennie et al.[16] assessed the post-intervention OPP in several PGAs, an indirect measure of the vascular perfusion of the posterior ocular segment that is linked with IOP. A comparison

of PGAs showed no statistically significant differences in post-treatment OPP, although only one study involved a comparison with BIM.

Rossetti et al.[26] compared BIM and LAT/TIM with respect to systolic and diastolic blood pressure (SBP, DBP), systolic and diastolic perfusion pressure (SPP, DPP). The two treatment groups did not show any statistically significant difference in OPP, DBP, and DPP; however, LAT/TIM was associated with a statistically significantly lower SPP and SBP compared to BIM. The authors reported that low SPP is a significant predictor of progression with a 50% higher risk, and that patients with low SPP tend to progress faster – a statement that underlines the clinical superiority of BIM over LAT/TIM.

Duration of effect

One meta-analysis by Diaconita et al.[17] assessed the washout duration of the PGAs of interest, which can be used to determine the long-term effect and duration of these drugs. However, a confirmatory washout period could not be determined for BIM due to the absence of an adequate number of high-quality studies and relevant data. Nevertheless, a 4-week washout period was suggested based on findings from one RCT in the meta-analysis.

Safety endpoints

Safety endpoints were evaluated in two systematic reviews (Tang et al.[15] and Rennie et al.[16]) and three RCTs (Moussa et al.,[22] Stalmans et al.,[24] and Maruyama et al.[25]).

Findings from one meta-analysis[15] showed conjunctival hyperemia was more frequent with BIM and TRA treatment compared with LAT, and BIM also resulted in an increased incidence of growth of lashes. The authors hypothesized that the relatively poor tolerability of BIM may be a function of its higher dose (0.03%) compared to LAT (0.005%) and TRA (0.004%), which conversely may result in the greater clinical effectiveness of BIM among all PGAs. The other meta-analysis, however, did not find any association between BIM and AEs, although only one primary study provided a relevant comparison.

Findings from the RCTs indicated that conjunctival hyperemia, keratitis, and follicular conjunctivitis were reported with BIM, LAT, TRA, and TAF treatment; however, the incidence of these AEs were generally similar among the PGAs. While iris hyperpigmentation, eyelash growth, and hyperemia may occur more frequently with LAT and BIM, the evidence is inconclusive, and limited by small sample size. Based on subjective assessment and in a small number of patients, TRA appeared to be more tolerated among the PGAs. Other AEs assessed in the RCTs, namely, visual field tests, Best Corrected Visual Acuity (BCVA) measurements, cup-to-disk ratio, area density classification (AD) score, tear film break-up time (BUT), heart rate or blood pressure did not show any statistically significant difference between the PGAs (including BIM).

Cost-effectiveness of bimatoprost versus other prostaglandin analogues for ophthalmic use

In the base case, treatment with BIM conferred the greatest QALY gain, followed by LAT, TAF, TRA, and TIM, irrespective of the perspective chosen. From a societal perspective, the ICER for BIM was negative against the other drugs, indicating that therapy with BIM dominated the alternative drug therapy by providing greater patient value and was less expensive than the alternative drug. From a third-party insurer cost perspective, the ICER, including the cost which the insurer would be expected to pay (ophthalmic direct medical

costs and nonophthalmic direct medical costs such as for depression, trauma, and facility admissions resulting from glaucoma therapy), was $15,334/QALY for BIM referent to LAT, and BIM dominated TAF, TRA, and TIM in this analysis. Using only direct ophthalmic medical costs, BIM dominated TAF and was cost-effective against TRA, LAT, and TIM. Even though TRA, LAT, and TIM were less expensive than BIM, more societal costs were saved by BIM due to its greater IOP-lowering effect that confers more years of good vision.

Evidence-based guidelines for the use of bimatoprost for elevated intraocular pressure

No relevant guidelines regarding the use of bimatoprost for the treatment of elevated intraocular pressure were identified; therefore, no summary can be provided.

Limitations

The scope of the five identified systematic reviews was broader than the research question of comparative benefits of BIM with other PGAs. Therefore, the number of primary studies comparing BIM to other PGAs within each of these systematic reviews were limited. The papers by Tang et al. and Li et al. were the most comprehensive reviews, each capturing over 10 articles with BIM as an intervention. Given the broader objective (assessment of the clinical effectiveness of all PGAs), there was significant clinical and statistical heterogeneity between the studies, resulting from differences in study design, patient population, previous and concurrent treatment history, assessment of clinical outcomes, and statistical analysis. The presence of heterogeneity is reinforced by the mixed quality of studies included, as reported by the systematic review authors based on the quality assessment of the included primary studies. Taken together, these factors limit the extrapolation of the findings of the systematic reviews. Aside from the methodological limitations, the dose of BIM (and other PGAs) was only reported in Tang et al., presenting additional difficulty in differentiating the comparative effectiveness of the two strengths of BIM (0.01% versus 0.03%).

The five systematic reviews included primary studies that were all published prior to 2015. Contrary to the systematic reviews, the seven RCTs that met the inclusion criteria of this report were all published in or after 2015. This may partially account for the pattern seen in the systematic reviews (i.e., BIM was consistently found to be better than other PGAs) that was not found in relatively newer RCTs. However, some of the RCTs had a number of major limitations, including no or unclear reporting of randomization and sample size estimation. Results of the RCTs may, in part, be attributed to Hawthorne effect, (i.e., behavioral change in reaction to patients' awareness of being observed) and regression towards the mean (i.e., the value of a variable tends to be closer to the mean with repeated measurement). The trials varied in study design, patient characteristics, interventions (including doses and formulation type), and clinical assessment of outcomes. Coupled with the relatively small sample size in the trials compared to the systematic reviews, these factors contribute to the uncertainty in the study findings.

The only cost-effectiveness study was well-conducted, used standardized clinical and cost inputs in the analysis, and included recently published data. However, the study was based in the US, with societal, third party insurance, and ophthalmic perspectives chosen; which may have limited applicability in the context of Canadian publicly funded healthcare.

Despite the availability of BIM in the market for the last decade, there is a paucity in published evidence-based guidelines that includes BIM in addition to other PGAs for the

treatment of glaucoma and other ophthalmic conditions. This constitutes a key gap in evidence.

Conclusions and Implications for Decision or Policy Making

A total of 13 relevant publications were included, which comprised five systematic reviews, four with a meta-analysis[15-18] and one with an NMA;[19] seven RCTs,[20-26] of which four were open-label and the remaining three were single or double-blinded trials; and one cost-effectiveness study.[27] No evidence-based guidelines were identified regarding the use of BIM for elevated IOP. Overall, the included studies were of mixed quality, some conducted with robust, well-reported methodology, whereas a number of studies had major limitations, including inadequate reporting of study procedure (e.g. randomization method) or data, inclusion of clinically heterogeneous studies in systematic reviews, assessment of publication bias, insufficient or uncertain power, and potential conflicts of interest.

Findings from three systematic reviews with meta-analysis or NMA showed BIM to be clinically superior in lowering IOP compared to other PGAs.[15,17,19] The clinical profile of LAT, TRA, and TAF were similar, all generally efficacious but typically weaker than BIM, though no significant differences were found between BIM and TAF in the systematic review that included this comparison in an NMA.[19] The clinical benefits of BIM (and other PGAs) appears to be associated with longer term used (3 to 6 months). The effect of PGAs on IOP was not translated to other indirect measures such as OPP. While no studies directly compared the PGAs with respect to the duration of effect, one meta-analysis confirmed a 4-week washout period for LAT and suggested the same period may apply for BIM.[17] BIM showed mixed results with respect to safety endpoints; one study reported an increased incidence of conjunctival hyperemia and growth of lashes associated with BIM treatment;[15] however, another study did not show any increased safety risk.[16]

Contrary to the findings of the systematic reviews, the RCTs generally showed a similar IOP-lowering activity in all PGAs, irrespective of the treatment duration (30 days to 6 months). In six parallel group RCTs, BIM showed no statistical difference compared with LAT, TRA, or TAF (either as monotherapy or in combination with TIM). However, one crossover trial reported that treatment with BIM was associated with a lower IOP compared to LAT. The IOP level appears to have diurnal fluctuation, although none of the PGAs showed a difference in diurnal or nocturnal IOP. BIM may also increase the SBP and SPP level compared with LAT, an effect that is believed to be protective against glaucoma progression. Safety endpoints showed BIM to be less tolerated than other PGAs; however, the difference may not be substantial.

The one included economic evaluation showed BIM to be superior to all other PGAs, a finding that was consistently found in different perspectives (societal, third party insurance, and ophthalmic). However, the context (US) and perspectives chosen may have limited applicability in Canadian health care setting.

The clinical evidence from the included systematic reviews and RCTs was associated with a number of factors that presented challenges in the interpretation of results across studies. The primary studies included in the systematic reviews were all published prior to 2015, whereas the seven RCTs provided more recent evidence, which may contribute to the differences in findings seen between the systematic reviews and RCTs. The number of primary studies within the included systematic reviews that incorporated BIM as a treatment was relatively small compared to the overall number of studies. Further, the included primary studies had varying degrees of clinical and statistical heterogeneity. Despite this,

results of the systematic reviews may be more robust, as the evidence was based on a larger sample set, which was found to be a limiting factor in the seven included RCTs. Additionally, a previously published CADTH report in 2015 had similar conclusions, with BIM found to be better or equivalent to TRA and LAT, in terms of clinical benefits, safety, and cost profile, although results were not conclusive.[9]

Overall, the comparative clinical benefits of BIM and other PGAs are not conclusive, although BIM appears to be at least as effective as LAT, TRA, or TAF (with or without TIM), with a similar safety profile. However, BIM was found to be the most cost-effective treatment of all PGAs. Based on the similar clinical profile and superior cost-effectiveness, BIM may be the preferred choice of PGAs. A number of key evidence gaps should be addressed in future research for a conclusive assessment of PGAs, including standardization of study design, protocol, and reporting, in a sufficiently large sample set, with robust methodology.

References

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA.* 2014;311(18):1901-1911.
2. Jacobs DS. Open-angle glaucoma: treatment. In: Post TW, ed. *UpToDate*. Waltham (MA): UpToDate; 2020: www.uptodate.com. Accessed 2020 Feb 11.
3. Canadian Ophthalmological Society Glaucoma Clinical Practice Guideline Expert Committee, Canadian Ophthalmological Society. Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of glaucoma in the adult eye. *Can J Ophthalmol.* 2009;44 Suppl 1:S7-93.
4. Types of glaucoma. Ashford (UK): International Glaucoma Association: https://www.glaucoma-association.com/about-glaucoma/types-of-glaucoma/. Accessed 2020 Feb 11.
5. Glaucoma care for adults. *Quality Standard*. Toronto (ON): Health Quality Ontario; 2019: https://www.hqontario.ca/Portals/0/documents/evidence/quality-standards/qs-glaucoma-qs-en.pdf. Accessed 2020 Feb 11.
6. Table: 13-10-0466-01. Healthy aging indicators. Ottawa (ON): Statistics Canada: http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=1051200. Accessed 2020 Feb 11.
7. Choi YM, Diehl J, Levins PC. Promising alternative clinical uses of prostaglandin F2alpha analogs: beyond the eyelashes. *J Am Acad Dermatol.* 2015;72(4):712-716.
8. Glaucoma: diagnosis and management. *(NICE guideline NG81)*. London (UK): National Institute for Health and Care Excellence; 2017: https://www.nice.org.uk/guidance/ng81. Accessed 2019 Feb 8.
9. Prostaglandin analogues for ophthalmic use: a review of the comparative clinical effectiveness and cost-effectiveness. *(CADTH rapid response report: summary with critical appraisal)*. Ottawa (ON): CADTH; 2015: https://www.cadth.ca/sites/default/files/pdf/htis/july-2015/RC0684-Prostaglandins-Final.pdf. Accessed 2020 Feb 11.
10. Lou H, Wang H, Zong Y, Cheng JW, Wei RL. Efficacy and tolerability of prostaglandin-timolol fixed combinations: an updated systematic review and meta-analysis. *Curr Med Res Opin.* 2015;31(6):1139-1147.
11. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008. http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf. Accessed 2020 Jan 20.
12. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health.* 2011;14(4):417-428.
13. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf. Accessed 2020 Jan 20.
14. Higgins JPT, Green S, editors. Figure 15.5.a: Drummond checklist (Drummond 1996). *Cochrane handbook for systematic reviews of interventions*. London (GB): The Cochrane Collaboration; 2011: http://handbook-5-1.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm. Accessed 2020 Jan 20.
15. Tang W, Zhang F, Liu K, Duan X. Efficacy and safety of prostaglandin analogues in primary open-angle glaucoma or ocular hypertension patients: a meta-analysis. *Medicine.* 2019;98(30):e16597.
16. Rennie G, Wilkinson A, White A, et al. Topical medical therapy and ocular perfusion pressure in open angle glaucoma: a systematic review and meta-analysis. *Curr Med Res Opin.* 2019;35(8):1421-1431.
17. Diaconita V, Quinn M, Jamal D, et al. Washout duration of prostaglandin analogues: a systematic review and meta-analysis. *J Ophthalmol.* 2018;2018:3190684.
18. Takagi Y, Santo K, Hashimoto M, Fukuchi T. Ocular hypotensive effects of prostaglandin analogs in japanese patients with normal-tension ocular hypotensive effects of prostaglandin analogs in Japanese patients with normal-tension glaucoma: a literature review. *Clin Ophthalmol.* 2018;12:1837-1844.
19. Li T, Lindsley K, Rouse B, et al. Comparative effectiveness of first-line medications for primary open-angle glaucoma: a systematic review and network meta-analysis. *Ophthalmology.* 2016;123(1):129-140.
20. Blondeau P, Hamid M, Ghalie Z. Prospective randomized clinical trial on the effects of latanoprost, travoprost and bimatoprost on latanoprost non-responders. *J Fr Ophtalmol.* 2019;42(8):894-899.
21. Nazir N, Ali Z, Latif E, Nazir I, Alvi Z. Comparison of efficacy of latanoprost 0.005% with bimatoprost 0.01% in patients with open angle glaucoma. *Med Forum Mon.* 2019;30(7):17-21.
22. El Hajj Moussa WG, Farhat RG, Nehme JC, et al. Comparison of efficacy and ocular surface disease index score between bimatoprost, latanoprost, travoprost, and tafluprost in glaucoma patients. *J Ophthalmol.* 2018:1319628.
23. Guven Yilmaz S, Degirmenci C, Karakoyun YE, Yusifov E, Ates H. The efficacy and safety of bimatoprost/timolol maleate, latanoprost/timolol maleate, and travoprost/timolol maleate fixed combinations on 24-h IOP. *Int Ophthalmol.* 2018;38(4):1425-1431.

24. Stalmans I, Oddone F, Cordeiro MF, et al. Comparison of preservative-free latanoprost and preservative-free bimatoprost in a multicenter, randomized, investigator-masked cross-over clinical trial, the SPORT trial. *Graefes Arch Clin Exp Ophthalmol.* 2016;254(6):1151-1158.

25. Maruyama Y, Ikeda Y, Mori K, et al. Comparison between bimatoprost and latanoprost-timolol fixed combination for efficacy and safety after switching patients from latanoprost. *Clin Ophthalmol.* 2015;9:1429-1436.

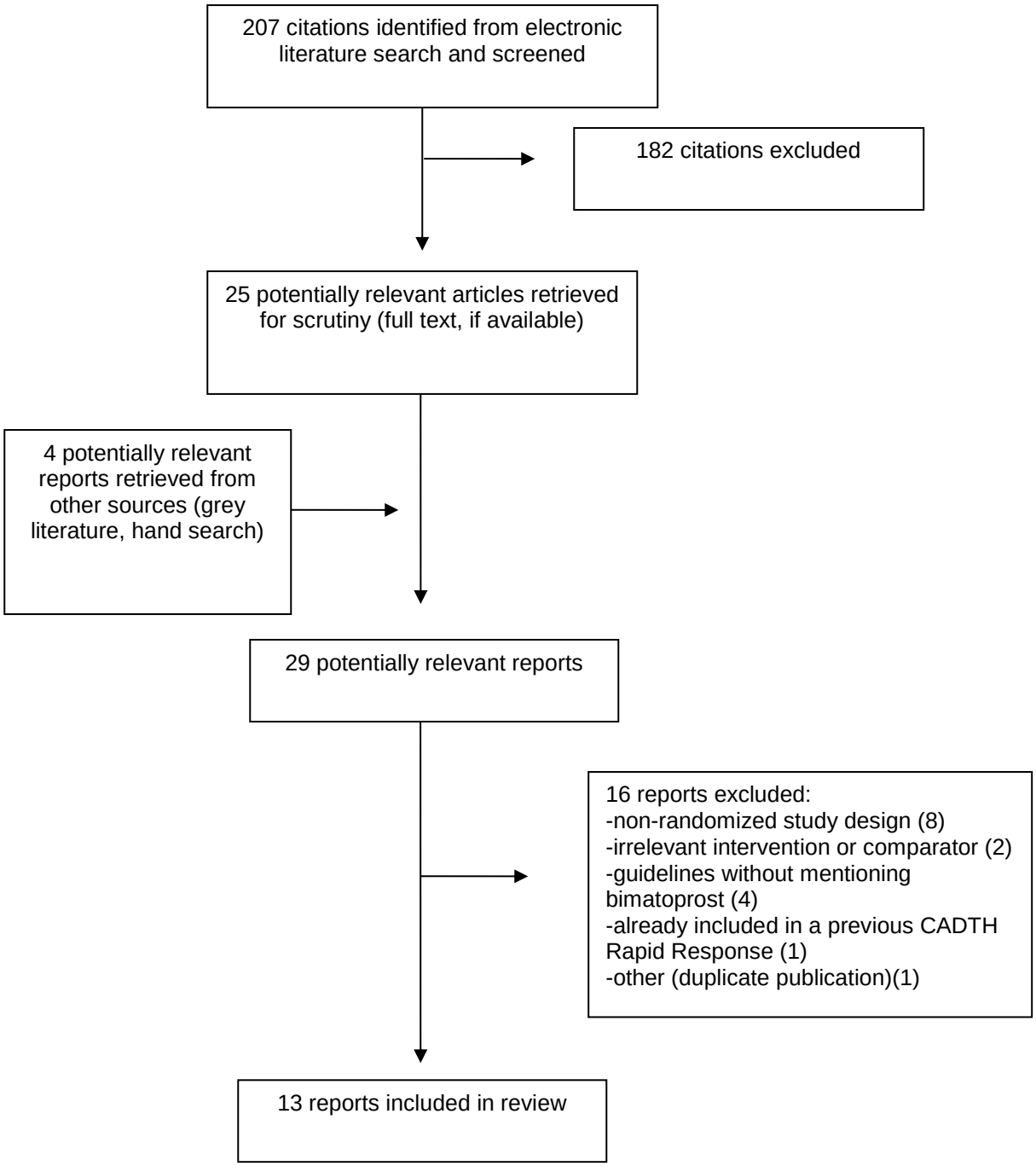
26. Rossetti L, Sacchi M, Karabatsas CH, et al. Comparison of the effects of bimatoprost and a fixed combination of latanoprost and timolol on 24-hour blood and ocular perfusion pressures: the results of a randomized trial. *BMC Ophthalmol.* 2015;15:7.

27. Brown GC, Brown MM. Patient preference-based comparative effectiveness and cost-utility analysis of the prostamides for open-angle glaucoma. *J Ocul Pharmacol Ther.* 2019;35(3):145-160.

28. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.

29. Guidelines for the economic evaluation of health technologies: Canada. *(CADTH methods and guidelines)*. Ottawa (ON): CADTH; 2017: https://www.cadth.ca/dv/guidelines-economic-evaluation-health-technologies-canada-4th-edition. Accessed 2020 Feb 11.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Reviews and Meta-Analyses

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
Tang et al. 2019[15] China	**Objective:** To compare the efficacy and safety of 0.005% LAT, 0.004% TRA and 0.03% BIM in the treatment of patients with POAG or OHT. **Lit search strategy:** Authors searched for literature published between January 1st 2000 and June 1st 2018 using PubMed, Embase, Cochrane library, Web of science, CNKI, Wanfang, and Vip databases. **No. of studies included:** 17 RCTs included in total, 15 compared BIM to other PGAs	Patients with POAG or OHT (average age 52 to 68 years).	LAT (0.005%), TRA (0.004%) and BIM (0.03%).	**Efficacy:** Mean IOP reduction from baseline to endpoint **Safety:** AEs including conjunctiva hyperemia, discomfort (itching, eye irritation, foreign body sensation), and growth of lashes **Length of follow-up:** Range 4–12 months
Rennie et al. 2019[16] Australia	**Objective:** To compare the benefits and harms of topical interventions for OPP in OAG. **Lit search strategy:** Authors conducted literature search on 30 March 2018 (no date restrictions) using MEDLINE, EMBASE and Cochrane register of controlled trials (CENTRAL). **No. of studies included:** 10 RCTs included in total, 1 compared BIM to other PGAs.	Patients with POAG, NTG, PXG and/or OHT (average age, 43 to 67 years across studies).	Topical monotherapies commonly used in glaucoma, including beta-blockers (TIM, timogel), PGA (LAT, BIM, TAF, TRA), alpha-2 adrenergic agonists (brimonidine) and carbonic anhydrase inhibitors (dorzolamide).	**Efficacy:** (Post-intervention) Mean OPP **Safety:** AEs **Length of follow-up:** Studies 8 weeks (median), interquartile range 4–10 weeks
Diaconita et al. 2018[17] Canada	**Objective:** To compare the washout of PGAs and quantify the duration and long-term effect of reported PGA washout periods. To investigate the long-term effects on IOP after discontinuation of topical PGAs.	Patients (>18 years of age) with glaucoma and OHT	Treatment with PGA (LAT, TRA, BIM, UNO) or prostamide monotherapy or combination therapy	**Clinical outcomes:** Long-term effects on IOP by assessing washout duration (pre- and postwashout IOP) **Length of follow-up:** NR

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
	Lit search strategy: Authors searched for literature published between 1985 and October 2016 using MEDLINE/PubMed, EMBASE, Cochrane Library, CINAHL, Web of Science, and BIOSIS Previews and conference proceedings. **No. of studies included:** 56 papers included in total, 1 RCT compared BIM to LAT.			
Takagi et al. 2018[18] Japan	**Objective:** To evaluate the IOP-lowering effects of PGAs **Lit search strategy:** Authors conducted a literature search on February 2018 using PubMed, Embase, ProQuest, and the Japanese databases JAPICDOC and JMEDPlus. **No. of studies included:** 27 papers included in total; 3 studies compared BIM to other PGAs or no treatment.	Japanese patients with NTG who were naïve to medication therapy or newly started on medications, with a comprehensive washout period.	Ocular hypotensive therapies, including UNO, LAT, TRA, TAF, BIM.	**Clinical outcomes:** Ocular hypotensive efficacy based on reduction in IOP values **Length of follow-up:** Range 4 weeks–6 months
Li et al. 2016[19] USA	**Objective:** To assess the comparative effectiveness of first line medical treatments **Lit search strategy:** Authors updated a previously conducted literature search on March 11, 2014 using CENTRAL, MEDLINE, EMBASE and the Food and Drug Administration's website. **No. of studies included:** 114 RCTs included in total; 13 studies compared BIM to other PGAs or no treatment.	Patients with POAG or OHT	First line topical medical interventions including beta blockers, carbonic anhydrase inhibitors, alpha-2 adrenergic agonists, and PGAs (BIM, LAT, TRA, TAF, UNO)	**Clinical outcomes:** IOP (difference in mean IOP, mean diurnal IOP, 24-hour mean IOP, peak IOP, morning IOP, and trough IOP) Visual field **Follow-up:** 3 months

AE = adverse events; BIM = bimatoprost; DB=double-blind; IOP = intraocular pressure; LAT=latanoprost; NTG = normal tension glaucoma; OHT = ocular hypertension; OPP = ocular perfusion pressure; P/OAG = Primary / open-angle glaucoma; PGA = prostaglandin analogue; PXG = pseudo-exfoliative glaucoma; TAF = tafluprost; TRA=travoprost; UNO = unoprostone.

Table 3: Characteristics of Included Primary Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
Blondeau et al. 2019[20] Canada	OL RCT with a switch design	Treatment-naive patients (≥ 18 years, mean age 68.8 years) (N = 83) newly diagnosed with OHT, POAG, NTG, PXG, or PG who were non-responder (≤ 20% IOP reduction) to LAT after one month of treatment	LAT pre-treatment for 1 month, single drop at bedtime Non-responders (< 20% IOP reduction from baseline) randomized to continue LAT (n = 29), or switch to BIM (n = 31) or TRA (n = 23) (no washout, no dosing information)	**Clinical outcomes:** IOP, number of new responders **Follow-up:** 1 month
Nazir et al. 2019[21] Pakistan	OL single-centre RCT	Patients (20-50 years) (N = 60) newly diagnosed with OAG presented with > 20 mmHg IOP	LAT (0.005%) vs BIM (0.01%) monotherapy (n = 30 in each group); one drop daily evening	**Clinical outcomes:** IOP **Follow-up:** 30 days
Moussa et al. 2018[22] Lebanon	OL single-centre RCT	Treatment-naive patients (age 30-82) (N = 32) with newly diagnosed POAG	BIM 0.01% (with BAK 0.02%) (n = 8), LAT 0.05% (with BAK 0.02%) (n = 7), TRA 0.004% (with 0.001% polyquad) (n = 8), and TAF 0.0015% (BAK-free) (n = 9); one drop in each eye every evening	**Clinical efficacy:** IOP **Safety:** Dry eye disease assessed using the original Ocular Surface Disease Index Questionnaire **Follow-up:** 6 months
Guven Yilmaz et al. 2018[23] Turkey	Observer-masked single-centre RCT	Patients (N = 54) with POAG who had failed to reach an IOP of < 21 mmHg while receiving bilateral monotherapy of hypotensive lipids and timolol maleate fixed combination treatment for ≥ 4 weeks	Switch from respective monotherapy to fixed combination of BIM/TIM (n =18), LAT/TIM (n = 14), and TRA/TIM (n = 18) once-daily before bedtime	**Clinical outcomes:** IOP, diurnal and nocturnal IOP variation **Follow-up:** IOP was monitored through 24-hr
Stalmans et al. 2016[24] 7 European centres (Belgium, Italy, UK, Austria, Portugal, Switzerland)	Investigator-masked, crossover, multi-centre RCT	Patients (age 39 to 85 years) (N = 67) with OHT, exfoliation glaucoma, or OAG who had an IOP of ≤ 21 mmHg on a preserved prostaglandin monotherapy for at least 6 weeks and needing treatment in both eyes	Preservative-free BIM (0.03%) vs LAT (dose not given), single drop once in the evening	**Clinical efficacy:** IOP (average of 3 diurnal measurements) at 3 and 6 months **Safety:** visual acuity, AEs, slit-lamp biomicroscopy, ocular tolerability, and optic nerve assessment. **Follow-up:** 3 months of each treatment, 6 weeks of washout
Maruyama et al. 2015[25] Japan	OL RCT	Japanese patients (N = 70) (mean age 66.9 years) with OAG who had used LAT	BIM (0.03%) (n = 30) or LAT/TIM fixed	**Clinical outcomes:** IOP, conjunctival injection score, corneal

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
		monotherapy for more than 4 weeks without adequate IOP control or had progressive glaucomatous changes in their visual field	combination (0.005%/0.5%) (n = 27)	epitheliopathy score (area density classification; AD score), tear film break-up time, heart rate, and blood pressure **Follow-up:** 12 weeks
Rossetti et al. 2015[26] 7 European University Eye Clinics (Italy, Greece, Germany, and Switzerland)	DB, multi-centre RCT	Adult patients (N = 200) with POAG/PEX or OHT, who were controlled (IOP <21 mmHg) on the unfixed combination of LAT and TIM for ≥ 3 months or eligible for dual therapy due to not being fully controlled on monotherapy (IOP >21 mmHg, or target IOP not reached)	BIM in the evening and fixed combination LAT/TIM in the morning, both with matching placebo	**Clinical outcomes:** 24-hr SBP and DBP, 24-hr OPP, 24-hr SPP and DPP **Follow-up:** 12 weeks

AE = adverse event; BAK = benzalkonium chloride; BIM = bimatoprost; DB=double-blind; IOP = intraocular pressure; LAT=latanoprost; NTG = normal tension glaucoma; OHT = ocular hypertension; OL = open-label; OPP = ocular perfusion pressure; P/OAG = Primary / open-angle glaucoma; PGA = prostaglandin analogue; PG = pigmentary glaucoma; PXG = pseudo-exfoliative glaucoma; RCT = randomized controlled trial; TAF = tafluprost; TIM = timolol; TRA=travoprost; UNO = Unoprostone

Table 4: Characteristics of Included Economic Evaluations

First Author, Publication Year, Country	Type of Analysis, Time Horizon, Perspective	Decision Problem	Population Characteristics	Intervention and Comparator(s)	Clinical and Cost Data Used in Analysis	Main Assumptions
Brown and Brown, 2019[27] USA	Value-Based Medicine, incremental cost-utility analysis, and average cost-utility analysis, 20-year model Ophthalmic, third party insurer, and societal cost perspectives	To perform patient preference-based comparative effectiveness and cost-effectiveness analyses to evaluate topical BIM, LAT, TRA, TAF, and TIM for the treatment of OAG	Patients with OAG with an untreated IOP of 26 mmHg	Topical BIM 0.01%, LAT 0.005%, TRA 0.004%, TAF 0.0015%, and TIM 0.5%	3-month clinical data based on published systematic review and NMA. AEs are based upon a double-masked, randomized BIM trial, 2 meta-analyses, and FDA randomized, double-masked clinical trial files. Time to end-stage glaucoma at different IOP, Visual field correlation with	All participants are adults with POAG Untreated people with glaucoma realize that they have glaucoma Patients have glaucoma in both eyes and are treated bilaterally Mean age of onset of POAG = 65 years Mean life expectancy at age 65 is 20 years

First Author, Publication Year, Country	Type of Analysis, Time Horizon, Perspective	Decision Problem	Population Characteristics	Intervention and Comparator(s)	Clinical and Cost Data Used in Analysis	Main Assumptions
					utilities are based on seminal studies Time tradeoff utilities are derived from patients utilizing the Center for Value-Based Medicine Utility Database Cost data: national average Medicare Fee Schedule Drug prices are taken from 2015 AWP data	Baseline IOP of glaucoma study cohorts = 26mm Hg Overall visual fields decrease 4 dB beginning 36 months before end-stage glaucoma. The associated utility is 0.902 3% annual discount rate for patient value outcomes (QALYs) and costs 20% miss rate for glaucoma drop administration

AE = adverse event; AWP = Average Wholesale Price; BIM = bimatoprost; DB=double-blind; IOP = intraocular pressure; LAT=latanoprost; FDA = Food and Drug Agency; NMA = network meta-analysis; NTG = normal tension glaucoma; OHT = ocular hypertension; P/OAG = Primary / open-angle glaucoma; PGA = prostaglandin analogue; RCT = randomized controlled trial; TAF = tafluprost; TIM = timolol; TRA=travoprost

Appendix 3: Critical Appraisal of Included Publications

Table 5: Strengths and Limitations of Systematic Reviews with Meta-Analyses using AMSTAR[11] and with Network Meta-Analyses using ISPOR[12]

Strengths	Limitations
Tang et al. 2019[15]	
• The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with predefined search strategy; in addition, a manual search from reference list of retrieved papers and review articles was also performed. The reporting of the search strategy followed the requirements of PRISMA statement. • Study selection, data extraction and study quality assessment was done independently by two investigators, disagreements were resolved by discussion or consensus involving a third investigator. • The Cochrane bias risk assessment tool (The Cochrane Collaboration) for RCTs was used to assess the risk of bias and quality of evidence, with discrepancies resolved by discussion between the authors. • Study and patient characteristics of each included trial were provided with adequate details. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	• With the exception of two included trials, the remaining 15 trials were judged by systematic review authors to be at high or unclear risk of bias in at least one item. Overall, the major potential sources of bias in the trials were selection and performance bias. • An assessment of publication bias was not done, and not justified.
Rennie et al. 2019[16]	
• A systematic and comprehensive literature search was conducted, with predefined search strategy. No date restrictions were applied, although studies published in English were included only. The reporting of the search strategy followed the requirements of PRISMA statement. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • Study selection, data extraction and study quality assessment were done independently by two investigators; disagreement in study selection was resolved with an arbitrator. • The Cochrane bias risk assessment tool (The Cochrane Collaboration) for RCTs was used to assess the risk of bias and quality of evidence. • The overall quality of the trials as assessed by systematic review authors was low to moderate, where 7 of 10 included studies had at least one Cochrane domain assessed as being at high risk of bias. Overall, the included studies performed poorly in the domains of selection, performance and detection bias.	• It is unclear if study quality assessment was done independently by the two authors, and how disagreements were resolved. • An assessment of publication bias was attempted by means of a Funnel plot; however, it was not possible due to the small number of studies included.

Strengths	Limitations
• Study and patient characteristics of each included trial were provided with adequate details. • Statistical methods for data analysis appeared to be appropriate, between-study heterogeneity was assessed, subgroup analyses were done that were aligned with the objective of the systematic review. • All authors and reviewers of the report declared no conflicts of interest. It was declared that the paper was not funded; however, a peer reviewer of the manuscript declared receiving unrestricted research fund from a Pharmaceutical, which is a distributor of BIM in Japan.	
Diaconita et al. 2018[17]	
• The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with a predefined search strategy. • The PRISMA flow diagram was provided that demonstrated the selection of studies included. However, there was no information if the reporting of the search strategy followed the requirements of PRISMA statement. • Study selection and data extraction were done independently by two investigators, disagreements were resolved by discussion or consensus. • The Downs and Black checklist was used to assess the qualities of included studies, without any information how discrepancies were resolved. A further quality check was performed to ensure the completeness of study methodology. • Statistical methods for data analysis appeared to be appropriate, between-study heterogeneity was assessed, subgroup analyses were done that were aligned with the objective of the systematic review. • Study and patient characteristics of each included trial were provided with somewhat adequate details. • Publication bias was assessed using a funnel plot; visual inspection of the funnel plot for both pre and post-washout IOP did not reveal any asymmetry. • All authors and reviewers of the report declared no conflicts of interest, although the source of funding for the report was not provided.	
Takagi et al. 2018[18]	
• A systematic and comprehensive literature search was conducted, with predefined search strategy. No date restrictions were applied, although studies published in English and Japanese were included only. • The PRISMA flow diagram was provided that demonstrated the selection of studies included. However, there was no information if the reporting of the search strategy followed the requirements of PRISMA statement.	• There is no information on whether study selection, data extraction and study quality assessment were done by one or multiple authors; and how disagreement was resolved. • No attempt was made to assess publication bias. • The study was mainly industry funded, 3 of the 4 authors are employees of a Pharmaceutical and received an advisor fee.

Strengths	Limitations
• The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • Study and patient characteristics of each included trial were provided with somewhat adequate details. • Statistical methods for data analysis appeared to be appropriate.	
Systematic review with network meta-analysis (appraised using the ISPOR questionnaire)[12]	
Li et al. 2016[19]	
Introduction • The rationale and objective of the review was clearly stated. **Methods** • A comprehensive and systematic search of the literature was conducted in accordance with the PRISMA statement and PRISMA extension statement for NMAs. • The search strategy and the study selection criteria were clearly stated. No date or language restrictions were in place, non-English language reports were assessed by a single individual who was a native or fluent speaker of the language. • Study selection, data extraction and study quality assessment were done independently by two reviewers; inconsistencies were resolved through discussion or consultation with a third reviewer. • The Cochrane collaboration risk of bias tool was used to appraise the methodological quality of included RCTs. The quality of the included studies varied, with low, medium, high, or unclear risk of bias found depending on the domains. The authors concluded that the overall risk of bias of the included RCTs was mixed at best. • The outcome measures were selected appropriately and clearly described. • Statistical method accounting for between-study heterogeneity was factored in the analysis. • The following 3 approaches were used to assess inconsistency: loop-specific approach, node-splitting approach, and a comparison of Bayesian model fit with and without assuming consistency in the network. In addition, the influence of selected trial characteristics (e.g., funding source, big effect size) was checked qualitatively when statistically significant inconsistency was detected and sensitivity analysis was conducted by removing studies seem to introduce statistical inconsistency. Upon finding statistical inconsistency, an “inconsistency” model was introduced to the data and a sensitivity analysis was conducted; however, this did not improve the model fit and did not change the conclusions. • The methods for indirect comparison (Bayesian network meta-analyses) and ranking treatment probabilities for each PGA were appropriate.	**Methods** • The models were conducted without covariate adjustment for patient or study characteristics, and hence control of potential bias could not be assessed. • There was no information on prior distributions for model parameters, and whether priors were informative or noninformative. • No assessment of publication bias was done, either visually using funnel plots or quantitatively using the Begg’s test and Egger’s test. **Discussion:** • Changes in visual field and optic nerve damage are more meaningful metrics of the effectiveness of glaucoma treatment compared to IOP. However, changes in visual field and optic nerve damage are difficult to quantify and effects on patient-centered outcomes such as visual function and blindness require a long follow-up time to observe. Since IOP reduction correlates with preservation of visual field, it is commonly used as a surrogate outcome in glaucoma trials. The authors attempted to quantify changes in visual field and optic nerve damage; however, only 11% of the included trials reported any analyzable visual field data and the data were measured and reported in many different ways, making a pair-wise meta-analysis or an NMA impossible for visual field and optic nerve damage. In addition, the short length of follow-up time of most trials (median 3 months), in contrast to what is needed to observe visual field change also precluded meaningful assessments of visual field outcome.

Strengths	Limitations
Results • Identification and selection of full-text studies for the NMA were well reported, as well as presented in a PRISMA flowchart. Additionally, a network diagram was provided. • A table with study and patient characteristics was provided; and summary effect estimates from each included trial was available. • Convergence of all models was assessed using the Gelman-Rubin diagnostic. **Discussion** • A description of the main findings was presented that highlighted the potential limitations of the results as well as possible explanations for discrepancies across studies. • The authors did not provide a discussion on the generalizability of findings; however, given the included studies combined a high number of patients, all available PGAs and commonly used anti-glaucomatous medications, and trials conducted across the world in various settings, the generalizability of the results is not likely to be a concern.	

BIM = bimatoprost; PGA = prostaglandin analogue; PRISMA = Preferred Reporting Items for Systematic Review and Meta-analysis Protocols statement; RCT = randomize controlled trial

Table 6: Strengths and Limitations of Clinical Studies using the Downs and Black Checklist[13]

Strengths	Limitations
Blondeau et al. 2019[20]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, interventions being compared, potential confounders, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • Randomization method was reported, although not in a clear manner.	**External validity** • Patients were excluded if they had previous intraocular surgery 6 months prior to recruitment, had other ophthalmic drop treatments, had contraindication to prostaglandins, corneal abnormalities, used cortisone medication, were monophthalmic, or were pregnant or nursing. Therefore, the results were not generalizable to these patients. However, these exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • During the course of the study, the two available doses of TRA and BIM became mixed, such that patients received Travatan and Travan Z, Lumigan and Lumigan RC at different periods of the study. • There was no washout period after LAT pre-treatment. • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. **Sample size/power**

Strengths	Limitations
• This was an open-label trial, therefore investigators and patients were unblinded. However, the measurement procedure appeared to be done in an objective manner. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics.	Even though the authors conducted a power calculation, it is unclear if the trial had sufficient power to detect a clinically important effect for the primary endpoint. The authors estimated that 47 patients per group would result in a 90% power; however, the maximum number of patients was in the BIM group (n = 31).
Nazir et al. 2019[21]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, interventions being compared, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • Randomization method was reported, although not in a clear manner. • This was an open-label trial; therefore, investigators and patients were unblinded. However, the measurement procedure appeared to be properly done. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate.	**Reporting** • Baseline characteristics were not provided; therefore, the distribution of the baseline characteristics could not be assessed. **Internal validity** • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. **Sample size/power** • The authors did not conduct a formal power calculation; therefore, it is not possible to assess if the trial had sufficient power to detect a statistically significant effect for the primary endpoint.
Moussa et al. 2018[22]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, interventions being compared, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with	**Internal validity** • Randomization method was not reported; therefore, the appropriateness of randomization could not be assessed. • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. **Sample size/power** • The authors did not conduct a formal power calculation; therefore, it is not possible to assess if the trial had sufficient power to detect a statistically significant effect for the primary endpoint.

Strengths	Limitations
serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • This was an open-label trial; therefore, investigators and patients were unblinded. However, the measurement procedure appeared to be done in an objective manner. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics.	
Guven Yilmaz et al. 2018[23]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, diagnosis of patients, interventions being compared, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • This was an observer-masked trial; therefore, investigators and staff were blinded to the treatment allocation. The concealment of treatment appeared to be maintained. • The measurement procedure for the study outcomes appeared to be properly done. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics.	**Internal validity** • Randomization method was not reported; therefore, the appropriateness of randomization could not be assessed. • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. **Sample size/power** • The authors did not conduct a formal power calculation; therefore, it is not possible to assess if the trial had sufficient power to detect a statistically significant effect for the primary endpoint.

Strengths	Limitations
Stalmans et al. 2016[24]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, diagnosis of patients, interventions being compared, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • This was an investigator-masked trial; therefore, investigators were blinded to the treatment allocation. The concealment of treatment appeared to be maintained, as patients were requested not to communicate the type of treatment to the investigator. • The measurement procedure for the study outcomes (e.g. IOP measurements, biomicroscopy, ophthalmoscopy) as well as washout appeared to be properly done. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics (e.g. best corrected visual acuity, visual field Mean Defect, cup-to-disc ratio of the optic disc (c/d ratio), screening IOP values and baseline IOP values, and prior treatment history). **Sample size/power** • The authors conducted a formal power calculation; results showed the study had over 80% power in detecting a difference in mean IOP of 1 mmHg between the two treatment groups (which was assumed to be a clinically relevant difference), with a type I error of 0.05, standard deviation of 2.8 mmHg, and assuming approximately 10% rate of withdrawals.	**Internal validity** • Randomization method was not reported; therefore, the appropriateness of randomization could not be assessed. • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes.
Maruyama et al. 2015[25]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, diagnosis of patients, interventions being compared, and main findings were clearly described.	**Internal validity** • Randomization method was not reported; therefore, the appropriateness of randomization could not be assessed.

Strengths	Limitations
• The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **Internal validity** • This was an open-label trial; therefore, investigators and patients were unblinded. However, the measurement procedure appeared to be done in an objective manner. • The measurement procedure for the study outcomes (e.g. IOP measurements, biomicroscopy, ophthalmoscopy) as well as washout appeared to be properly done. • The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics (e.g. age, sex).	• Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. • Thirteen of the 70 patients dropped out, primarily due to poor compliance, side effects. This may have biased the results negatively, since the sample size was small to begin with. **Sample size/power** • The authors conducted a formal power calculation; results showed 35 patients per group would allow a mean IOP difference of 1.7 mmHg and standard deviation of 2.5 mmHg between the treatment groups to be detected at a 5% significance level and 80% power. However, no provision was made for dropouts, and 13 out of the 70 initially recruited patients dropped out, which may have lowered the study power. **External validity** • It is known that hearth rate and blood pressure may be affected by β-blockers, among other factors. Since β-blockers cannot be used in patients with heart or pulmonary disease, these patients were excluded from the trial. The authors anticipated that this exclusion may have affected the results.
Rossetti et al. 2015 [26]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, diagnosis of patients, interventions being compared, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients with serious or interfering comorbid conditions, history of or currently receiving confounding treatments, allergy to any of the study drugs, and those unwilling or unlikely to be adherent to the study protocol were excluded. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • Randomization method was reported; and was done in an appropriate manner. • This was a double-blind trial; therefore, both investigators and patients were blinded to the treatment allocation. The concealment of treatment appeared to be maintained, as matching placebo was used with the study drug. • The measurement procedure for the study outcomes (e.g. IOP measurements, biomicroscopy, ophthalmoscopy) as well as washout appeared to be properly done.	**External validity** • The authors suspected that “Hawthorne effect”, i.e. the alteration of behavior by patients in a clinical trial setting due to their awareness of being observed, and the highly selected patient population may have contributed to the study findings to some extent, which may not be generalizable to the overall population.

Strengths	Limitations
• The outcomes measured were valid and reliable, and outcome definitions were consistent with widely accepted criteria. • The statistical tests were appropriate; the effect of confounders was assessed using appropriate analytical method. • There were no major imbalances between treatment arms in terms of baseline characteristics (e.g. age, sex, distribution of glaucoma type, central corneal thickness, concomitant medications). **Sample size/power** • The authors conducted a formal power calculation based on the 24-hour mean change in IOP. Results showed the study had over 80% power in detecting a difference in mean IOP of 1.5 mmHg (with a SD of 3.8 mmHg) between the two treatment groups at a 5% significance level, with 81 patients per group (1-sided test) and 102 patients per group (2-sided test). Further, P-values were adjusted for multiple comparisons (Bonferroni corrections).	

BIM = bimatoprost; LAT=latanoprost; PGA = prostaglandin analogue; TAF = tafluprost; TRA=travoprost;

Table 7: Strengths and Limitations of Economic Studies using the Drummond Checklist[14]

Strengths	Limitations
Brown and Brown, 2019[27]	
Study design • The research question, its economic importance, the perspective of the analyses, and the rationale for choosing the interventions were stated. • The comparators were clearly described, and rationale provided. • The form of economic evaluation was stated and justified in relation to the questions addressed. **Data collection** • The source of clinical effectiveness data, description and results of the source study, extrapolation method for the total time horizon, primary outcomes measures, and utility values were provided. • The study utilized a number of standardized inputs and outputs. Standardized methodology for cost-utility analysis was useful in allowing comparison across studies. • The study included drug and supportive care cost, which as justified since the perspective was from a societal as well as US third party payer. The study included relevant costs given the context and perspective, it covered other areas of costs including health care resources, non-medical costs such as costs of capital, operation, caregivers' and individuals' due to time lost from usual activity.	**Analysis and interpretation of results** • In the analysis, OAG was a modeled as a bilateral disease that progresses symmetrically, even though first and second eye involvement is done with the more commonly used Markov modeling. Nonetheless, a sensitivity analysis was done by halving the patient value gain from each prostamide, a feature that more than compensated for changes that would be induced by Markov modeling of first and second eyes, still revealed good cost-effectiveness for each drug.

Strengths	Limitations
• Methods for the estimation of quantities and unit costs were described. • Details of the analysis method was provided, including key parameters of the model. **Analysis and interpretation of results** • The time horizon for costs and benefits was stated, and accounted for the lifetime of the patients following diagnosis. • The discount rate was stated. • Details of statistical tests were provided, incremental analysis was reported. • A number of sensitivity and scenario analyses were performed to test the robustness of the model assumptions. • The answer to the study question was given, conclusions were based on the data reported, and accompanied by the appropriate caveats.	

OAG = open angle glaucoma

Appendix 4: Main Study Findings and Authors' Conclusions

Table 8: Summary of Findings Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusion
Tang et al. 2019 [15]	
Efficacy: The comparative efficacy of the PGAs was reported by the treatment period of the study drug, and comparisons with BIM are reported here only. The mean IOP reduction from baseline between treatment groups was expressed as WMD, with values > 0 favoring left PGAs. A total of 3, 9, and 5 trials were pooled assessing LAT versus BIM at 1, 3, and 6 months post-treatment, respectively. Data indicated that BIM was more effective for IOP control in the third and sixth month. • At month 1, the WMD between LAT and BIM was -0.11 mmHg (95% CI, -0.97 to 0.76, P = 0.81, I^2 = 0%, heterogeneity P = 0.1) • At month 3, the WMD between LAT and BIM was -0.75 mmHg (95% CI, -1.05 to -0.45, P < 0.00001, I^2 = 61%, heterogeneity P = 0.009) • At month 6, the WMD between LAT and BIM was -0.82 mmHg (95% CI, -1.55 to -0.09, P = 0.03, I^2 = 35%, heterogeneity P = 0.19) A total of 5, 9, and 6 trials were pooled assessing TRA versus BIM at 1, 3, and 6 months post-treatment, respectively. Data indicated that BIM showed greater efficacy in lowering IOP in the third month. • At month 1, the WMD between TRA and BIM was -0.64, 95% CI -1.64 to 0.37, P = 0.21, I^2 = 43%, heterogeneity P = 0.13) • At month 3, the WMD between TRA and BIM was -0.93, 95% CI -1.25 to -0.60, P < 0.00001, I^2 = 1%, heterogeneity P = 0.43) • At month 6, the WMD between TRA and BIM was -0.71, 95% CI -1.65 to 0.23, P = 0.14, I^2 = 49%, heterogeneity P = 0.08) **Safety:** The safety endpoints were dichotomous outcomes and expressed as OR, with values < 1 favoring left PGAs. • Based on the fixed model, LAT generally showed a better safety profile compared to BIM, with overall OR 0.4 (95% CI; 0.30 to 0.53, P < 0.00001, I^2 = 0%, heterogeneity P = 0.84), ORs for conjunctival hyperemia, discomfort, and growth of lashes 0.36 (95%CI; 0.26 to 0.51, P < 0.00001, I^2 = 0%, heterogeneity P = 0.71), 0.63 (95% CI; 0.35 to 1.13, P = 0.12, I^2 = 0%, heterogeneity P = 0.96), and 0.27 (95%CI; 0.11 to 0.69, P = 0.006, I^2 = 0%, heterogeneity P = 0.78), respectively. • Based on the fixed model, TRA showed a better safety profile compared to BIM overall, with overall OR 0.73 (95% CI; 0.56 to 0.95, P = 0.02, I^2 = 0%, heterogeneity P = 0.64), ORs for conjunctival hyperemia, discomfort, and growth of lashes 0.64 (95%CI; 0.46 to 0.88, P = 0.007, I^2 = 0%, heterogeneity P = 0.89), 1.01 (95% CI; 0.62 to 1.65, P = 0.97, I^2 = 16%, heterogeneity P = 0.31), and 0.59 (95%CI; 0.14 to 2.50, P = 0.47, I^2 =39%, heterogeneity P = 0.20), respectively. • A total of 508 patients received BIM and reported ocular AEs. There were 259 (50.98%) reported cases of ocular AEs, including 204 (40.16%), 35 (6.89%), and 20 (3.94%) cases of conjunctival hyperemia, discomfort, and growth of lashes, respectively.	"*0.03% bimatoprost appears more effective following long time use (3 and 6 month post-treatment) for IOP control compared to 0.005% latanoprost, and is more effective compared to 0.004% travoprost after being used for a certain period of time (3 months post-treatment); nevertheless, 0.005% latanoprost is better tolerated in patients with POAG or OHT*" pg 1

Main Study Findings	Authors' Conclusion
Rennie et al. 2019[16]	
Efficacy: WMD of the post-intervention mean OPP was calculated. Results comparing BIM to other PGAs are presented only. There was no statistically significant difference in mean OPP at the end of treatment with comparisons of prostaglandins, including comparisons involving BIM. • BIM compared to LAT had a mean OPP difference of 0.00mmHg (1 trial, 40 patients, 95% CI; -4.06 to 4.06, test for heterogeneity NA) **Safety:** Relative risk and SE was used for dichotomous harms data. There was no statistically significant difference in the risk of developing AEs between LAT and BIM, RR (SE) 0.77 (0.21) P = 0.2	"*Bimatoprost increases mean ocular perfusion pressure when compared to timolol. As a class, prostaglandins increase mean ocular perfusion pressure.* *Prostaglandins may provide beneficial ocular perfusion pressure profiles compared to alternative agents*" pg 2
Diaconita et al. 2018[17]	
The mean and SD of pre and post-washout IOP was extracted for all included studies to compute the mean IOP reduction and % of IOP reduction, which was then used to calculate weighted SMD as the treatment effect. • Meta-analysis based on 8 studies (including 5 RCTs) revealed an average 4-week discontinuation of LAT leads to a return to patients' baseline IOP. • One RCT compared the washout duration of LAT and BIM in 73 patients previously treated with IOP-lowering medication and found similar pre and post-washout IOP at 4 weeks with no statistically significant change after washout.	"*A significant IOP-lowering effect of latanoprost was not observed beyond 4 weeks, suggesting this may be an appropriate washout period for latanoprost. We could not identify appropriate washout periods for either travoprost or bimatoprost, although a majority of articles had 4-week washout durations for the two drugs*" pg 1
Takagi et al. 2018[18]	
The mean and SD/SE of IOP reduction at postdosing final evaluation point was extracted or calculated. • Based on 3 articles for LAT, 2 articles for TRA, 1 article for BIM, and 5 articles for TAF, the mean (SD) change in IOP was -3.03 (2.18) mmHg with LAT, -3.33 (1.79) mmHg with TRA, -3.05 (1.99) mmHg with TAF, and -4.00 (1.90) mmHg with BIM. Based on this, BIM had the strongest ocular hypotensive efficacy and LAT, TRA, and TAF had similar levels of ocular hypotensive efficacy and all three were weaker than BIM. • Based on 14 articles for LAT, 7 articles for TRA, 3 articles for BIM, and 7 articles for TAF, a linear regression analysis was performed which showed the r^2 values ranged between 0.85 (LAT) and 0.93 (TRA), with similar slopes for all PGAs. The steepness of the slopes revealed BIM had the strongest ocular hypotensive profile, followed by similar efficacy of TAF and TRA, and LAT was the weakest. • The percentage IOP reductions ranged between 15% and 20% with different PGAs. • The ocular hypotensive efficacy of the PGAs varied according to pretreatment IOP values. At lower predosing IOP values (12-15 mmHg), the ocular hypotensive efficacies of TAF and TRA were greater than that of LAT; whereas the IOP reduction was almost same with all three PGAs at higher predosing IOP (17–18 mmHg).	"*In the rank order of IOP-lowering efficacy of PGAs, bimatoprost was the strongest and latanoprost the weakest. Travoprost and tafluprost had almost the same level of ocular hypotensive effect, and both were stronger than latanoprost. The scatter plot analysis showed that all PGAs reduced IOP by 15%–20%. At higher IOP (17–18 mmHg), the ocular hypotensive effect was almost the same with latanoprost, travoprost, and tafluprost. In contrast, at lower IOP (12–15 mmHg), the IOP reduction with latanoprost was weaker than with travoprost or tafluprost.*" pg 1837
Li et al. 2016[19] USA	
• Based on direct comparison between PGAs, the mean difference between treatment groups at 3 months were reported, with a mean	"*Bimatoprost, latanoprost, and travoprost are among the most efficacious drugs, although the*

Main Study Findings	Authors' Conclusion
difference of > 0 indicative of a greater IOP reduction by the first drug, and therefore is more efficacious than the comparator drug. Conversely, a mean difference of < 0 favors the drug on the right. ○ BIM vs LAT (6 studies): 0.87 (95% CI; 0.01 to 1.73, I^2 = 76%) ○ BIM vs TRA (8 studies): 0.59 (95% CI; −0.13 to 1.30, I^2 = 74%) ○ LAT vs TRA (7 studies): −0.06 (95% CI; −0.46 to 0.34, I^2 = 0%) ○ LAT vs TAF (1 study): −0.90 (95% CI; −3.40 to 1.60, I^2 = NA) • The mean reductions from baseline (95% CI) in IOP in mmHg at 3 months for BIM, LAT, TRA, TAF were 5·61 (4·94 to 6·29), 4·85 (4·24 to 5·46), 4·83 (4·12 to 5·54), and 4·37 (2·94 to 5·83), respectively. • Based on indirect comparison between PGAs, the mean difference between treatment groups at 3 months were reported, with a mean difference of < 0 favoring the drug on the left; whereas a mean difference > 0 favoring the drug on the right. Results from indirect multiple-treatment comparison: ○ BIM vs LAT: −0.77 (95% CrI; −1.26 to −0.27) ○ BIM vs TRA: −0.78 (95% CrI; −1.3 to −0.26) ○ BIM vs TAF: −1.24 (95% CrI; −2.65 to 0.18) ○ LAT vs TRA: −0.02 (95% CrI; −0.53 to 0.5) ○ LAT vs TAF: −0.48 (95% CrI; −1.83 to 0.91) ○ TRA vs TAF: −0.46 (95% CrI; −1.87 to 0.98)	*within class differences were small and may not be clinically meaningful*" pg 2

AE = adverse events; BIM = bimatoprost; CI = confidence interval; CrI = credible interval; DB=double-blind; IOP = intraocular pressure; LAT=latanoprost; OHT = ocular hypertension; OPP = ocular perfusion pressure; OR = odds ratio; P/OAG = Primary / open-angle glaucoma; PGA = prostaglandin analogue; RCT = randomized controlled trial; RR = relative risk; SD/SE = standard deviation/error' SMD = standardized mean difference; TAF = tafluprost; TRA=travoprost; WMD = weighted mean difference

Table 9: Summary of Findings of Included Primary Clinical Studies

Main Study Findings	Authors' Conclusion
Blondeau et al. 2019[20]	
• The mean IOP at baseline was 23.7±4.7 mmHg. The mean IOP among non-responders at randomization was 21.5±4.5 mmHg. There was no statistically significant difference in the mean IOP reduction between the treatment groups, and were -0.9 mmHg, -2.1 mmHg, and -2,5 mmHg in the LAT, BIM and TRA, respectively (P = 0.15). • Of the 83 non-responders, 32 (38.5%) became new responders (> 20% IOP reduction) 1 month after treatment, with a similar proportion of responders across the treatment groups, 9 (31%), 13 (41.9%) and 10 (43.5%) in LAT, BIM, and TRA, respectively (P = 0.58).	"*There is no added benefit of switching latanoprost to another topical prostaglandin for patients who are initially non-responders*" pg 895
Nazir et al. 2019[21]	
• Among male patients, the difference between baseline and 30-day IOP in the right eye was 6.75±0.52 and 7.58±0.24 in the LAT and BIM group, respectively (within group P = 0.001) and 6.6±0.7 and 7.28±0.38 in the left eye between the two groups, respectively (within group P = 0.0031). • Among female patients, the difference between baseline and 30-day IOP in the right eye was 6.18±0.01 and 7.06±0.06 in the LAT and BIM group, respectively (within group P = 0.001) and 6.36±0.1 and 8.0±0.31 in the left eye between the two groups, respectively(within group P = 0.001). • Difference in IOP between baseline and at day 30 for both eyes were analyzed by age group (20-35 years, and 36-50 years), and showed a consistent pattern favoring BIM (results not presented).	"*….mean change in reducing the intraocular pressure with Bimatoprost 0.01% is more significant than Latanoprost 0.05%*" pg 17

Main Study Findings	Authors' Conclusion

Moussa et al. 2018[22]

Efficacy:

- All four drugs equally and statistically significantly reduced the IOP from baseline at each follow-up visits with respect to baseline (P < 0. 01).
- There was no statistically significant difference in IOP reduction among the PGAs (P = 0.112).

IOP mmHg (mean±SD)	BIM (n=8)	LAT (n=7)	TRA (n=8)	TAF (n=9)	P value
IOP at inclusion	26.13±6.15	24.71±2.36	24.38±1.85	25.22±2.28	0. 79
IOP at 1-month (% reduction)	17.12±3.42 (34.48%)	16.32±2.96 (33.95%)	16.32±2.01 (33.06%)	17.22±3.09 (31.72%)	P<0.01
IOP at 3-month	15.58±4.20 (39.34%)	17.02±3.76 (31.12%)	17.43±1.32 (28.51%)	16.97±2.07 (32.71%)	P<0.01
IOP at 6-month	15.50±2.93 (40.68%)	17.43±2.57 (29.46%)	16.88±1.13 (30.76%)	18.11±2.42 (28.19%)	P<0.01

P values comparing IOP at each follow-up time compared to baseline

Safety:

- At month 6, the OSDI score was statistically significantly superior for TRA (10.68 ± 5.73) compared to the other three drugs, 21.76 ± 11.10 for BIM, 32.13 ± 24.10 for LAT, and 25.60 ± 6.25 for TAF (P < 0.05).
- LAT caused the most significant eyelash growth and iris discoloration.
- Conjunctival hyperemia, follicular conjunctivitis and superficial keratitis occurrence were similar in the four groups, a significant difference between the four groups was not found.
- None of the patients developed prostaglandin-associated orbitopathy or any severe acute side effects (acute allergy and angioedema).

AE, n (%)	BIM (n=8)	LAT (n=7)	TRA (n=8)	TAF (n=9)	P value
Conjunctival hyperemia	6 (75.0)	5 (71.4)	4 (50.0)	9 (100.0)	0.13
Superficial keratitis	4 (50.0)	4 (57.1)	3 (3.8)	3 (33.3)	0.76
Follicular conjunctivitis	2 (25.0)	4 (57.1)	3 (3.8)	5 (55.5)	0.51
Iris hyperpigmentation	0 (0)	2 (28.6)	0 (0)	0 (0)	0.05
Eyelash growth	1 (12.5)	4 (57.1)	2 (25.0)	0 (0)	0.05
Herpetic reactivation	0 (0)	1 (14.3)	0 (0)	0 (0)	0.30
OSDI at 6-month	21.76±11.10	32.13±24.10	10.68±5.73	25.60±6.25	

P values comparing incidence of AEs between treatment groups

"*All prostaglandin analogues equally and significantly reduce the IOP in patients with POAG. According to the results of the OSDI score, latanoprost seems to be the least tolerated among the four drugs*" pg 1

"*Latanoprost was more associated with iris pigmentation and eyelash growth*" pg 5

Guven Yilmaz et al. 2018[23]

- All three groups reached target level of IOP ≤ 21 mmHg before 24-hr IOP monitorization. The mean 24-hr IOP did not differ significantly between the three groups, BIM: 14.6 ± 2.9 mmHg; LAT: 14.1 ± 3.7 mmHg and TRA: 15.8 ± 2.0 mmHg, P > 0.05.
- There were no statistically significant differences in diurnal or nocturnal IOP variation between the three groups (P > 0.05). Mean diurnal IOP variation was 4.6 ± 2.3 mmHg, 5.8 ± 2.4 mmHg, and 4.3 ± 1.7 mmHg in the BIM, LAT, and TRA group, respectively; and mean nocturnal IOP variation was 3.2 ± 2.8 mmHg, 2.9 ± 1.9 mmHg, and 3.0 ± 1.6 mmHg in the three groups.

"*All three fixed combinations effectively controlled IOP for 24-h and had a similar effect on diurnal and nocturnal IOP variations*" pg 1425

Main Study Findings	Authors' Conclusion

- IOP was highest at 02:00 in all three groups, after which time it consistently decreased. IOP in the BIM/TIM and LAT/TIM group was lowest at 14:00 versus 18:00 in the TRA/TIM group. While the BIM/TIM group provided significantly lower IOP at 10:00 and 14:00 versus the other two groups, the LAT/TIM group demonstrated significantly better IOP-lowering efficacy at 02:00 compared with the other two groups.

IOP mmHg (mean±SD)	BIM/TIM (n=18)	LAT/TIM (n=14)	TRA/TIM (n=18)	P value
IOP at baseline	13.5 ± 4.2	14.9 ± 2.5	14.4 ± 2.8	0.39
24-hr IOP	14.15 ± 3.63	13.98 ± 2.56	15.45 ± 1.92	0.29
Diurnal IOP	4.16 ± 2.97	5.91 ± 4.05	4.11 ± 2.86	0.27
Nocturnal IOP	4.38 ± 2.93	4.41 ± 4.37	5.41 ± 3.01	0.62
IOP levels through 24-hr				
At 10.00	12.83 ± 4.84	15.17 ± 3.53	15.76 ± 3.15	0.02
At 14.00	12.22 ± 2.77	13.08 ± 3.20	14.94 ± 1.98	0.003
At 18.00	13.61 ± 4.56	13.83 ± 4.66	13.29 ± 2.25	0.90
At 22.00	14.83 ± 4.14	13.25 ± 3.69	14.12 ± 3.60	0.55
At 02.00	16.89 ± 4.93	14.16 ± 4.02	18.41 ± 3.30	0.02

P values comparing mean IOP levels between treatment groups

Stalmans et al. 2016[24]

Efficacy:

- There was a statistically significant difference in IOP for both drugs at 3 and at 6 months compared to baseline**:** −4.0 ± 0.5 mmHg for both BIM/LAT and LAT/BIM at 3 months (P < 0.01 for both drugs); −5.2 ± 0.5 and −3.4 ± 0.5 mmHg for BIM/LAT and LAT/BIM, respectively (both P < 0.01), at 6 months.
- At month 6, the estimated mean IOP was 1.6 ± 0.5 mmHg higher in patients treated with LAT compared with BIM (P < 0.01).
- The intra-patient mean IOP was 0.9 ± 0.2 mmHg higher with LAT comparedwith BIM (P < 0.01)
- The difference between the two treatments at the end of 3 months was not statistically significant (P = 0.32), the estimated difference was 0.5 mmHg and the two mean IOPs for BIM/LAT and LAT/BIM were very similar (15.3 ± 2.2 and 15.1 ± 2.2, respectively).

	IOP at screening	IOP at baseline	IOP at 3 months (uncorrected)	IOP at 3 months (corrected)	IOP at 6 months (uncorrected)	IOP at 6 months (corrected)
BIM/LAT (n = 33)	15.18±2.26	19.2±3.32	15.3±2.22	14.64±0.65	15.9±2.80	14.06±0.5
LAT/BIM (n = 34)	15.09±2.75	19.2±4.61	15.1±2.21	15.20±0.64	13.9±2.61	15.66±0.5

Safety:

- Visual field tests, BCVA measurements, and cup-to-disk ratio evaluation were performed throughout the study. No significant changes were detected between the beginning and the end of the study for any of these measures.
- No relevant side effects or AEs were encountered throughout the study.
- Hyperemia scores were mild in both treatment groups; although the difference between BIM/LAT and LAT/BIM was statistically significant (0.85 ± 0.49 vs 0.71± 0.45, respectively), with a mean intra-patient difference of 0.14 (P < 0.05).

"*This study demonstrates a superior efficacy of [BIM] over [LAT] in lowering IOP. The results are consistent both in the parallel comparison between the two treatment groups at 6 months as well as in the intra-subject pressure comparison*" pg 1152

Main Study Findings	Authors' Conclusion

AE, mean scores	Pre-trial	Washout	BIM/LAT	LAT/BIM
Lids	1.03	1.05	1.02	1.02
Cornea	1	1	1.02	1.02
Conjunctiva	1.02	1	0.99	1.02
Hyperemia score	0.92	0.77	0.85	0.71
Iris	1	1	1	1
Anterior chamber	1	1.02	1	1
Lens	1.58	1.58	1.55	1.55

Maruyama et al. 2015[25]

Efficacy:

- There was a statistically significant decrease in mean IOP at 4 weeks compared with week 0 in both groups (both P < 0.0001). However, there were no IOP differences at 12 weeks compared with the values at 4 weeks in both groups (P = 0.85 for the BIM group and P = 0.84 for the LAT/TIM group)
- Comparisons between the two groups showed no statistically significant differences at any timepoints. Similarly, a comparison for the ratio of patients with % in IOP changes showed no statistically significant differences in any of the ratios between the treatment groups.

IOP±SD (mmHg)	LAT/TIM (n=27)	BIM (n=30)	P value
Baseline	13.3±2.7	13.2±3.0	0.72
Week 4	11.5±2.9	11.6±2.7	0.67
Week 12	11.6±2.7	11.6±2.5	0.91
% IOP change, n (%)			
> 10% IOP increase	1 (3.7%)	1 (3.3%)	1.00
±10% IOP change	10 (37%)	13 (43.3%)	0.79
10% < IOP ≤ 20%	7 (25.9%)	9 (30%)	0.78
20% < IOP ≤ 30%	6 (22.2%)	4 (13.3%)	0.49
>30% IOP reduction	3 (11.1%)	3 (10%)	1.00

P values comparing mean IOP levels between treatment groups

Safety

- The conjunctival injection score was higher in the BIM group than in the LAT/TIM group at 12 weeks (P=0.0091).
- There were no statistically significant differences between the two drugs with respect to AD score, tear film break-up time, heart rate, and blood pressure at any timepoint.

	Baseline			Week 4			Week 12		
Assessment	LAT/TIM (n=27)	BIM (n=30)	P value	LAT/TIM (n=27)	BIM (n=30)	P value	LAT/TIM (n=27)	BIM (n=30)	P value
Injection	0.7±0.5	0.6±0.5	0.50	0.7±0.6	0.9±0.7	0.40	0.6±0.5	1.2±0.8	0.0091
AD score	1.6±1.4	1.9±1.5	0.63	1.6±1.4	1.5±1.4	0.70	1.7±1.4	1.6±1.4	0.82
BUT (seconds)	5.2±4.0	6.0±4.4	0.57	5.0±4.5	6.0±4.7	0.42	5.2±3.9	5.4±4.3	0.82
HR (BPM)	77±13	80±12	0.21	76±13	79±11	0.37	77±15	77±10	0.88
SBP (mmHg)	130±23	127±23	0.40	128±24	127±24	0.83	129±23	129±23	0.91
DBP (mmHg)	75±15	73±13	0.29	76±15	74±13	0.30	74±15	75±13	0.47

P values comparing the outcome between treatment groups

"*Bimatoprost and [LAT/TIM] exhibited similar efficacy for reduction of IOP. Safety results indicated that only the conjunctival injection score at 12 weeks was higher in the bimatoprost group compared with the [LAT/TIM] group*" pg 1429

Rossetti et al. 2015[26]

- The IOP levels at baseline (i.e. under the unfixed combination of LAT and TIM) and at 12 weeks were very similar in the two groups and no statistically significant difference was shown at any comparison.

"*Bimatoprost and [LAT/TIM] had similar DBPs and OPPs; SBP was significantly lower with [LAT/TIM]. In this study, the percentage of "dippers"*

Main Study Findings	Authors' Conclusion
• SBP and DBP were very similar at baseline in the two groups and, as expected, were significantly lower during the night (P = 0.01). • Holter SBP and SPP that were significantly higher in the BIM group compared to LAT/TIM (135.1 mmHg vs 128.1 mmHg and 119.0 mmHg vs 111.8 mmHg, respectively, P = 0.04, p = 0.03).	*was considerably higher than the one described in previous studies on the role of perfusion pressure in glaucoma*" pg 1

	BIM (n=101)	LAT/TIM (n=99)	P value
Baseline IOP (sd)	16.3 (3.3)	15.5 (2.9)	0.2
12-week IOP (sd)	16.1 (2.5)	16.3 (3.7)	0.7
Baseline nocturnal IOP (sd)	16.9 (3.6)	16.0 (3.3)	0.3
12-week nocturnal IOP (sd)	16.1 (2.6)	16.1 (3.9)	0.8
Baseline SBP (sd)	136.5 (18.3)	134.2 (20.1)	0.1
Holter SBP (sd)	135.1 (16.7)	128.1 (15.3)	0.04
Baseline DBP (sd)	79.1 (10.2)	78.2 (10.1)	0.4
Holter DBP (sd)	79.5 (8.3)	78.7 (11.8)	0.4
Baseline nocturnal SBP (sd)	121.0 (13.8)	122.1 (15.8)	0.3
Holter nocturnal SBP (sd)	124.8 (14.4)	120.0 (14.5)	0.08
Baseline nocturnal DBP (sd)	72.7 (7.9)	73.2 (9.5)	0.4
Holter nocturnal DBP (sd)	71.7 (7.9)	70.6 (11.3)	0.2
Baseline SPP (sd)	120.2 (15.7)	118.7 (16.8)	0.3
Holter SPP (sd)	119.0 (10.8)	111.8 (15.3)	0.03
Baseline DPP (sd)	62.8 (6.9)	62.7 (8.2)	0.6
Holter DPP (sd)	63.4 (8.0)	62.4 (11.1)	0.1
Baseline nocturnal SPP (sd)	104.1 (13.1)	106.1 (16.4)	0.2
Holter nocturnal SPP (sd)	108.7 (14.4)	103.9 (17.3)	0.07
Baseline nocturnal DPP (sd)	55.8 (8.0)	57.2 (12.1)	0.1
Holter nocturnal DPP (sd)	55.6 (7.4)	54.5 (12.3)	0.2

P values comparing mean values of the outcome between treatment groups

AD = area density classification; BIM = bimatoprost; BUT = tear film break-up time; DBP = diastolic blood pressure; DPP = diastolic perfusion pressure; IOP = intraocular pressure; LAT=latanoprost; OHT = ocular hypertension; OPP = ocular perfusion pressure; OSDI = Ocular Surface Disease Index; P/OAG = Primary / open-angle glaucoma; PGA = prostaglandin analogue; RCT = randomized controlled trial; RR = relative risk; SBP = systolic blood pressure; SPP =systolic perfusion pressure; SD/SE = standard deviation/error; SMD = standardized mean difference; TAF = tafluprost; TRA = travoprost; WMD = weighted mean difference

Table 10: Summary of Findings of Included Economic Evaluations

Main Study Findings	Authors' Conclusion
Brown and Brown, 2019[27]	
• BIM conferred a mean 2.56 QALY gain (22.9% patient quality-of-life gain). • LAT conferred a 2.00 QALY gain (17.8% quality-of-life gain). • TAF a 1.99 QALY gain (17.9% quality-of-life gain) • TRA a l.92 QALY gain (17.2% quality-of-life gain). • TIM a 1.42 QALY gain (12.8% quality-of-life gain). • The ophthalmic cost-perspective, incremental cost-utility ratio of bimatoprost referent to travoprost was $6,034/QALY, to latanoprost was $27,973/QALY, and to timolol was $16,063/QALY. Bimatoprost dominated tafluprost, meaning that it conferred greater patient value for lesser cost than tafluprost. The QALY gain of each drug versus no therapy is shown in the table below.	"*Topical bimatoprost delivers greater patient value than the other prostamides and topical timolol for the treatment of OAG. Bimatoprost is incrementally cost-effective referent to the other prostamides and timolol*" pg 145

Societal cost perspective, ICER referent to BIM				
Drug	QALY gain	QOL (value) gain	Societal costs	Incremental $/QALY
BIM	2.56	22.9%	(-$446,289)	NA
LAT	2.00	17.8%	(-$413,738)	BIM dominant

Main Study Findings	Authors' Conclusion

TAF	1.99	17.9%	(-$345,803)	BIM dominant
TRA	1.92	17.2%	(-$386,280)	BIM dominant
TIM	1.42	12.8%	(-$323,286)	BIM dominant
Third party insurer cost perspective, ICER referent to BIM				
Drug	QALY gain	QOL (value) gain	Third party insurer costs	Incremental $/QALY
BIM	2.56	22.9%	(-$51,487)	NA
TAF	1.99	17.9%	(-$26,876)	BIM dominant
TRA	1.92	17.2%	(-$45,856)	BIM dominant
TIM	1.42	(-$47,491)	(-$47,491)	BIM dominant
LAT	2.00	17.8%	(-$60,074)	$15,334
Ophthalmic cost perspective, incremental cost-utility ratios referent to BIM				
Drug	QALY gain	QOL (value) gain	Ophthalmic costs	Incremental $/QALY
BIM	2.56	22.9%	$27,358	NA
TAF	1.99	17.9%	$38,607	BIM dominant
TRA	1.92	17.2%	$23,569	$6,034
TIM	1.42	12.8%	$10,171	$16,063
LAT	2.00	17.8%	$11,691	$27,973

The average cost-utility ratio compares therapy to no therapy; the incremental cost-utility ratio compares the drug conferring the greatest patient value (bimatoprost) to the other drugs.
Minus values are in parentheses.
When a drug is dominant, it confers greater patient value and has a less expensive price than a comparator drug.

The 20-year societal (direct medical, direct nonmedical, and indirect medical) costs in 2015 U.S. real dollars are shown in the table below.

- The 20-year drug cost and total glaucoma (ophthalmic) medical costs was the lowest for TIM and highest for TAF.
- Third party insurance cost, and nonophthalmic cost varied between drugs.
- Total societal costs (including the ophthalmic direct medical costs of glaucoma therapy) showed each drug conferring a large net savings for society versus no therapy, ranging from (-$446,289) for BIM therapy to (-$323,386) for TIM therapy.
- Based on the annual, societal ROI for each of the drugs for the direct glaucoma medical costs expended, TIM and LAT showed the highest annual and 20-year ROI. BIM showed 15.3% or 1631% annual and 20-year ROI, respectively.

Costs	TIM	LAT	TRA	BIM	TAF
Drug cost	$1,070	$2,590	$14,468	$18,257	$29,505
Direct ophthalmic (glaucoma) medical costs	$9,101	$9,101	$9,101	$9,101	$9,101
Total direct glaucoma medical costs	$10,171	$11,691	$23,569	$27,358	$38,607
Nonophthalmic direct medical costs (depression, trauma, etc.)	(-$57,662)	(-$71,765)	(-$69,425)	(-$78,846)	(-$65,483)

Main Study Findings					Authors' Conclusion
Third party insurer (direct medical oph + nonophthalmic	(-$47,491)	(-$60,074)	(-$45,856)	(-$51,487)	(-$26,876)
Caregiver	($241,976)	(-$312,775)	(-$300,179)	(-$350,244)	(-$281,190)
Nursing home	(-$22,301)	(-$27,925)	(-$26,942)	(-$30,901)	(-$24,441)
Employment	(-$11,517)	(-$12,964)	(-$13,304)	(-$13,657)	(-$12,296)
Total societal costs	(-$323,286)	(-$413,738)	(-$386,280)	(-$446,289)	(-$345,803)
Annual financial ROI for the direct medical glaucoma costs expended	19.1%	19.7%	15.3%	15.3%	2.2%
Total, 20-year, financial ROI for the direct medical glaucoma costs expended	3,178%	3,539%	1,639%	1,631%	896%

Negative costs in parentheses () indicate dollars returned to patients, the government, and insurers from the treatment of POAG. The better vision associated with glaucoma therapy obviates many of these costs, especially those associated with end-stage glaucoma. These costs, which accrue against the direct ophthalmic medical costs, do not accrue until 20/70 vision first appears in association with visual field loss at 24 months before end-stage glaucoma.
Direct, nonophthalmic medical costs obviated by timolol therapy maintaining good vision long-term include those for: depression, injury, skilled Nursing Facility admissions, and other, as yet, unidentified medical costs.

The 20-year, macroeconomic costs saved by a 1-year cohort of glaucoma patients treated with a prostamide (vs. no treatment) are shown in the table below, as is the distribution of savings.

- The greatest savings occurred with BIM. The greatest beneficiaries of the savings were patients, with over 80% of savings accrued to patients, no matter the drug used.

2015 U.S. Real Dollars Saved by Glaucoma Therapy Over 20 Years by a One-Year Glaucoma Cohort (66,880 People) Aged 65 Years							
Drug	Per patient savings	Total U.S. Savings* (millions)	Medicare savings (millions)	Medicaid savings (millions)	Commercial insurance savings (millions)	Other savings (millions)	Patient Savings** (millions)
BIM	446,289	29,848	2,865	1,442	459	165	24,930
LAT	413,738	27,671	3,267	1,299	522	193	22,357
TRA	386,280	25,835	2,546	1,229	399	147	21,484
TAF	345,803	23,127	1,583	1,127	246 86	86	20,075
TIM	323,386	21,621	2,585	1,037	421	153	17,409
Percentage of total dollars saved							
Drug	Per patient savings	Total. Savings* (66,880 people)	Medicare savings	Medicaid savings	Commercial insurance savings	Other savings	Patient Savings** (% of all savings)
BIM	100%	100%	9.6%	4.8%	1.5%	0.6%	83.5%
LAT	100%	100%	11.8%	4.7%	1.9%	0.7%	80.9%

Main Study Findings							Authors' Conclusion
TRA	100%	100%	9.9%	4.8%	1.5%	0.6%	83.3%
TAF	100%	100%	6.9%	4.9%	1.0%	0.4%	86.9%
TIM	100%	100%	12.0%	4.8%	1.9%	0.7%	80.6%

*The total savings include the 20-year savings conferred by treatment with each of the drugs for a calculated U.S. annual cohort of 66,880 new patients with open angle glaucoma and a mean age of 65 years.48

**Patient savings occur in the form of decreased caregiver costs, decreased out-of-pocket costs for medical care and decreased employment (salary loss) costs.

The mean patient saving was 83.0% of all savings and the average Medicare savings was 10.04% of all savings. The average saving per patient from prostamide usage, calculated by averaging the per patient saving for bimatoprost, tafluprost, travoprost, and latanoprost was $383,099.

BIM = bimatoprost; LAT=latanoprost; QALY = quality adjusted life years; QoL = quality of life; ROI = return on investment;TAF = tafluprost; TRA = travoprost

Appendix 5: Additional References of Potential Interest

Non-randomized studies excluded from the report

Deshpande SS, Sonty S, Ahmad A. Evaluating intraocular pressure-lowering solutions for the treatment of open-angle glaucoma: comparison between bimatoprost 0.03% and bimatoprost 0.01% - an observational switch study. *Clin Ophthalmol*. 2017 Jul 27;11:1371-1376.

El Ameen A, Vandermeer G, Khanna RK, Pisella PJ. Objective ocular surface tolerance in patients with glaucoma treated with topical preserved or unpreserved prostaglandin analogues. *Eur J Ophthalmol*. 2019 Nov;29(6):645-653.

Heo JH, Rascati KL, Wilson JP, Lawson KA, Richards KM, Nair R. Comparison of prostaglandin analog treatment patterns in glaucoma and ocular hypertension. *J Manag Care Spec Pharm*. 2019 Sep;25(9):1001-1010.

Kara C, Şen EM, Elgin K, Serdar K3, Yilmazbaş P. Does the intraocular pressure-lowering effect of prostaglandin analogues continue over the long term? *Int Ophthalmol*. 2017 Jun;37(3):619-626.

Kim HW, Choi YJ, Lee KW, Lee MJ. Periorbital changes associated with prostaglandin analogs in Korean patients. *BMC Ophthalmol*. 2017 Jul 17;17(1):126.

Ohyama K, Kawakami H, Inoue M. Blood pressure elevation associated with topical prostaglandin F2α analogs: an analysis of the different spontaneous adverse event report databases. *Biol Pharm Bull*. 2017;40(5):616-620.

Rodríguez-Agramonte F, Jiménez JC, Montes JR. Periorbital changes associated with topical prostaglandins analogues in a Hispanic population. *P R Health Sci J*. 2017 Dec;36(4):218-222.

Tamçelik N, Izgi B, Temel A, Yildirim N, Okka M, Özcan A, Yüksel N, Elgin U, Altan Ç, Ozer B. Prospective, non-interventional, multicenter study of the intraocular pressure-lowering effects of prostaglandin analog/prostamide-containing therapies in previously treated patients with open-angle glaucoma or ocular hypertension. *Clin Ophthalmol*. 2017 Apr 19;11:723-731.