

Concept Paper

Not peer-reviewed version

Metformin for Age-Related Macular Degeneration: Moving Beyond Observational Studies to Causal Inference Through Target Trial Emulation and Advanced Analytics

[Amr Ahmed](#)*

Posted Date: 4 March 2026

doi: 10.20944/preprints202603.0202.v1

Keywords: metformin; age-related macular degeneration; causal inference; target trial emulation; propensity score matching; causal artificial intelligence; observational studies



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Concept Paper

Metformin for Age-Related Macular Degeneration: Moving Beyond Observational Studies to Causal Inference Through Target Trial Emulation and Advanced Analytics

Amr Ahmed

Public Health Department, Riyadh First Health Cluster, Ministry of Health, Riyadh, Saudi Arabia

Abstract

Background: Age-related macular degeneration (AMD) represents the leading cause of irreversible vision loss in elderly populations globally. While metformin has emerged as a promising candidate for AMD prevention based on multiple observational studies, the causal relationship remains uncertain due to inherent limitations of observational research designs. **Objective:** This comprehensive review critically evaluates the current evidence base for metformin in AMD prevention and treatment, with particular emphasis on methodological approaches that address causal inference, including target trial emulation frameworks, propensity score methods, and emerging applications of causal artificial intelligence. **Methods:** We conducted a systematic review of recent literature (2019-2025) focusing on studies employing advanced causal inference methodologies. Particular attention was given to the largest meta-analysis to date (2.68 million participants) and studies utilizing target trial emulation, propensity score matching, instrumental variable analysis, and causal AI approaches. **Results:** Recent meta-analytic evidence demonstrates a statistically significant protective association (pooled OR = 0.86, 95% CI: 0.79–0.93, $p < 0.001$) between metformin use and AMD development across 18 observational studies [1]. However, substantial heterogeneity ($I^2 = 90\%$) and inherent biases in observational designs—including immortal time bias, disease latency bias, and confounding by indication—limit causal interpretation. Studies employing propensity score matching and dose-response analyses reveal protective effects primarily at low-to-moderate cumulative doses (270–600g over 2 years) [2]. Critically, no adequately powered randomized controlled trial has yet definitively established causality. **Conclusions:** While observational evidence suggests potential benefit, the causal effect of metformin on AMD prevention remains unproven. Rigorous application of target trial emulation frameworks, coupled with advanced causal AI methodologies, offers a pathway to strengthen causal inference from existing observational data. However, definitive evidence requires prospective randomized trials specifically designed to test metformin's efficacy in non-diabetic populations at risk for AMD.

Keywords: metformin; age-related macular degeneration; causal inference; target trial emulation; propensity score matching; causal artificial intelligence; observational studies

1. INTRODUCTION

Age-related macular degeneration (AMD) constitutes a major public health challenge, affecting approximately 200 million individuals worldwide with projections suggesting this figure will reach 288 million by 2040 [1]. As the leading cause of irreversible blindness in individuals over 55 years in developed nations, AMD imposes substantial burdens on healthcare systems and significantly diminishes quality of life for affected individuals. Current therapeutic interventions remain limited, with anti-VEGF therapy addressing only the neovascular (wet) form of AMD, while no effective treatment exists for the more prevalent geographic atrophy (dry AMD).

Metformin, a biguanide antihyperglycemic agent widely prescribed for type 2 diabetes mellitus, has garnered increasing attention for potential therapeutic applications beyond glycemic control [9]. Mechanistic studies demonstrate that metformin exerts pleiotropic effects relevant to AMD pathophysiology, including activation of AMP-activated protein kinase (AMPK), enhancement of autophagy in retinal pigment epithelium cells, reduction of oxidative stress, anti-inflammatory properties through modulation of NF- κ B signaling, and inhibition of pathological angiogenesis. These biological mechanisms provide a compelling rationale for investigating metformin as a potential preventive or therapeutic agent for AMD.

1.1. The Causal Inference Challenge in Observational AMD Research

A critical distinction must be drawn between *association* and *causation*. While numerous observational studies have documented associations between metformin use and reduced AMD incidence [2,3,4,8,9], establishing causality requires overcoming several fundamental challenges inherent to observational research:

- Confounding by indication: Patients prescribed metformin differ systematically from non-users in baseline characteristics, comorbidities, and healthcare utilization patterns.
- Immortal time bias: In retrospective cohort studies, patients must survive long enough to receive metformin, creating artificial survival advantages.
- Disease latency bias: AMD manifests subclinical changes years before clinical diagnosis, potentially affecting medication prescribing patterns before formal AMD diagnosis [7].
- Healthy user bias: Metformin users may engage in healthier behaviors and have better healthcare access compared to non-users.
- Depletion of susceptibles: Long-term metformin users represent a selected population that has tolerated the medication, introducing selection bias.

These biases cannot be fully addressed through conventional statistical adjustment or propensity score methods alone, as they may not capture unmeasured confounders. This necessitates more sophisticated causal inference frameworks, particularly target trial emulation and advanced causal AI approaches, to approximate randomized trial conditions using observational data [5,6].

2. EVIDENCE FROM OBSERVATIONAL STUDIES

2.1. Recent Meta-Analytic Evidence

The most comprehensive synthesis to date, published in 2025, analyzed 18 observational studies encompassing 2,683,234 individuals [1]. This meta-analysis demonstrated a statistically significant protective association between metformin use and AMD development (pooled OR = 0.86, 95% CI: 0.79–0.93, $p = 0.0002$) [1]. However, substantial heterogeneity across studies ($I^2 = 90\%$) suggests considerable variability in effect estimates, potentially reflecting differences in study populations, metformin dosing regimens, AMD definitions, and unmeasured confounding.

Notably, when examining AMD progression (rather than incidence), the pooled hazard ratio was 0.82 (95% CI: 0.55–1.23, $p = 0.34$), failing to reach statistical significance with even greater heterogeneity ($I^2 = 99\%$) [1]. This divergence between incidence and progression outcomes warrants careful interpretation and suggests that metformin's potential protective effects may be limited to early AMD stages or that methodological limitations obscure progression effects.

2.2. Dose-Response Relationships

Multiple large-scale case-control studies have revealed intriguing dose-response patterns [2,4]. The University of Chicago group analyzed nearly 400,000 matched case-control pairs, demonstrating that protective effects were most pronounced at low-to-moderate cumulative doses (270-600 grams over 2 years), with odds ratios of 0.73-0.79 [2]. Interestingly, higher cumulative doses (>1080 grams

over 2 years) showed attenuated protective effects (OR = 0.84), suggesting a potential plateau or even diminishing returns at higher exposures [2].

These dose-response findings raise important questions about optimal dosing strategies and potential mechanisms. The non-linear relationship may reflect competing biological processes—protective metabolic effects versus potential mitochondrial toxicity at higher doses—or may simply represent residual confounding with higher doses correlating with more severe diabetes and associated comorbidities.

3. TARGET TRIAL EMULATION: A FRAMEWORK FOR STRENGTHENING CAUSAL INFERENCE

3.1. Conceptual Foundation

Target trial emulation (TTE), pioneered by Hernán and Robins, provides a rigorous framework for designing observational studies that mimic randomized controlled trials [5,6]. The approach requires explicit specification of seven core protocol components:

- Eligibility criteria
- Treatment strategies
- Treatment assignment
- Start of follow-up (time zero)
- End of follow-up
- Outcomes
- Statistical analysis plan

Critically, TTE addresses multiple sources of bias that plague conventional observational analyses [5]. By explicitly aligning eligibility, treatment assignment, and follow-up initiation at a defined 'time zero,' TTE eliminates immortal time bias. By specifying treatment strategies rather than treatment received, it avoids post-treatment selection bias. By requiring clear outcome definitions and analysis plans, it reduces outcome misclassification and selective reporting.

3.2. Application to Metformin-AMD Research

Despite the theoretical appeal of TTE, its application to metformin-AMD research remains limited [5,6]. A properly specified target trial for metformin in AMD prevention would include:

Eligibility: Individuals aged ≥60 years without AMD at baseline (confirmed by multimodal imaging), free from contraindications to metformin, with normal renal function (eGFR ≥45 mL/min/1.73m²).

Treatment strategies: (1) Initiate metformin 500mg twice daily with dose escalation to 1000mg twice daily as tolerated; (2) No metformin (placebo or standard care).

Treatment assignment: Random assignment at time zero.

Time zero: Date of first ophthalmologic examination after age 60 confirming no AMD.

Follow-up: Annual comprehensive eye examinations with optical coherence tomography (OCT) and fundus photography for 5-10 years.

Outcome: Incident AMD (early, intermediate, or advanced) based on standardized classification systems.

Analysis: Intention-to-treat with per-protocol sensitivity analyses using inverse probability weighting to account for non-adherence and informative censoring.

Emulating this target trial using real-world data requires addressing several challenges: identifying appropriate time zero in electronic health records, measuring and adjusting for time-varying confounders (such as diabetes development, renal function changes, or other medication initiations), handling non-adherence and treatment switching, and addressing missing data in ophthalmologic examinations [6].

4. ADVANCED CAUSAL INFERENCE METHODOLOGIES

4.1. Propensity Score Methods

Propensity score (PS) methods attempt to balance measured confounders between treatment groups by condensing multidimensional baseline characteristics into a single score representing the probability of treatment. Recent metformin-AMD studies have employed PS matching, inverse probability weighting (IPW), and stratification [2,3,4,8].

The largest PS-matched study analyzed approximately 200,000 matched pairs from the Merative MarketScan database [3]. After 1:1 matching on demographics, comorbidities, and healthcare utilization, metformin users demonstrated 14% lower odds of developing AMD (adjusted OR = 0.86, 95% CI: 0.82–0.90) [3]. Critically, this effect persisted in diabetic patients without diabetic retinopathy, suggesting the protective association is independent of diabetes-related ocular complications [4].

However, PS methods share a fundamental limitation: they can only balance *measured* confounders. Unmeasured factors—such as genetic predisposition, dietary patterns, smoking intensity, or subclinical systemic inflammation—remain potential sources of bias. The assumption of no unmeasured confounding (strong ignorability) is untestable and likely violated in metformin-AMD research.

4.2. Instrumental Variable Analysis

Instrumental variable (IV) analysis offers a potential solution to unmeasured confounding by identifying variables that affect treatment assignment but have no direct effect on the outcome except through the treatment. Valid instruments must satisfy three assumptions: (1) relevance—strong association with treatment; (2) independence—no association with unmeasured confounders; and (3) exclusion restriction—no direct effect on outcome.

In metformin-AMD research, potential instruments include: physician prescribing preferences (calendar time trends in metformin vs. alternative diabetes medications), geographic variation in metformin prescribing patterns, or policy changes affecting metformin accessibility. However, identifying truly valid instruments remains challenging. Physician preference instruments may correlate with unmeasured patient characteristics (selection bias), while geographic instruments may reflect regional differences in AMD risk factors (violation of exclusion restriction).

To our knowledge, no published metformin-AMD study has successfully employed IV methods, highlighting the difficulty of identifying valid instruments in this context [1]. Future research might leverage natural experiments—such as formulary changes, guideline updates, or regional policy variations—to construct more defensible instrumental variables.

4.3. Causal Artificial Intelligence and Machine Learning Approaches

Causal AI represents an emerging frontier in causal inference, leveraging machine learning techniques while maintaining theoretical rigor regarding causal relationships [10]. Several approaches show promise for metformin-AMD research:

Doubly robust estimation: Combines outcome regression and PS models using machine learning algorithms (random forests, gradient boosting, neural networks) to estimate treatment effects. These methods remain unbiased if either the outcome model or PS model is correctly specified—a weaker assumption than requiring both models to be correct.

Targeted maximum likelihood estimation (TMLE): Provides doubly robust, efficient estimation of causal parameters while accommodating high-dimensional confounders through machine learning. TMLE has been successfully applied to pharmaceutical epidemiology but not yet to metformin-AMD research.

Causal forests and causal Bayesian networks: Enable discovery of heterogeneous treatment effects and complex causal pathways. These methods could identify subpopulations (defined by age, genetic markers, baseline drusen burden, or systemic comorbidities) most likely to benefit from metformin.

Deep learning for confounding adjustment: Neural networks can learn complex, non-linear relationships between covariates and outcomes, potentially capturing confounding patterns missed by traditional regression. However, these models require large sample sizes, careful validation, and transparent interpretation to avoid overfitting and spurious associations [10].

A recent proof-of-concept study in AMD prognosis demonstrated the feasibility of hybrid neuro-symbolic frameworks combining mechanistic disease knowledge with deep learning [10]. Such approaches achieved >85% explainability while maintaining high predictive accuracy (AUROC 0.94) [10]. Adapting these frameworks to causal estimation of metformin effects—by integrating biological pathway knowledge, multimodal imaging data, and genomic information—represents a promising direction for future research.

5. CRITICAL LIMITATIONS AND BIASES IN CURRENT EVIDENCE

5.1. Disease Latency Bias

Disease latency bias (DLB) poses a particularly insidious threat to validity in metformin-AMD research [7]. AMD exhibits a prolonged preclinical phase during which subclinical retinal changes (drusen accumulation, RPE alterations) develop years before clinical diagnosis. If these early changes cause subtle visual symptoms or lead to healthcare encounters that influence medication prescribing, DLB can create spurious protective associations [7].

For example, patients developing early AMD may experience difficulty reading or driving, prompting physician visits where metformin is *not* prescribed due to age-related changes in renal function or polypharmacy concerns. This creates a selection bias where patients destined to develop AMD are *less likely* to receive metformin, artifactually inflating protective associations [7].

Addressing DLB requires [7]: (1) Ensuring truly AMD-free status at baseline through multimodal imaging rather than relying solely on diagnostic codes; (2) Implementing lag periods between exposure assessment and outcome ascertainment; (3) Using negative control outcomes (conditions unrelated to metformin) to detect unmeasured confounding patterns; and (4) Conducting sensitivity analyses excluding patients diagnosed with AMD shortly after metformin initiation.

5.2. Confounding by Diabetes and Its Complications

The relationship between diabetes, diabetic retinopathy, and AMD introduces complex confounding [8]. Diabetes itself may affect AMD risk through multiple pathways: chronic hyperglycemia, advanced glycation end-products, microvascular dysfunction, and systemic inflammation. Diabetic retinopathy shares pathophysiologic features with AMD, potentially affecting outcome classification.

Studies addressing this challenge through stratification (diabetics without retinopathy vs. non-diabetics) or by demonstrating protective effects in non-diabetic metformin users strengthen causal arguments [3,4]. However, concerns persist about residual confounding, particularly regarding:

- Diabetes severity and duration
- Glycemic control trajectory over time
- Other diabetes medications with potential AMD effects
- Diabetes-related comorbidities (cardiovascular disease, nephropathy)

6. RECOMMENDATIONS FOR FUTURE RESEARCH

6.1. Methodological Imperatives

To advance causal understanding of metformin's effects on AMD, future observational studies must incorporate rigorous causal inference frameworks [5,6]:

Explicit target trial protocols: All observational analyses should begin by specifying the target randomized trial being emulated, including clear eligibility criteria, treatment strategies, time zero definition, and analysis plans [5,6].

Active comparator designs: Rather than comparing metformin users to non-users (which maximizes confounding), compare metformin to specific alternative diabetes medications (e.g., DPP-4 inhibitors, SGLT2 inhibitors) to reduce confounding by indication [8].

Clone-censor-weight approach: Create 'clones' of each patient at time zero, assign them to different treatment strategies, censor when their actual treatment deviates from assigned strategy, and use inverse probability weighting to adjust for informative censoring [6].

Negative control outcomes: Include outcomes known to be unrelated to metformin (e.g., acoustic neuroma, inguinal hernia) to detect unmeasured confounding patterns. Protective associations with negative controls indicate bias.

Triangulation across data sources: Replicate analyses across multiple independent databases (commercial insurance, Medicare, Veterans Affairs, international registries) to assess consistency and detect database-specific biases [1,8].

6.2. The Imperative for Randomized Trials

Ultimately, observational evidence—no matter how methodologically sophisticated—cannot definitively establish causality. Randomized controlled trials remain the gold standard for causal inference. Given metformin's favorable safety profile, widespread availability, and low cost, a properly designed RCT is both feasible and ethically justified.

The METforMIN trial, currently underway, represents a critical step forward. This phase 2 randomized trial investigates metformin for minimizing geographic atrophy progression in non-diabetic AMD patients. While small (target enrollment approximately 100 participants), it will provide crucial proof-of-concept data on safety, tolerability, and preliminary efficacy signals.

However, definitive efficacy testing requires a large-scale, multicenter RCT with:

- Enrollment of $\geq 5,000$ participants without AMD or with early AMD
- Stratification by diabetes status (diabetic and non-diabetic cohorts)
- Standardized, multimodal imaging outcomes (OCT, fundus autofluorescence)
- Follow-up duration of ≥ 5 years to capture adequate AMD events
- Pre-specified subgroup analyses by age, genetic risk (CFH, ARMS2 variants), baseline drusen burden, and smoking status
- Comprehensive safety monitoring including renal function, vitamin B12 levels, and gastrointestinal tolerability

7. CONCLUSIONS

The accumulated observational evidence suggests a potential protective association between metformin use and AMD development, with the most recent meta-analysis demonstrating a 14% reduction in odds across nearly 3 million participants [1]. Dose-response analyses indicate maximal benefit at moderate cumulative doses [2], and protective effects appear independent of diabetic retinopathy status [4]. These findings align with plausible biological mechanisms involving AMPK activation, autophagy enhancement, oxidative stress reduction, and angiogenesis inhibition.

However, critical gaps in causal evidence persist. Despite the volume of observational studies [2,3,4,8,9], none have rigorously applied target trial emulation frameworks or employed instrumental variable methods to address unmeasured confounding. Substantial biases—including disease latency bias [7], immortal time bias, confounding by indication, and healthy user effects—likely inflate protective associations. High statistical heterogeneity across studies ($I^2 = 90\%$) suggests that effect estimates vary substantially based on population characteristics, study design, and residual confounding [1].

The path forward requires a two-pronged approach. First, observational research must embrace rigorous causal inference methodologies—target trial emulation [5,6], active comparator designs,

negative control outcomes, and causal AI techniques [10]—to strengthen causal arguments from existing data. Second, and more importantly, definitive answers demand properly powered randomized controlled trials specifically designed to test metformin's efficacy in AMD prevention among non-diabetic populations at risk.

Given the enormous public health burden of AMD [1], the favorable safety profile of metformin, and its low cost and widespread availability, investment in such trials is justified and urgent. Until then, clinicians should view current observational evidence as hypothesis-generating rather than practice-changing, recognizing the fundamental distinction between association and causation.

References

1. Han MD, Lee SN, Warad D, et al. Potential Efficacy of Metformin for Age-Related Macular Degeneration: A Systematic Review and Meta-Analysis. *Ophthalmology Science*. 2025;5(2):100399.
2. Blitzer AL, Ham SA, Colby KA, Skondra D. Association of Metformin Use With Age-Related Macular Degeneration: A Case-Control Study. *JAMA Ophthalmol*. 2021;139(3):302-309.
3. Aggarwal S, Moir J, Hyman MJ, et al. Metformin Use and Age-Related Macular Degeneration in Patients Without Diabetes. *JAMA Ophthalmol*. 2024;142(1):53-57.
4. Kaufmann GT, Hyman MJ, Gonnah R, Hariprasad S, Skondra D. Association of Metformin and Other Diabetes Medication Use and the Development of New-Onset Dry Age-Related Macular Degeneration: A Case-Control Study. *Invest Ophthalmol Vis Sci*. 2023;64(11):22.
5. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA*. 2022;328(24):2446-2447.
6. Fu EL, Clase CM, Evans M, et al. Target Trial Emulation to Improve Causal Inference from Observational Data: What, Why, and How? *J Am Soc Nephrol*. 2023;34(8):1305-1318.
7. Ling JYM, Mansournia MA, Etminan M. Disease latency bias and the protective effect of metformin against age-related macular degeneration. *Eye*. 2024;38:1616-1617.
8. Gokhale KM, Adderley NJ, Subramanian A, et al. Metformin and risk of age-related macular degeneration in individuals with type 2 diabetes: a retrospective cohort study. *Br J Ophthalmol*. 2023;107(7):980-986.
9. Brown EE, Ball JD, Chen Z, Khurshid GS, Prosperi M, Ash JD. The Common Antidiabetic Drug Metformin Reduces Odds of Developing Age-Related Macular Degeneration. *Invest Ophthalmol Vis Sci*. 2019;60(5):1470-1477.
10. Chen Z, Lu Q, Wang M, et al. Explainable Artificial Intelligence Framework for Predicting Treatment Outcomes in Age-Related Macular Degeneration. *Sensors*. 2024;25(22):6879.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.