Open Access Indonesian Journal of Medical Reviews
Volume 6, Issue 4, 2026 | Article ID: 930 Pages: 362–377 | e-ISSN 2807-6257
OAIJMR

ORIGINAL ARTICLE

One-Month Anatomical and Functional Response to Intravitreal Bevacizumab in Diabetic Macular Edema: A Retrospective Single-Arm Study in Padang, Indonesia

Yolanda Wulandari Erwen[1*], Weni Helvinda[2], Khalilul Rahman[2]

[1]Ophthalmology Resident of Andalas University/M. Djamil General Hospital, Padang, Indonesia
[2]Staff of Vitreoretina Subdivision, M. Djamil General Hospital, Padang, Indonesia

ARTICLE INFO

Article history
Received: July 9, 2026
Revised: August 1, 2026
Accepted: August 29, 2026
Available Online: September 4, 2026
Published: September 6, 2026

Keywords
Bevacizumab; Diabetic retinopathy; Macular edema; Tomography, optical coherence; Visual acuity

***Corresponding author**
Yolanda Wulandari Erwen
E-mail: yolaawulandari.erwen@gmail.com

DOI
https://doi.org/10.37275/oaijmr.v6i4.930

ABSTRACT

Background: Intravitreal bevacizumab is the anti-vascular endothelial growth factor agent available for diabetic macular edema (DME) in most Indonesian public hospitals, but West Sumatran data are scarce.

Objective: We quantified the one-month response and bounded what these data support.

Methods: This was a re-analysis (STROBE) of the aggregate tabulations of a retrospective single-arm pre-post series of all eyes given bevacizumab 1.25 mg/0.05 mL for DME at Dr. M. Djamil General Hospital, Padang (January 2022 to December 2024). Central macular thickness (CMT) and Snellen acuity before and one month after injection were compared by exact McNemar test with Wilson 95% confidence intervals (CI); unidentified quantities were bounded by exhaustive enumeration rather than assumed.

Result: Fourteen eyes of 11 patients were analysed, and mean CMT fell from 527 to 318 µm (−40%; Glass's Δ 1.37). The anatomical paired table is identified: CMT reached 300 µm or less in 9 of 14 eyes (64.3%, 95% CI 38.8–83.7; exact McNemar P = 0.004), holding under complete clustering (worst case P = 0.031). The acuity transition table is not identified; across its four admissible forms the exact P spans 0.008 to 0.057, widening to 0.227 once clustering is added. Effect modification by glycaemic control was not identified (risk difference bounded only between −0.833 and +1.000 across 24 reconstructions), and no between-stratum difference is detectable at 80% power.

Conclusion: The anatomical benefit at one month is therefore large and robust, whereas the functional benefit is suggestive but not established, and the roles of diabetes duration and glycaemic control remain undetermined.

1. Introduction

Diabetic retinopathy (DR) is the commonest microvascular complication of diabetes mellitus and the leading cause of preventable vision loss in working-age adults. The most recent pooled analysis of population-based studies places the global prevalence of DR among people with diabetes at 22.3%, vision-threatening DR at 6.2% and diabetic macular edema (DME) at 6.8%[1], and regional syntheses show that these fractions vary several-fold between populations and health systems[2]. Retinopathy is frequently present before diabetes is even recognised: a substantial minority of people already have DR at the time of type 2 diabetes diagnosis[3], and community surveys in low- and middle-income settings continue to report both high prevalence and low awareness[4]. Contemporary cohorts confirm that DME prevalence rises steeply with diabetes duration[5], and access to the standards of diabetes care that would slow that trajectory is itself patterned by residence and glycaemic control[6]. In Indonesia, clinic-based series report DR in a large fraction of people attending with type 2 diabetes[7], and Indonesian-led evidence synthesis has contributed directly to the global picture[8].

DME arises when chronic hyperglycaemia destabilises the neurovascular unit of the retina. Oxidative and inflammatory injury converge on upregulation of vascular endothelial growth factor A (VEGF-A), which phosphorylates the tight-junction proteins occludin, claudin-5 and zonula occludens-1 and so breaches the inner blood-retinal barrier. Experimental work shows that microglia increase tight-junction permeability in coordination with Müller cells under diabetic conditions[9], and clinical imaging links hyperreflective foci and subretinal fluid to the same microglial activation and barrier breakdown[10]. The intraocular milieu is not VEGF alone: aqueous humour biochemistry differs measurably between eyes with and without DME[11], and cytokine profiles differ with prior vitrectomy[12], which is one reason a proportion of eyes respond incompletely to VEGF blockade. Fluid accumulates within and beneath the neurosensory retina, and the resulting increase in central macular thickness (CMT) is the imaging correlate of the disease, with 300 µm a widely used threshold for clinically significant central thickening[13].

Anti-VEGF therapy is first-line treatment for DME with central involvement. A Cochrane network meta-analysis finds that the available agents all reduce macular thickness and improve acuity, with differences between them modest relative to the difference from no treatment[14], and the Protocol T randomised comparison of aflibercept, bevacizumab and ranibizumab remains the reference for head-to-head anatomical and functional effect[15]. Newer molecules extend the treatment interval rather than the size of the effect[16,17], and combination strategies with corticosteroid or laser produce incremental rather than transformative gains[18-20]. Real-world registries consistently report smaller mean gains than pivotal trials, with treatment intensity rather than biological non-response the dominant explanation[21,22]. Bevacizumab, a recombinant humanised monoclonal antibody binding all VEGF-A isoforms, is used off-label but is by a wide margin the least costly option per loading course[23], which is why it remains the dominant first-line agent wherever health financing is constrained[24].

That last point defines the Indonesian context. Intravitreal bevacizumab is the anti-VEGF agent available at Dr. M. Djamil Central General Hospital in Padang, the tertiary referral centre for West Sumatra, where it is delivered within the Jaminan Kesehatan Nasional scheme administered by BPJS Kesehatan and under the national clinical practice guideline (PNPK) for diabetic retinopathy, with glycaemic targets set by the PERKENI consensus. Indonesian investigators have reported adjunctive strategies in this setting[25], but practice differs from trial conditions in ways that matter: injections are batched to meet procedural quotas, drug supply is intermittent, and patients travel long distances for review, so the number of eyes treated in

any period is small and follow-up is compressed. Loss to follow-up is a documented driver of poor outcomes in diabetic eye disease[26], and national-scale analyses show how heavy the cumulative treatment burden of anti-VEGF therapy becomes[27].

Published Indonesian data on the short-term response to intravitreal bevacizumab in DME remain sparse, and no contemporary series from a West Sumatran tertiary centre has reported paired OCT and acuity outcomes. A second gap is methodological, and it is the one this paper is principally about. Small single-arm series frequently attach causal language about effect modifiers, most often diabetes duration and glycaemic control, to datasets that cannot support it. The usual objection to such claims is low power. That objection is too weak. When a study reports only marginal tabulations, some quantities are not merely imprecisely estimated but are not estimated at all: no sample size would sharpen them, because the data contain no information about them whatever[28]. Distinguishing non-identification from imprecision is therefore not a pedantic refinement — it determines whether the remedy is a larger study or a different data-collection form. Our aims were accordingly (i) to quantify the change in CMT and best-corrected visual acuity (BCVA) one month after treatment with exact small-sample inference, (ii) to establish which of the paired tables underlying that inference are identified by the available tabulations and to bound those that are not, and (iii) to derive sharp bounds for the association between glycaemic control and anatomical response, together with the sample size a future study would need.

2. Methods

Study design, data provenance and reporting

This is a re-analysis of the aggregate tabulations of a retrospective, single-arm, uncontrolled pre-post (before-after) clinical series. We describe the design in these terms deliberately. Because every eye received the same intervention and each eye serves as its own control, the study has no concurrent comparison group and cannot support between-group causal inference; the label 'cohort' is not used.

The provenance of the data requires an explicit statement, because it determines what can be estimated. The analysis presented here was performed on the summary tabulations held in the departmental record of this series — counts, percentages, means and standard deviations at the level of the group — and not on a re-abstracted eye-by-eye dataset. No eye-level file was available to the analysis. That is the direct cause of the non-identification documented in sections 3.3, 3.5 and 3.6: the linkage of each eye to its patient, of HbA1c to eye, and of the anatomical to the functional outcome is present in the clinical notes but absent from the tabulations. Where such linkage could be restored from the primary record, it should be, and the corresponding bounds would collapse to point estimates. We report the bounds because they are what the available data support, and we state plainly that they are a property of the data extraction rather than of the clinical material.

Reporting follows the STROBE statement for observational studies[29], and the unit-of-analysis conventions recommended for ophthalmic research in which both eyes of some patients contribute data[30]. All analyses were performed in Python 3.11 with NumPy and SciPy; the original tabulations were produced in SPSS, and the two are reconciled in section 3.8.

Setting and participants

The study was conducted at the Vitreoretina Subdivision, Department of Ophthalmology, Dr. M. Djamil Central General Hospital, Padang, West Sumatra, a tertiary referral hospital and the teaching hospital of the Faculty of Medicine, Universitas Andalas. All patients with diabetic retinopathy and DME who attended between January 2022 and December 2024 were identified from the departmental record. Total (consecutive) sampling was used; no sampling fraction was applied.

Eyes were eligible if DME had been diagnosed by a vitreoretinal consultant ophthalmologist, the eye had received intravitreal bevacizumab, and complete paired OCT and BCVA data were available before and one month after injection. Eyes were excluded for co-existing macular disease (retinal vein occlusion, uveitis, retinitis pigmentosa or pseudophakic cystoid macular edema), vitreomacular traction, or an OCT scan of signal strength below 6. The resulting flow of patients and eyes is set out in Figure 1: of 27 patients who attended with DR and DME, 11 patients (14 eyes) received bevacizumab and had complete paired data. Three of the 11 patients contributed both eyes, giving a mean cluster size of 1.27 eyes per patient. As Figure 1 also records, the exclusion criteria were applied to the treated eyes only; the 16 untreated patients were not assessed against them, so the 40.7% figure is a treatment fraction among attenders, not a treatment rate among confirmed-eligible patients.

Intervention and what is not recorded about it

Each treated eye received intravitreal bevacizumab 1.25 mg in 0.05 mL, administered off-label in the operating theatre under topical anaesthesia and aseptic technique through the pars plana 3.5-4.0 mm posterior to the limbus. Bevacizumab binds all isoforms of VEGF-A and its vitreous concentration remains pharmacologically active for approximately one month, which is the reason the response window was fixed at that interval. Reported rates of acute endophthalmitis after intravitreal bevacizumab are of the order of one per several thousand injections[31,32].

The following are not recorded in the tabulations available to us and are therefore absent from every analysis below, and are listed for the reader in the final row of Table 1: the number of injections each eye received and whether the index injection was the first; diabetic retinopathy severity grade and panretinal photocoagulation status; DME morphological subtype, disorganisation of the retinal inner layers, ellipsoid-zone integrity, hyperreflective foci and subretinal fluid; lens status; prior focal or grid laser; blood pressure, lipid profile and renal function; the OCT device model and segmentation software; and the interval between the baseline scan and the injection. Each of these influences the response to anti-VEGF therapy[33,34,35]. Consequently we do not describe the exposure as 'a single injection', and we do not use the term 'centre-involved', since no operational criterion for central involvement beyond CMT above 300 µm is documented.

Outcomes, definitions and the endpoint hierarchy

The anatomical outcome was CMT measured on spectral-domain OCT. Three pre-specified definitions were applied and are reported separately in Table 2, because conflating them is a common source of inconsistency: (A) crossing the clinical threshold, that is CMT ≤ 300 µm at one month; (B) any absolute reduction from baseline; and (C) a relative reduction exceeding 20% of baseline. Definition A is the primary anatomical endpoint. Definitions B and C are supersets of A by construction, and we report the nesting explicitly.

The functional outcome was BCVA measured on a Snellen chart and categorised as good (better than 6/18), moderate (6/18 to 6/60) or poor (worse than 6/60). Two functional quantities are reported and they are different in kind. A shift to a better acuity category is a paired state comparison and is testable by McNemar; it is designated the primary functional endpoint. A gain of at least two Snellen lines is a change score, not a binary state measured at two times, so McNemar is not applicable to it; moreover the available dichotomy of 'improved' against 'fixed or decreased' merges stability with deterioration, which would make an apparent absence of worsening a definitional artefact. It is therefore reported as a single proportion with a Wilson interval and an exact one-sample binomial test against pre-specified null response rates of 20%, 33% and 50%. Snellen notation compresses change relative to ETDRS letter scores and is a conservative but coarse instrument[36,37,38].

Covariates available in the tabulations were age, sex, duration of diabetes mellitus and the most recent HbA1c before injection, dichotomised at 6.5% in line with the PERKENI consensus; long-term glycaemic exposure is an established determinant of

retinopathy progression[39]. Age, sex, diabetes duration and HbA1c are recorded once per patient; CMT and BCVA once per eye. The two denominators are therefore different (11 patients against 14 eyes) and every proportion in this report states which one it uses, as shown throughout Table 1. Because OCT central subfield thickness has an axial resolution of roughly 5 μm, CMT values are quoted to the nearest micrometre in the text even where the source tabulation carries two decimal places.

Statistical analysis

With 14 eyes, asymptotic procedures are not appropriate. Every test, interval, power calculation and bound reported below is exact: no normal approximation and no simulation is used anywhere in the analysis, and all enumerations are exhaustive rather than sampled.

Paired dichotomous outcomes were tested with the exact (binomial) McNemar test[40]. For discordant counts b and c the two-sided P value is $2 \times P(X \le \min(b, c))$ with X distributed as Binomial(b + c, 0.5). Because no eye deteriorated on the anatomical outcome, the McNemar odds ratio there is infinite; we therefore report its exact conditional lower 95% bound, obtained by solving $P(X \ge b \mid b + c, \pi) = 0.025$ for $\pi = \psi/(1 + \psi)$, and verified against the closed form $\psi_L = \pi_L/(1 - \pi_L)$ with $\pi_L = 0.025^{1/(b+c)}$. This conditions on the upper tail at the observed b; conditioning on the lower tail returns the bound for the reciprocal parameter. Where both discordant counts are non-zero the same conditional construction is solved numerically.

Confidence intervals for single proportions are Wilson score intervals, which have markedly better coverage than the Wald interval at these sample sizes; Clopper-Pearson intervals are given alongside them in Table 2 as a conservative check. Where no event was observed, the upper bound follows the exact rule of three. Intervals for means are Student t intervals with n − 1 degrees of freedom throughout.

Two co-primary endpoints are tested, one anatomical and one functional. No multiplicity adjustment is applied to them and none is needed for the anatomical endpoint, which remains significant at a Bonferroni-corrected α of 0.025 under every robustness scenario examined; for the functional endpoint we report the bound on its exact P and note where it crosses either threshold. All other P values in this paper are non-inferential: they index members of an enumerated admissible set rather than test hypotheses, and they must not be counted, corrected or interpreted as evidence.

Because three patients contributed both eyes, the independence assumption underlying the McNemar test is violated to an unknown degree. Conditional logistic regression stratified on patient and generalised estimating equations are the textbook remedies, but neither is estimable here: with no discordant pair in the reverse direction the conditional likelihood is monotone and the maximum likelihood estimate diverges, and a generalised estimating equation over 11 clusters with a degenerate outcome does not converge[30]. We therefore bound the effect of clustering exactly, as reported in part C of Table 3. Under complete intra-patient correlation the second eye of a bilateral patient carries no independent information, so the number of independent discordant units is bounded between b − 3 and b, and the exact McNemar test is recomputed at every point of that range. We note that deflating the discordant count by a design effect of the form 1 + (m − 1) × ICC and flooring to an integer, a procedure sometimes seen in this situation, has no coverage interpretation and is anti-conservative; it is not used here.

Robustness to individual observations was quantified by the fragility index, reported in part D of Table 3: the smallest number of discordant pairs that would have to reverse direction, holding the total number of discordant pairs fixed, for the exact P value to reach 0.05[41].

For the continuous CMT outcome only the marginal means and standard deviations were recoverable, so the within-eye correlation ρ between baseline and one-month measurements is not identified. We report the range of the paired t statistic, the two-sided P value and Cohen's d_z across ρ from −1 to +1 using $SD_d^2 = s_1^2 + s_2^2 - 2 \times \rho \times s_1 \times s_2$, and display that range in Figure 2b. This is an outer sensitivity range over a one-parameter family, not a sharp identified set: the lower corner at ρ = −1 requires a one-month measurement near 112 μm, which is not a physiologically possible macula, so we also report the range over the empirically plausible interval ρ = 0.3 to 0.7. Because the two standard deviations differ by more than a factor of two, Glass's Δ, computed on the baseline standard deviation alone, is the preferred correlation-free effect size; d_{av} is given for comparison.

Identification and bounding of unestimable quantities

Four quantities in this study are not identified by the available tabulations. Each was bounded by exhaustive enumeration of its admissible set rather than resolved by assumption, using the Fréchet-Hoeffding inequalities that constrain a joint distribution when only its margins are known[28], and each bound is reported in place of a point estimate.

First, the paired BCVA transition table. Only the marginal category distributions are recorded (3 good and 11 moderate at baseline; 11 good and 3 moderate at one month), so the number of eyes good at both times is bounded between max(0, 3 + 11 − 14) = 0 and min(3, 11) = 3, giving four admissible transition tables. Every quantity derived from that table — the exact McNemar P value, the odds-ratio bound and the fragility index — is reported over all four in part B of Table 3. By contrast the anatomical paired table IS identified, because all 14 eyes exceeded 300 μm at baseline, which forces the cell for eyes at or below 300 μm at both times to zero and leaves a unique table; part A of Table 3 sets the two side by side. This asymmetry between the two endpoints is the reason they are reported differently.

Second, the concordance between anatomical and functional response. The marginal counts are known (9 of 14 eyes reached CMT ≤ 300 μm; 9 of 14 gained at least two Snellen lines) but the joint distribution is not. We enumerated the complete admissible set for the number of eyes achieving both outcomes and computed Cohen's κ, the φ coefficient and Fisher's exact P for every admissible table; the full enumeration is given in Table 4.

Third, effect modification by glycaemic control. HbA1c is recorded per patient (5 patients at or below 6.5%, 6 above) whereas response is recorded per eye, and which patients contributed two eyes is not recorded in the tabulations. We enumerated every admissible eye-to-patient allocation together with every admissible joint table, and report in Table 5 the sharp bounds on the risk difference, the odds ratio and Fisher's exact P over that set. Because the identified set is a finite collection of points rather than an interval, we state explicitly whether the null value is itself an admissible reconstruction or merely lies in the convex hull of the set.

Fourth, the number of independent discordant units under clustering, treated as described in section 2.5. Diabetes duration was tabulated only as a single mean with no stratification, so no bound tighter than the logically vacuous interval is obtainable for it, and we say so rather than estimating it.

Precision, power and minimum detectable effect

No a priori sample-size calculation was performed, because the series comprises every eligible eye treated in the study period. We therefore report the achieved precision instead: the half-width of the Wilson interval for the primary anatomical endpoint, and the same quantity for hypothetical sample sizes up to 200 eyes.

The power of Fisher's exact test to detect effect modification by glycaemic control was computed by full enumeration of the joint binomial support, not by simulation, so the values carry no Monte Carlo error. Because the eye-to-patient allocation is unknown, power and the minimum detectable effect were computed at every one of the four admissible allocations (6 against 8, 7 against 7, 8 against 6, and 9 against 5 eyes) and are reported as a range in Table 5. The minimum detectable risk difference at 80% power was obtained by exact bisection on the same enumeration; a normal-approximation formula understates

it by about seven percentage points at these sample sizes and is not used. We report the calculation anchored both at an assumed reference response rate of 90% in the better-controlled stratum and at the observed overall response rate of 64.3%.

Sample sizes required for a future adequately powered study are reported at the foot of Table 5 for a range of plausible between-stratum differences, inflated by the design effect for two-eye correlation.

Arithmetic consistency audit of the source tabulations

Because this report re-derives every statistic from tabulations rather than from an eye-level file, an arithmetic consistency audit of those tabulations was run first, and its findings are collected in Table 6. Each printed mean was tested for granularity attainability by the GRIM criterion, that is whether any total on the instrument's plausible recording grid divides at the stated denominator to a value that rounds to the printed figure[42]. The test uses exact rational arithmetic with a half-open rounding interval throughout — a floating-point implementation decides boundary cases arbitrarily — and the search range over denominators is stated wherever a generalisation is made. Each mean-and-standard-deviation pair was then tested jointly by the GRIMMER extension, which asks whether any integer sum of squares is compatible with both the reported mean and the reported standard deviation, subject to the parity constraint that the sum of squares and the sum of a set of integers share the same parity. Each printed percentage was checked against its stated denominator, and each printed P value against every contrast it could plausibly belong to.

Two assumptions underlie the granularity tests and are stated as assumptions rather than facts. The first is that CMT is recorded in whole micrometres, which holds for the central subfield readout of the common spectral-domain platforms but would fail if a value were averaged across repeat scans or across subfields, or if two devices with different segmentation software were used across the three-year window. The second is that age is recorded in whole years, which fails if age was computed from date of birth. Where an assumption fails the corresponding finding lapses, and we say so alongside it.

Ethics

This study received ethical approval from the Ethics Committee of CMHC Research Center, Indonesia (Ref no. CMHC/VII/2022/108). The study was conducted in accordance with the Declaration of Helsinki. Because the present analysis used aggregate tabulations derived from data already recorded during routine clinical care, no additional procedure was performed for research purposes and no individually identifiable record was accessed by the analysis.

3. Results

Cohort assembly and baseline characteristics

Twenty-seven patients with diabetic retinopathy and DME attended the Vitreoretina Subdivision during the three-year study window. As detailed in Figure 1, eleven patients (40.7%) received intravitreal bevacizumab and had complete paired imaging and acuity data; the remaining 16 did not receive an injection, principally because of loss to ophthalmic follow-up, intermittent drug availability and the requirement to assemble a procedural quota before an injection list could run. Whether all 16 would have satisfied the exclusion criteria is not recorded, so 40.7% is a treatment fraction among attenders rather than a treatment rate among confirmed-eligible patients. The analysed sample was 14 eyes of 11 patients, of whom three contributed both eyes; Figure 1 also records the two endpoints on which those eyes were assessed.

Baseline characteristics are shown in Table 1. Mean age was 53.75 ± 5.3 years (95% CI 50.19-57.31) and mean duration of diabetes 6.9 ± 2.23 years (95% CI 5.40-8.40). Seven patients (63.6%, 95% CI 35.4-84.8) were male and four (36.4%, 95% CI 15.2-64.6) female. Six patients (54.5%, 95% CI 28.0-78.7) had a most recent HbA1c above 6.5%, that is above the PERKENI target, and five (45.5%, 95% CI 21.3-72.0) were at or below it. All 14 eyes had a baseline CMT above 300 µm, with a mean of 527 ± 153 µm (95% CI 438.4-614.9), and 11 of 14 eyes (78.6%) presented with moderate visual impairment. The closing row of Table 1 lists the clinical variables that the tabulations do not contain, and that omission constrains every analysis that follows.

Table 1. Baseline characteristics of the study population

Characteristic	Unit of analysis	Value	95% CI
Patient-level characteristics (n = 11 patients)			
Age, years, mean ± SD	Patient	53.75 ± 5.3	50.19–57.31
Male sex, n (%)	Patient	7 (63.6)	35.4–84.8
Female sex, n (%)	Patient	4 (36.4)	15.2–64.6
Duration of diabetes, years, mean ± SD	Patient	6.9 ± 2.23	5.40–8.40
HbA1c ≤ 6.5%, n (%)	Patient	5 (45.5)	21.3–72.0
HbA1c > 6.5%, n (%)	Patient	6 (54.5)	28.0–78.7
Patients contributing both eyes, n (%)	Patient	3 (27.3)	9.7–56.6
Eye-level characteristics (n = 14 eyes)			
Central macular thickness, µm, mean ± SD	Eye	526.67 ± 152.87	438.4–614.9
CMT > 300 µm, n (%)	Eye	14 (100.0)	78.5–100.0
BCVA good (better than 6/18), n (%)	Eye	3 (21.4)	7.6–47.6
BCVA moderate (6/18 to 6/60), n (%)	Eye	11 (78.6)	52.4–92.4
BCVA poor (worse than 6/60), n (%)	Eye	0 (0.0)	0.0–21.5
Mean eyes per patient (cluster size)	Eye	1.27	—
Variables not recorded in the available tabulations			
Diabetic retinopathy severity grade; panretinal photocoagulation status; DME morphological subtype; disorganisation of the retinal inner layers; ellipsoid-zone integrity; lens status; prior focal or grid laser; number of prior intravitreal injections; blood pressure; lipid profile; renal function; OCT device model and segmentation software; baseline-scan-to-injection interval	—	Not recorded	—

SD = standard deviation; CI = confidence interval (Wilson score for proportions, Student t with n − 1 degrees of freedom for means); CMT = central macular thickness; BCVA = best-corrected visual acuity. Patient-level and eye-level denominators differ and are stated for every row. The printed means for age and baseline CMT are the values held in the source tabulation; their arithmetic attainability is examined in Table 6. The final row lists variables whose absence constrains the interpretation of every analysis below.

Anatomical outcome

Mean CMT fell from 527 ± 153 µm at baseline to 318 ± 73 µm (95% CI 275.6-359.4) at one month, an absolute reduction of 209 µm and a relative reduction of 40%, as shown in Figure 2a and summarised in the first block of Table 2. The within-eye correlation between the two measurements is not recoverable from summary data. Across the full mathematical range of that correlation the paired t statistic lies between 3.47 and 9.74 and the two-sided P value between 2.4×10^{-7} and 0.004, so the reduction is statistically significant whatever the correlation; Figure 2b plots that range and shades the empirically plausible interval within it. We present this as an outer sensitivity range rather than a sharp identified set, because its lower corner requires a one-month measurement near 112 µm, which is not an anatomically possible macula. Over the plausible correlation range of 0.3 to 0.7, t lies between 5.28 and 6.84 and Cohen's d_z between 1.41 and 1.83. The correlation-free effect size Glass's Δ, computed on the baseline standard deviation alone, is 1.37; d_{av} is 1.86 but is not the preferred anchor here because the two standard deviations differ by more than a factor of two. By any of these measures the anatomical effect is large.

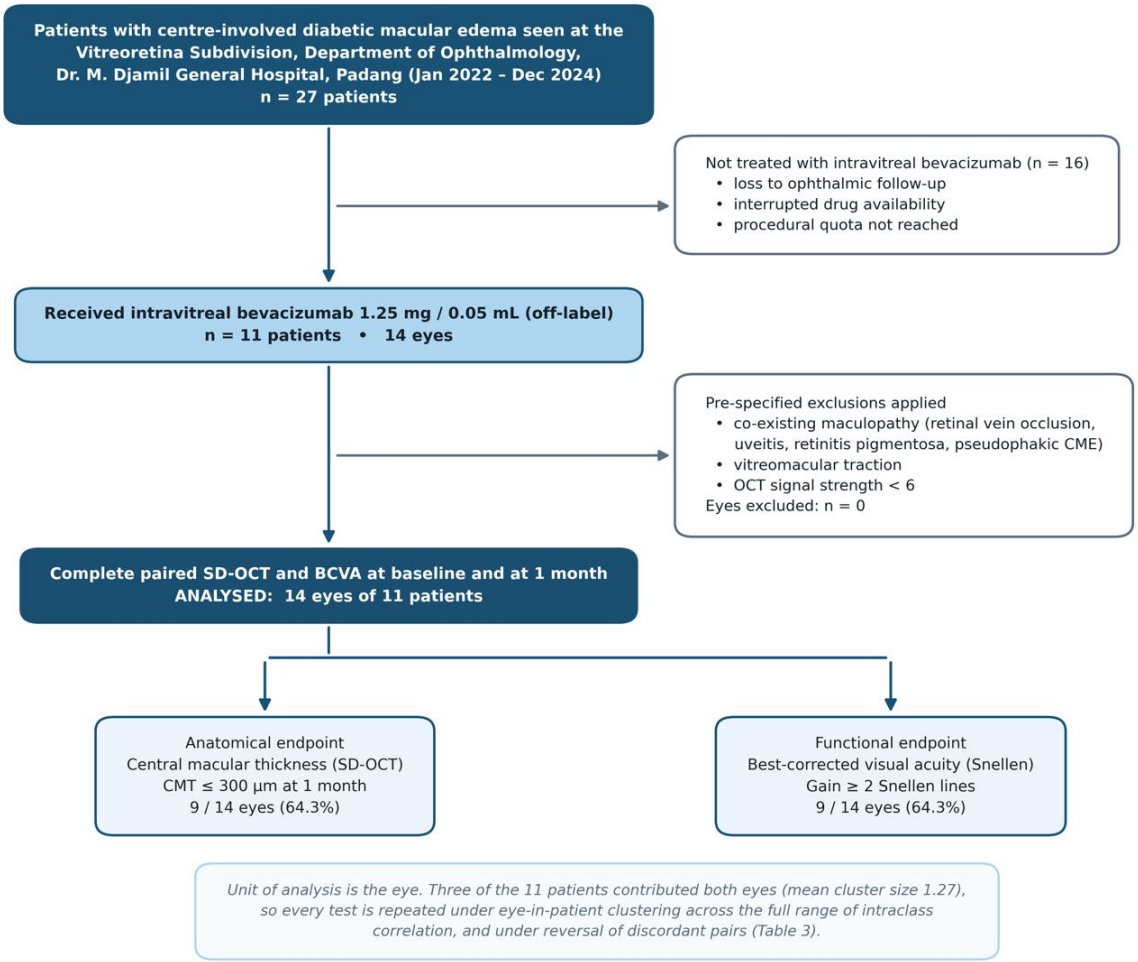

Figure 1. Assembly of the study population. Of 27 patients with diabetic macular edema attending the Vitreoretina Subdivision of Dr. M. Djamil General Hospital between January 2022 and December 2024, 11 patients (14 eyes) received intravitreal bevacizumab and had complete paired imaging and acuity data. Reasons for non-treatment were organisational rather than clinical, and the exclusion criteria were applied to the treated eyes only, so 40.7% is a treatment fraction among attenders rather than a treatment rate among confirmed-eligible patients. Three patients contributed both eyes.

The anatomical paired table is identified. Because every eye exceeded 300 µm at baseline, the cell for eyes at or below 300 µm at both times is structurally zero, and the observed margins admit exactly one transition table: 9 eyes crossed below the threshold and none crossed back, as illustrated in Figure 2c. CMT reached 300 µm or less in 9 of 14 eyes (64.3%, Wilson 95% CI 38.8-83.7; Clopper-Pearson 35.1-87.2), the exact McNemar P value is 0.004, and the exact conditional lower 95% bound for the McNemar odds ratio is 1.97 against an infinite point estimate; these are the values given in Table 2. Five eyes (35.7%, 95% CI 16.3-61.2) thinned but remained above 300 µm at one month.

The three pre-specified anatomical definitions behaved as expected and are reported separately in Table 2. Every eye showed some absolute reduction (definition B: 14/14, 100%, 95% CI 78.5-100), and the mean relative reduction of 40% is consistent with, but does not by itself establish, the claim that all 14 eyes exceeded a 20% relative reduction (definition C). Because eye-level percentage changes are not tabulated, definition C is reported as the unverified assertion it is rather than as a result.

Table 2. Primary anatomical and functional outcomes one month after intravitreal bevacizumab (14 eyes of 11 patients)

Outcome	Baseline	1 month	Effect measure (95% CI)	Exact P
Anatomical outcome — paired table IDENTIFIED				
Mean CMT, µm	527 ± 153	318 ± 73	Δ −209 (−40%); Glass's Δ 1.37; d_{av} 1.86	See note
Paired effect size d_z, plausible ρ = 0.3–0.7	—	—	1.41 to 1.83	—
Paired effect size d_z, outer range over ρ ∈ [−1, 1]	—	—	0.93 to 2.60	—
CMT ≤ 300 µm, n (%)	0 (0.0)	9 (64.3)	Wilson 38.8–83.7; Clopper-Pearson 35.1–87.2	0.004
CMT > 300 µm, n (%)	14 (100.0)	5 (35.7)	Wilson 16.3–61.2	—
Any absolute reduction in CMT, n (%)	—	14 (100.0)	Wilson 78.5–100.0	—
McNemar odds ratio, anatomical threshold	—	—	infinite (exact conditional lower bound 1.97)	0.004
Functional outcome — paired table NOT identified (see Table 3)				
BCVA good (better than 6/18), n (%)	3 (21.4)	11 (78.6)	Wilson 52.4–92.4	—
BCVA moderate (6/18 to 6/60), n (%)	11 (78.6)	3 (21.4)	Wilson 7.6–47.6	—
BCVA poor (worse than 6/60), n (%)	0 (0.0)	0 (0.0)	Wilson 0.0–21.5	—
Shift to a better acuity category (primary functional endpoint)	—	—	Exact OR lower bound 0.97 to 1.71 over 4 admissible tables	0.008 to 0.057
Gain of ≥ 2 Snellen lines, n (%)	—	9 (64.3)	Wilson 38.8–83.7	<0.001 vs 20%; 0.017 vs 33%; 0.212 vs 50%
Any recorded loss of acuity, n (%)	—	0 (0.0)	Wilson 0.0–21.5 (rule of three 23.2)	—

All P values are two-sided exact tests unless stated. The anatomical McNemar test is a single value because every eye exceeded 300 µm at baseline, which forces one cell of the transition table to zero and leaves a unique table; the corresponding acuity table is not identified, so its exact P is reported as a bound over the four admissible forms (Table 3). The mean CMT row carries no single P value because the within-eye correlation is unidentified: the two-sided paired P lies between 2.4×10^{-7} and 0.004 over ρ ∈ [−1, 1] and between 1.2×10^{-5} and 1.5×10^{-4} over the plausible range ρ = 0.3–0.7. A gain of two Snellen lines is a change score, not a state measured twice, so it is tested as a one-sample proportion against pre-specified null response rates (one-sided exact binomial). Glass's Δ is preferred to d_{av} because the two standard deviations differ by more than a factor of two.

Functional outcome, and why its paired table is not identified

At baseline 3 eyes (21.4%, 95% CI 7.6-47.6) had good acuity and 11 (78.6%) moderate acuity; no eye was in the poor category. At one month those figures were reversed, with 11 eyes (78.6%, 95% CI 52.4-92.4) good and 3 (21.4%) moderate. Figure 3a shows the two distributions side by side, and the corresponding counts appear in the second block of Table 2.

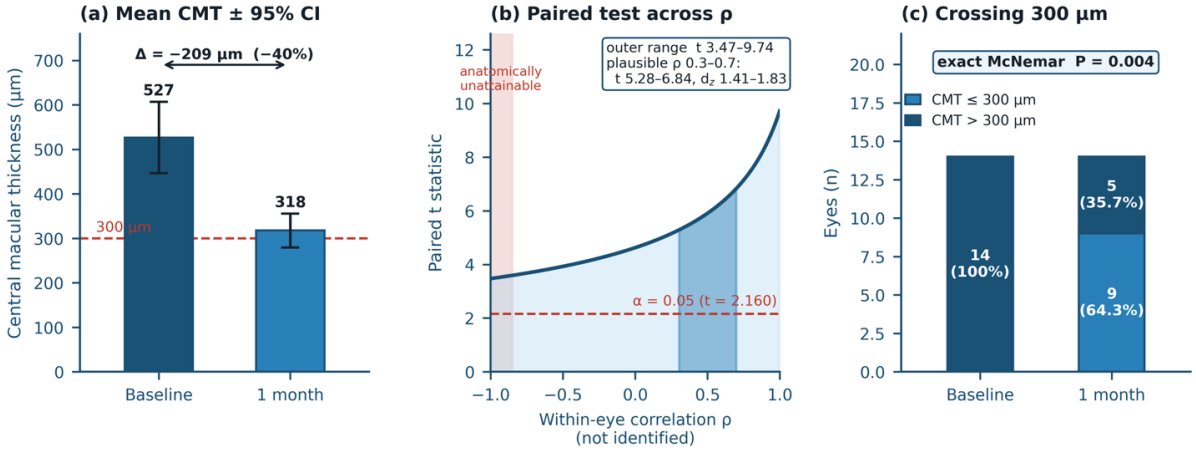

Figure 2. Anatomical outcome. (a) Mean central macular thickness fell from 527 to 318 µm, a reduction of 209 µm (40%); error bars are 95% confidence intervals and the dashed line marks the 300 µm clinical threshold. (b) The within-eye correlation between baseline and one-month measurements is not identified from summary data, so the paired t statistic is shown across its full mathematical range; it exceeds the critical value throughout, but the shaded band marks the empirically plausible correlation range of 0.3 to 0.7, and the corner at ρ = −1 is anatomically unattainable. (c) Nine of 14 eyes crossed below 300 µm and none crossed back; because every eye began above the threshold this transition table is uniquely determined, and the exact McNemar P is 0.004.

Unlike the anatomical endpoint, the functional paired table is NOT determined by these margins. The number of eyes good at both times is bounded between 0 and 3, which admits four transition tables; these are enumerated in part B of Table 3 and displayed in Figure 3b. Their exact McNemar P values are 0.057, 0.039, 0.021 and 0.008 respectively, so the exact P for the category shift is bounded only within 0.008 to 0.057 — an interval that straddles the conventional threshold — and the exact conditional lower bound for the odds ratio within 0.97 to 1.71. As Figure 3b makes plain, three of the four admissible tables reach P < 0.05 and one does not. Reporting the most favourable corner alone, as though 8 eyes had crossed upward and none downward, would assume precisely the kind of unidentified quantity this paper is about; we therefore report the bound. The point estimate most consistent with the source summary lies at that favourable corner, but it is an inference from a P value rather than an observation, and it is not treated as established here.

For the second functional quantity, 9 of 14 eyes (64.3%, 95% CI 38.8-83.7) gained at least two Snellen lines, as shown in Figure 3c. Because a two-line gain is a change score rather than a state

measured twice, it is tested as a single proportion against pre-specified null response rates: the exact one-sided binomial P values are < 0.001 against a 20% null, 0.017 against a 33% null and 0.212 against a 50% null, and these three tests are listed in Table 2. In other words the observed response rate is convincingly greater than a fifth of eyes, marginally greater than a third, and indistinguishable from a half. We emphasise the distinction between the two functional figures, because they are easily conflated: 78.6% is the prevalence of good acuity after treatment, whereas 64.3% is the proportion of eyes that improved.

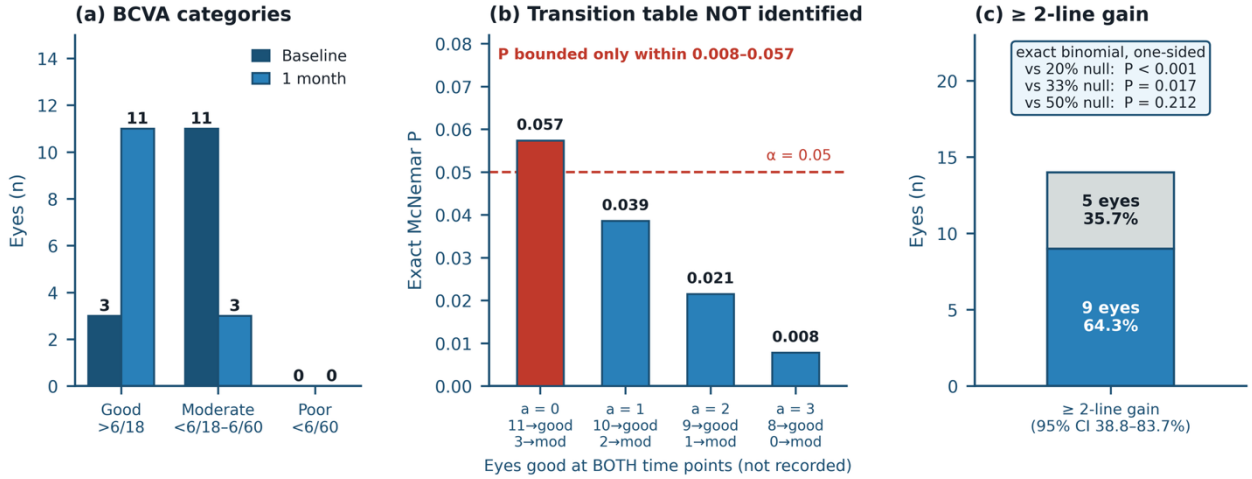

Figure 3. Functional outcome. (a) Distribution of best-corrected visual acuity categories before and one month after injection. (b) The paired transition table is NOT identified by these margins: the number of eyes good at both times is bounded between 0 and 3, giving four admissible tables whose exact McNemar P values span 0.008 to 0.057. (c) Nine of 14 eyes (64.3%) gained at least two Snellen lines, tested as a one-sample proportion against pre-specified null response rates because a two-line gain is a change score rather than a state measured twice.

No eye was recorded as losing acuity, although the available dichotomy merges stability with deterioration, so the absence of worsening is partly definitional. No adverse event was recorded. With zero events in 14 eyes, the exact upper 95% bound on the risk of any given complication, including endophthalmitis, is 23.2%, as the final row of Table 2 records; the series is uninformative about safety and must not be cited as evidence of it.

Robustness to clustering and to individual observations

Because three patients contributed both eyes, the exact McNemar test may overstate the information content of the sample. Part C of Table 3 bounds both primary tests exactly. Under complete intra-patient correlation the second eye of a bilateral patient carries no independent information, so up to three discordant units are lost. For the anatomical endpoint the number of independent discordant units is bounded between 6 and 9 and the exact P value between 0.004 and 0.031: the anatomical conclusion survives the strongest possible clustering assumption. For the functional endpoint, taking even the most favourable admissible transition table, the corresponding bound is 5 to 8 units and a P value between 0.008 and 0.063, which crosses the conventional threshold. Combining the two sources of uncertainty — the unidentified transition table and complete clustering — gives 16 scenarios of which 6 (37.5%) reach P < 0.05 and the worst attains P = 0.227, as part D of Table 3 records. The functional result is therefore suggestive but not established.

Table 3. Identification and robustness of the paired inference

Item	Structure	Discordant pairs	Exact P	Interpretation
Part A. Is the paired transition table identified?				
Anatomical endpoint (CMT ≤ 300 µm)	Yes — unique table	b = 9, c = 0	0.004	Every eye exceeded 300 µm at baseline, forcing one cell to zero
Functional endpoint (acuity category shift)	No — 4 admissible tables	see Part B	0.008 to 0.057	Both categories occupied at both time points
Part B. The four admissible acuity transition tables				
Good at both times = 0	moderate→good = 11; good→moderate = 3	b = 11, c = 3	0.057	Exact conditional OR lower bound 0.97; fragility index 0
Good at both times = 1	moderate→good = 10; good→moderate = 2	b = 10, c = 2	0.039	Exact conditional OR lower bound 1.07; fragility index 1
Good at both times = 2	moderate→good = 9; good→moderate = 1	b = 9, c = 1	0.021	Exact conditional OR lower bound 1.25; fragility index 1
Good at both times = 3	moderate→good = 8; good→moderate = 0	b = 8, c = 0	0.008	Exact conditional OR lower bound 1.71; fragility index 1
Part C. Exact patient-level clustering bound (3 patients contributed both eyes)				
Anatomical endpoint — 0 independent unit(s) lost	b = 9, c = 0	—	0.004	Significant
Anatomical endpoint — 1 independent unit(s) lost	b = 8, c = 0	—	0.008	Significant
Anatomical endpoint — 2 independent unit(s) lost	b = 7, c = 0	—	0.016	Significant
Anatomical endpoint — 3 independent unit(s) lost	b = 6, c = 0	—	0.031	Significant

Table 3. Identification and robustness of the paired inference

Item	Structure	Discordant pairs	Exact P	Interpretation
Functional endpoint, best corner — 0 independent unit(s) lost	b = 8, c = 0	—	0.008	Significant
Functional endpoint, best corner — 1 independent unit(s) lost	b = 7, c = 0	—	0.016	Significant
Functional endpoint, best corner — 2 independent unit(s) lost	b = 6, c = 0	—	0.031	Significant
Functional endpoint, best corner — 3 independent unit(s) lost	b = 5, c = 0	—	0.062	NOT significant
Part D. Combined worst case and fragility				
Functional endpoint: unidentified table combined with clustering	16 scenarios	—	0.008 to 0.227	6 of 16 (37.5%) reach P < 0.05
Anatomical endpoint fragility index	2 of 9 discordant pairs	7 vs 2	0.180	One reversal only: P = 0.039, still significant
Functional endpoint fragility index	1 under 3 of 4 admissible tables	—	0.021 to 0.070	0 under the fourth table, which is not significant to begin with

Part A distinguishes an endpoint whose paired table is determined by the observed margins from one whose table is not. Part B enumerates the four Fréchet-Hoeffding admissible acuity transition tables. Part C bounds the effect of eye-in-patient clustering exactly: under complete intra-patient correlation the second eye of a bilateral patient carries no independent information, so up to three discordant units are lost. Deflating the discordant count by a design effect and flooring to an integer, a procedure sometimes used for this purpose, has no coverage interpretation, is anti-conservative, and is not used here. Part D combines the two sources of uncertainty for the functional endpoint.

Fragility reinforces the same asymmetry, and the relevant rows are also in part D of Table 3. Reversing two of the nine anatomical discordant pairs would raise the exact P value to 0.180, so the anatomical result has a fragility index of 2; a single reversal leaves P at 0.039. For the functional endpoint the fragility index is 1 under three of the four admissible transition tables and 0 under the fourth, in which the result is not significant to begin with. This is an intrinsic property of a 14-eye series and not a defect of the analysis, but it bounds the confidence that can reasonably be placed in the functional result[41].

Concordance between anatomical and functional response is not identified

Nine eyes achieved the anatomical endpoint and nine achieved the functional endpoint, but whether these were the same nine eyes cannot be recovered from the tabulated data. Enumerating the complete admissible set, as set out in Table 4 and plotted in Figure 4a, shows that between 4 and 9 eyes could have achieved both, that is between 28.6% and 64.3% of the sample. Across that set Cohen's κ ranges from −0.556 to 1.000, the φ coefficient over the same interval, and the odds ratio from 0 to infinity. Fisher's exact P takes the values 0.086, 0.580, 1.000, 0.266, 0.023 and 0.0005 as the dual-response count rises from 4 to 9; as the final column of Table 4 shows, it is not monotone, and the two significant values sit at the two extremes of the concordance range while the middle of that range is uniformly non-significant. The published data are thus compatible with perfect agreement between anatomical and functional response and equally compatible with substantially worse-than-chance agreement, and no statement about the correspondence between the two endpoints is supportable without eye-level linkage.

Table 4. Complete Fréchet-Hoeffding admissible set for the concordance between anatomical and functional response

Both (a)	Anat. only (b)	Funct. only (c)	Neither (d)	Dual (%)	κ	Φ	OR	Fisher P
4	5	5	0	28.6	-0.556	-0.556	0.00	0.0859
5	4	4	1	35.7	-0.244	-0.244	0.31	0.5804
6	3	3	2	42.9	0.067	0.067	1.33	1.0000
7	2	2	3	50.0	0.378	0.378	5.25	0.2657
8	1	1	4	57.1	0.689	0.689	32.00	0.0230
9	0	0	5	64.3	1.000	1.000	infinite	0.0005

Nine of 14 eyes met the anatomical endpoint and nine met the functional endpoint, but the joint distribution is not recorded. Every row is a logically possible 2 × 2 table consistent with those margins; there are exactly six. Cohen's κ spans −0.556 to 1.000 and the odds ratio 0 to infinity across the set, so no statement about the correspondence between anatomical and functional response is identified by these data. Note that Fisher's exact P is NOT monotone in a: the two values below 0.05 sit at the two extremes of the concordance range while the middle of that range is uniformly non-significant, so the minimum and maximum alone would misrepresent the set. a = eyes meeting both endpoints; b = anatomical endpoint only; c = functional endpoint only; d = neither; Dual (%) = percentage of the 14 eyes meeting both; κ = Cohen's kappa; Φ = phi coefficient; OR = odds ratio.

Effect modification by glycaemic control and diabetes duration is not identified

HbA1c stratum is recorded per patient and response per eye, and the eye-to-patient allocation is not in the tabulations, so the two-by-two table relating glycaemic control to anatomical response is not determined. Enumerating every admissible allocation of the three second eyes together with every admissible joint table yields 24 admissible reconstructions, which are summarised at the head of Table 5 and plotted individually in Figure 4b. Across them the risk difference for anatomical response in the HbA1c above 6.5% stratum relative to the at-or-below stratum spans −0.833 to +1.000 and the odds ratio spans 0 to infinity. The identified set is a finite collection of 24 points, not an interval: as Figure 4b shows, no reconstruction has a risk difference of exactly zero (the closest is +0.042) and none has an odds ratio of exactly one, so the null lies within the convex hull of the identified set rather than within the set itself. The association is unidentified: the data are equally consistent with poorer glycaemic control being strongly harmful, strongly protective, or irrelevant to one-month anatomical response.

For diabetes duration the position is starker still. Duration is tabulated only as a single mean of 6.9 ± 2.23 years with no stratification of any kind, so no cross-tabulation against response exists and no bound narrower than the logically vacuous interval can be derived, as the second block of Table 5 records. We therefore make no claim about it. This matters because both variables are well-established risk factors for the development and progression of retinopathy[39,5]; that they are genuine risk factors for disease does not entitle a 14-eye series to assert that they modify the response to treatment.

Table 5. Partial identification of the effect-modification questions, what this design could detect, and what a future study would require

Quantity	Value or bound	Interpretation
Sharp bounds on effect modification by glycaemic control (24 admissible reconstructions)		
Risk difference (HbA1c > 6.5% vs ≤ 6.5%)	−0.833 to +1.000	Not identified
Odds ratio	0 to infinite	Not identified
Reconstruction with a risk difference of exactly zero	No such reconstruction (closest +0.042)	The identified set is 24 discrete points, not an interval
Reconstruction with an odds ratio of exactly one	No such reconstruction	The null lies in the convex hull of the set, not in the set
Effect modification by duration of diabetes		
Strata published in the source tabulation	0	Not identifiable at all
Derivable bound on the odds ratio	0 to infinite (logically vacuous)	—
What this design could detect (exact enumeration, α = 0.05, no simulation)		
Power at a 70-point difference (90% vs 20%)	69.6% at 8 vs 6 eyes	64.8–69.6% across the 4 admissible allocations
Power at a 60-point difference (90% vs 30%)	52.4% at 8 vs 6 eyes	44.2–52.4% across the 4 admissible allocations
Power at a 50-point difference (90% vs 40%)	35.5% at 8 vs 6 eyes	26.3–35.5% across the 4 admissible allocations
Power at a 40-point difference (90% vs 50%)	20.6% at 8 vs 6 eyes	13.3–20.6% across the 4 admissible allocations
Power at a 30-point difference (90% vs 60%)	9.5% at 8 vs 6 eyes	5.3–9.5% across the 4 admissible allocations
Minimum detectable risk difference at 80% power	76.2 percentage points (8 vs 6 eyes)	75.9–77.7 points across allocations
Exact power at the normal-approximation answer (68.9 points)	67.7%	The asymptotic formula is optimistic by about 7 points and is not used
Maximum attainable power anchored at the observed 64.3% response rate	75.4%	No between-stratum difference at all is detectable at 80% power
Sample size required for a future study (80% power, α = 0.05, design effect 1.27)		
Difference of 20 percentage points (90% vs 70%)	62 eyes per stratum	158 eyes in total
Difference of 30 percentage points (90% vs 60%)	32 eyes per stratum	82 eyes in total
Difference of 40 percentage points (90% vs 50%)	20 eyes per stratum	51 eyes in total
Difference of 20 percentage points (80% vs 60%)	82 eyes per stratum	209 eyes in total
Difference of 30 percentage points (80% vs 50%)	39 eyes per stratum	100 eyes in total
Difference of 25 percentage points (75% vs 50%)	58 eyes per stratum	148 eyes in total
Sample size on the continuous scale (difference in CMT reduction between strata)		
Difference of 50 µm (d = 0.34)	139 eyes per stratum	278 eyes in total
Difference of 75 µm (d = 0.51)	62 eyes per stratum	124 eyes in total
Difference of 100 µm (d = 0.68)	35 eyes per stratum	70 eyes in total
Difference of 125 µm (d = 0.84)	23 eyes per stratum	46 eyes in total
The remedy that requires no additional patients		
Record laterality, per-eye HbA1c and per-eye outcome status	0 additional eyes	Removes the non-identification in Tables 3, 4 and 5 entirely, at any sample size

Sharp bounds are obtained by enumerating every admissible allocation of the three second eyes to the two HbA1c strata together with every admissible joint table. Power and the minimum detectable effect were computed by full enumeration of the joint binomial support of Fisher's exact test, so they carry no Monte Carlo error; because the eye-to-patient allocation is unknown, both are reported across all four admissible allocations. Sample sizes on the dichotomous scale are inflated by the design effect for two-eye correlation; those on the continuous scale assume a paired-difference standard deviation of 148 µm, the mid-range value across the admissible within-eye correlations. The P values in this table index members of an enumerated admissible set; they are not hypothesis tests and must not be counted or corrected as such.

Precision, power and what a future study would require

The Wilson 95% interval for the primary anatomical endpoint runs from 38.8% to 83.7%, a half-width of 22.5 percentage points. Figure 4d traces how that half-width shrinks with sample size: reaching 9.3 percentage points would require approximately 100 eyes and 6.6 points approximately 200 eyes.

For the effect-modification question the design is weaker still, and weaker than a normal-approximation calculation would suggest. By exact enumeration at the mid-range allocation of 8 against 6 eyes, a 40-percentage-point difference in response rates (50% against 90%) is detected with 20.6% power at a two-sided α of 0.05, and the minimum risk difference detectable with 80% power is 76.2 percentage points; across the four admissible allocations that minimum ranges from 75.9 to 77.7 points and power at a 40-point difference ranges from 13.3% to 20.6%. Figure 4c plots the whole power curve with its allocation band, and the corresponding rows of Table 5 give the numbers. A normal-approximation formula returns 68.9 points, which is optimistic by about seven points and has only 67.7% exact power at its own claimed minimum. More striking still, if the reference stratum is anchored at the observed overall response rate of 64.3% rather than at an assumed 90%, the maximum attainable exact power across all allocations — even against a risk difference so large that no eye in one stratum responds — is 75.4%. No between-stratum difference whatever is detectable at 80% power in this design.

A future study designed to answer the same question at 80% power would need approximately 82 eyes to detect a 30-percentage-point difference and approximately 158 eyes to detect a 20-percentage-point difference once the design effect for two-eye correlation is applied. Framed on the continuous scale, detecting a 100 µm difference in CMT reduction between glycaemic strata would require approximately 70 eyes. All of these figures are tabulated in Table 5. The more immediate remedy, however, costs nothing and is recorded in the final row of that table: recording laterality, per-eye HbA1c and per-eye outcome status on the clinical form would remove the non-identification described in sections 3.3, 3.5 and 3.6 entirely, at any sample size.

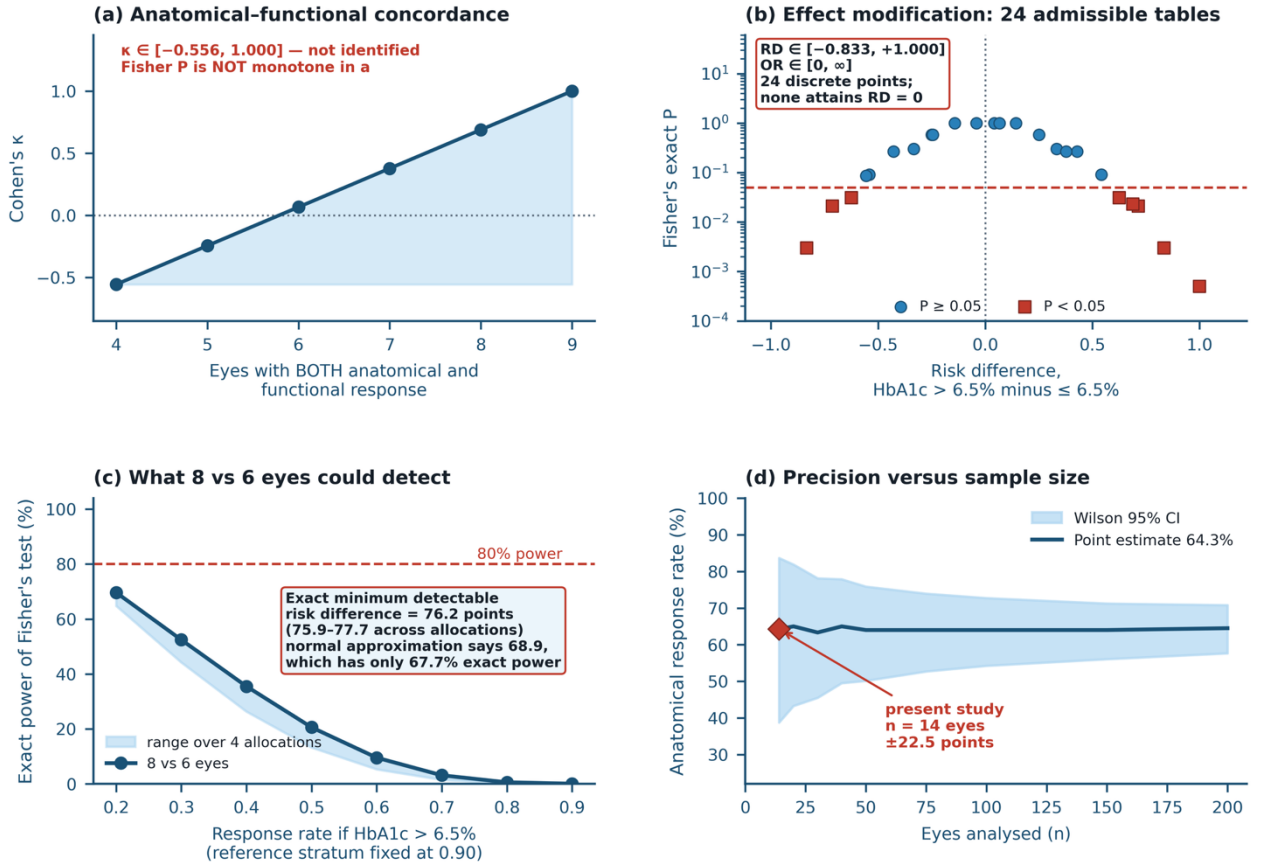

Figure 4. Partial identification and precision. (a) The complete admissible set for the concordance between anatomical and functional response: Cohen's κ is bounded only between −0.556 and 1.000. (b) All 24 admissible reconstructions of the association between glycaemic control and anatomical response; the risk difference spans −0.833 to +1.000 and no reconstruction attains the null exactly. (c) Exact power of Fisher's exact test, computed by enumeration across all four admissible eye allocations: the minimum detectable risk difference at 80% power is 76.2 percentage points at the mid-range allocation, and the normal-approximation answer of 68.9 points is optimistic. (d) Width of the Wilson confidence interval for the anatomical response rate as a function of sample size.

Arithmetic consistency of the source tabulations

Table 6 lists the discrepancies identified in the tabulations from which this analysis was built, together with the corrected values used here. Three are arithmetic. The printed baseline CMT mean of 526.67 µm cannot be produced by any set of 14 whole-micrometre measurements: the nearest attainable value at n = 14 is 526.64 µm, and searching denominators from 1 to 60 shows that 526.67 is attainable at n = 3, 6, 9, 12 and other multiples of three, and also at n = 43, 46, 49, 52, 55 and 58. The most economical explanation is not that the OCT grid is finer than 1 µm but that the mean was computed on a denominator of 3, 6, 9 or 12; twelve is a suggestive number in a series that reports 14 eyes, and the authors should confirm the denominator. The printed mean age of 53.75 years is likewise unattainable from 11 whole-year ages and from 11 half-year values, though attainable at quarter-year granularity and at denominators including 4, 8, 12, 51, 53, 55, 57 and 59; if age was computed from date of birth, as is usual, the whole-year premise fails and this finding lapses. A figure of 37.5% appears in narrative text for the 5 of 14 eyes that remained above 300 µm, where the correct value is 35.7%. The difference between the two CMT means is 209.17 µm as printed and 209.14 µm at the nearest attainable baseline, so no substantive conclusion depends on the discrepancy.

A fourth arithmetic finding, recorded in the GRIMMER block of Table 6, emerges only from the joint mean-and-standard-deviation test. The pair (6.9, 2.23) for diabetes duration passes the mean-only check — 76/11 = 6.909 rounds to 6.9 — but fails the GRIMMER criterion: with the total fixed at 76, the reported sample standard deviation of 2.23 requires an integer sum of squares of 575, whereas the parity constraint that the sum of squares of a set of integers shares the parity of their sum forces an even value. The pair is therefore jointly unattainable from 11 whole-year durations. The same test passes for the one-month CMT pair (317.50, 72.52), which admits nine compatible sums of squares, and is not applicable to the age and baseline-CMT pairs because their means already fail.

Table 6. Arithmetic consistency audit of the source tabulations, with the corrected values used in this report

Item	As held in the source tabulation	Audit finding	Value used here
Granularity attainability of the printed means (exact rational arithmetic, n searched 1–60)			
Baseline mean CMT (n = 14 eyes)	526.67 µm	Not attainable from 14 whole-µm values. Nearest attainable at n = 14 is 526.64. Attainable at n = 3, 6, 9, 12, 15 and other multiples of 3, and also at n = 43, 46, 49, 52, 55 and 58. Assumes whole-µm recording, which fails if the value is averaged across scans or subfields, or if two OCT devices were used.	526.67 retained; the authors should confirm the denominator, since 12 is a plausible candidate
Mean age (n = 11 patients)	53.75 years	Not attainable from 11 whole-year ages, nor at half-year granularity; attainable at quarter-year granularity and at n = 4, 8, 12, 51, 53, 55, 57, 59. Assumes whole-year recording; if age was computed from date of birth, as is usual, this finding lapses.	53.75 retained; recording granularity to be confirmed
Mean duration of diabetes (n = 11)	6.9 years	Mean attainable (76/11 = 6.909 rounds to 6.9). See the GRIMMER row below.	6.9

Table 6. Arithmetic consistency audit of the source tabulations, with the corrected values used in this report

Item	As held in the source tabulation	Audit finding	Value used here
Mean CMT at 1 month (n = 14 eyes)	317.50 µm	Attainable exactly (total 4,445 µm). No discrepancy.	317.50
Joint mean-and-standard-deviation consistency (GRIMMER, with the integer parity constraint)			
Duration of diabetes (6.9 ± 2.23, n = 11)	Passes the mean-only check	FAILS jointly. With the total fixed at 76, a sample SD of 2.23 requires an integer sum of squares of 575, but the sum of squares of a set of integers shares the parity of their sum, which is even. No admissible sum of squares exists.	Reported as printed with the inconsistency flagged; the authors should re-derive it
CMT at 1 month (317.50 ± 72.52, n = 14)	—	PASSES: 9 compatible integer sums of squares exist.	317.50 ± 72.52
Age and baseline CMT pairs	—	Not applicable: the means themselves already fail the granularity test.	—
Denominators and percentages			
Demographic percentages	63.6 / 36.4 / 45.5 / 54.5	All computed on 11 patients, not 14 eyes; correct, but the change of denominator is not declared.	Denominator stated in every row of Table 1
Eyes remaining above 300 µm	5 eyes (37.5%)	5/14 = 35.7%; 37.5% corresponds to 3/8.	35.7% (95% CI 16.3–61.2)
Attribution of P values			
P = 0.004	Printed beside 'all 14 eyes showed a reduction'	Reproduces exactly as the exact McNemar test on the 9 eyes crossing 300 µm ($2 \times 0.5^9 = 0.0039$). Applied to the 14/0 contrast it would be 0.0001.	Reported against the threshold-crossing contrast
P = 0.008	Printed beside 'gain of ≥ 2 Snellen lines in 9 eyes'	Reproduces exactly as the exact McNemar test on 8 upward and 0 downward category transitions ($2 \times 0.5^8 = 0.0078$). Applied to the 9/5 contrast it would be 0.004. This match cannot be used as evidence that the (8, 0) table was observed, since the same audit finds the value misattributed.	Reported as a bound over the four admissible transition tables (Table 3)
Editorial and software consistency			
Recruitment window	January 2022 to December 2024 / to November 2024	Stated inconsistently in two places.	January 2022 to December 2024
Response definition	'decrease of > 20% from baseline' / 'decreased CMT value'	Two different definitions used interchangeably; per-eye percentage changes are not tabulated.	Three definitions specified and reported separately
Analysis software	SPSS 29	The present analysis uses Python 3.11 with NumPy and SciPy.	Every value recomputed from first principles rather than carried across
Checks that passed			
Feasibility of the 1-month summary	mean 317.50, SD 72.52, 9 eyes ≤ 300	Minimum SD compatible with those constraints is 24.37; the reported SD exceeds it.	Feasible
Feasibility of the baseline summary	mean 526.67, SD 152.87, all eyes > 300	Implies a maximum single measurement of about 1,058 µm, physiologically plausible.	Feasible

Granularity attainability was tested by exact rational arithmetic over a half-open rounding interval; a floating-point implementation decides boundary cases arbitrarily and was not used. The GRIMMER test asks whether any integer sum of squares is compatible with both the printed mean and the printed standard deviation, subject to the constraint that the sum of squares of a set of integers shares the parity of their sum. Both tests rest on stated assumptions about recording granularity; where an assumption fails, the corresponding finding lapses. No finding alters the direction or the statistical significance of the anatomical outcome.

Two findings are attributional. The value P = 0.004 reproduces exactly and only as the exact McNemar test on the 9 eyes that crossed the 300 µm threshold ($2 \times 0.5^9 = 0.0039$); applied instead to the statement that all 14 eyes showed some reduction it would be 0.0001. Likewise P = 0.008 reproduces exactly and only as the exact McNemar test on 8 upward and 0 downward category transitions ($2 \times 0.5^8 = 0.0078$); applied to the 9 eyes gaining at least two Snellen lines it would be 0.004. Both were correct computations reported alongside a different contrast from the one they tested, and Table 6 restores each to its proper contrast, while noting the circularity that must be avoided: the fact that P = 0.008 matches one of the four admissible transition tables cannot be used as evidence that this table was observed, precisely because the same audit finds the value misattributed. Two further inconsistencies were editorial: the recruitment window is stated as ending in December 2024 in one place and November 2024 in another, and demographic percentages are computed on 11 patients while outcome percentages are computed on 14 eyes without the change of denominator being declared.

Two consistency checks passed, and they close Table 6. Given that 9 eyes finished at or below 300 µm and 5 above, the minimum standard deviation compatible with a one-month mean of 317.50 µm is 24.37 µm, well below the reported 72.52, so the one-month summary is internally feasible. At baseline, the reported mean and standard deviation imply a maximum single measurement of about 1,058 µm, which is physiologically plausible for severe DME, so the baseline summary is feasible as well.

4. Discussion

In 14 eyes of 11 patients treated with intravitreal bevacizumab 1.25 mg at the Vitreoretina Subdivision of Dr. M. Djamil Central General Hospital, Padang, mean central macular thickness fell by 209 µm, or 40%, within one month. Just under two-thirds of eyes crossed the 300 µm clinical threshold (64.3%, 95% CI 38.8-83.7). The anatomical finding, summarised in Table 2 and Figure 2, is identified, large by every effect-size measure, and robust: it survives complete correlation between fellow eyes (worst-case exact P = 0.031) and has a fragility index of 2. The functional finding is weaker than it first appears. Its paired table is not identified by the available tabulations, and across the four admissible forms enumerated in Table 3 the exact P spans 0.008 to 0.057; adding worst-case clustering widens the range to 0.227. We therefore describe the functional benefit as suggestive rather than established. Neither the correspondence between anatomical and functional response nor any effect modification by glycaemic control or diabetes duration is identified by these data, and we quantify rather than assume that limitation.

The demographic profile of the series, set out in Table 1, is unremarkable and consistent with the wider literature. Mean age was 53.75 years, at the younger end of the range most often

reported for DME, which is expected in an Indonesian referral population where type 2 diabetes presents earlier than in European cohorts[7,3]. Men accounted for 63.6% of patients, in line with reports of a higher incidence of retinopathy and of progression to DME in men[1,2]. Mean diabetes duration of 6.9 years places this cohort within the window in which DME incidence rises most steeply[5]. With 11 patients none of these comparisons is a test of anything; they are offered only to establish that the series is not obviously atypical.

The magnitude of anatomical response sits at the upper end of the published range, and the reasons matter more than the number. In the Protocol T randomised comparison the bevacizumab arm achieved a mean reduction in central subfield thickness of roughly 100 µm over the first year[15], and contemporary real-world series report smaller mean reductions still[21,43]. Bevacizumab-specific series document both the microvascular change that accompanies that reduction[44] and the proportion of eyes reaching near-normal acuity after it[45]. Our 209 µm reduction at one month, plotted in Figure 2a, exceeds both, but the quantities are not commensurable and four explanations are more plausible than a superior drug effect. First, baseline thickness and baseline choroidal thickness are among the most consistent predictors of the magnitude of anatomical response, and at 527 µm our eyes were thicker at entry than the trial populations[46,47,48]. Second, and we regard this as the most serious alternative explanation rather than a footnote: every eye was selected on a threshold of CMT above 300 µm on a single baseline scan, and central subfield thickness in DME fluctuates meaningfully from visit to visit on top of instrument repeatability. With selection on a threshold and no repeat pre-injection measurement, regression to the mean cannot be separated from drug effect at all in these data. Third, the anatomical floor compresses the comparison: normal central subfield thickness is roughly 250 to 290 µm, so the achievable reduction is bounded below by anatomy and the mean fall is driven disproportionately by the thickest eyes. Fourth, one month is close to the point of maximal early effect, whereas the trial figures are annual values under continued protocol-driven treatment; early anatomical response is itself only a partial predictor of eventual outcome[49]. The correct reading is that our result is consistent with, and not superior to, the established evidence base, and that this series adds no new efficacy evidence to it.

The functional result is similarly consistent with the literature once measurement scale is taken into account. Snellen categorisation is coarse: acuity measured on a Snellen chart differs systematically and variably from ETDRS letter scores, and category boundaries at 6/18 and 6/60 make a shift look dramatic when the underlying change may be modest[36-38]. Real-world registry analyses of anti-VEGF therapy in DME consistently report smaller mean acuity gains than pivotal trials, with treatment intensity rather than biological non-response the dominant explanation[21,22,50]. Our 64.3% two-line gain rate, shown in Figure 3c, should therefore be read as an early-response signal in a favourable subgroup, not as an achievable steady-state outcome, and it is entirely compatible with reports in which sustained benefit requires a full loading phase followed by structured re-treatment[51,52].

Five eyes (35.7%) thinned but remained above 300 µm at one month. Persistent central thickening after initial anti-VEGF therapy is common rather than exceptional, and refractory DME is best identified against a defined treatment course rather than after a single injection[53]. Mechanistically, incomplete response reflects the fact that VEGF-A is not the only mediator of barrier failure; inflammatory and microglial signalling persists after VEGF neutralisation[10-12]. Structural factors matter too: disorganisation of the retinal inner layers and ellipsoid-zone disruption limit both anatomical and functional recovery[33,34,55,56], and macular ischaemia sets a ceiling on acuity that no amount of fluid resolution can lift[57-59]. We stress that none of these features was recorded in the present series, so they are candidate explanations drawn from the literature, not findings of this study. Where thickening persists after an adequate loading phase, the evidence supports reassessment of systemic status and consideration of an agent switch, an intravitreal corticosteroid or adjunctive laser rather than mechanical repetition of the same injection[60-62]; renal impairment in particular is associated with a poorer treatment course[63,64]. Figure 5b sets out the corresponding loading-phase pathway. It is drawn from the cited literature and not from the data in this paper, its response assessment is deliberately placed after three injections rather than one, and its safety nodes are explicit because a corticosteroid branch in a phakic eye carries avoidable cataract and a material risk of intraocular pressure rise. The upper panel of Figure 5 places that pathway in its mechanistic context, tracing the route from chronic hyperglycaemia through VEGF-A upregulation and tight-junction failure to the central thickening that bevacizumab is given to reverse.

The central methodological contribution of this report is the demonstration that three quantities the study appears to estimate are not estimable from its data, and that this is a different problem from low power. The distinction is worth stating plainly, because the shorthand objection — 'n = 14, no subgroup analysis is interpretable' — is true but leads to the wrong remedy. A low-powered estimate becomes sharp as the sample grows. A non-identified quantity does not: enrolling 1,400 eyes and tabulating them the same way would leave the bounds in Tables 3, 4 and 5 exactly as wide[28]. The remedy for the first is a bigger study; the remedy for the second is a different data-collection form, and it costs nothing.

The clearest instance is the contrast between our two primary endpoints, which part A of Table 3 sets out directly. The anatomical table is identified — every eye began above 300 µm, which forces one cell of the transition table to zero — so a single exact P value of 0.004 is legitimate. The functional table is not, because both acuity categories were occupied at both times, and its exact P is bounded only within 0.008 to 0.057. Two endpoints in the same 14 eyes, analysed by the same test, differ not in precision but in whether they are estimable at all. The same logic applies to the anatomical-functional concordance in Table 4, where Cohen's κ is bounded only within −0.556 to 1.000 — a range that is clinically consequential, because the dissociation between anatomical and functional response is precisely what determines whether CMT is a usable surrogate in an individual eye[59].

The third instance concerns the effect modifiers named in the original framing of this work. HbA1c is recorded per patient and response per eye, and the linkage between them is absent from the tabulations, so the association is bounded only by the logically vacuous interval: risk difference from −0.833 to +1.000 across all 24 admissible reconstructions in Figure 4b. Diabetes duration is worse still, being tabulated only as an unstratified mean. Both are unquestionably risk factors for the development and progression of diabetic retinopathy, and long-term glycaemic variability in particular tracks progression closely[39] — but a risk factor for incident disease is not automatically a modifier of short-term treatment response, and this series could not have detected such modification even had it existed: as Figure 4c shows, no risk difference at all is detectable at 80% power when the reference rate is anchored at the observed response rate. Stating this explicitly, rather than reporting a plausible-sounding subgroup narrative, is the more useful contribution.

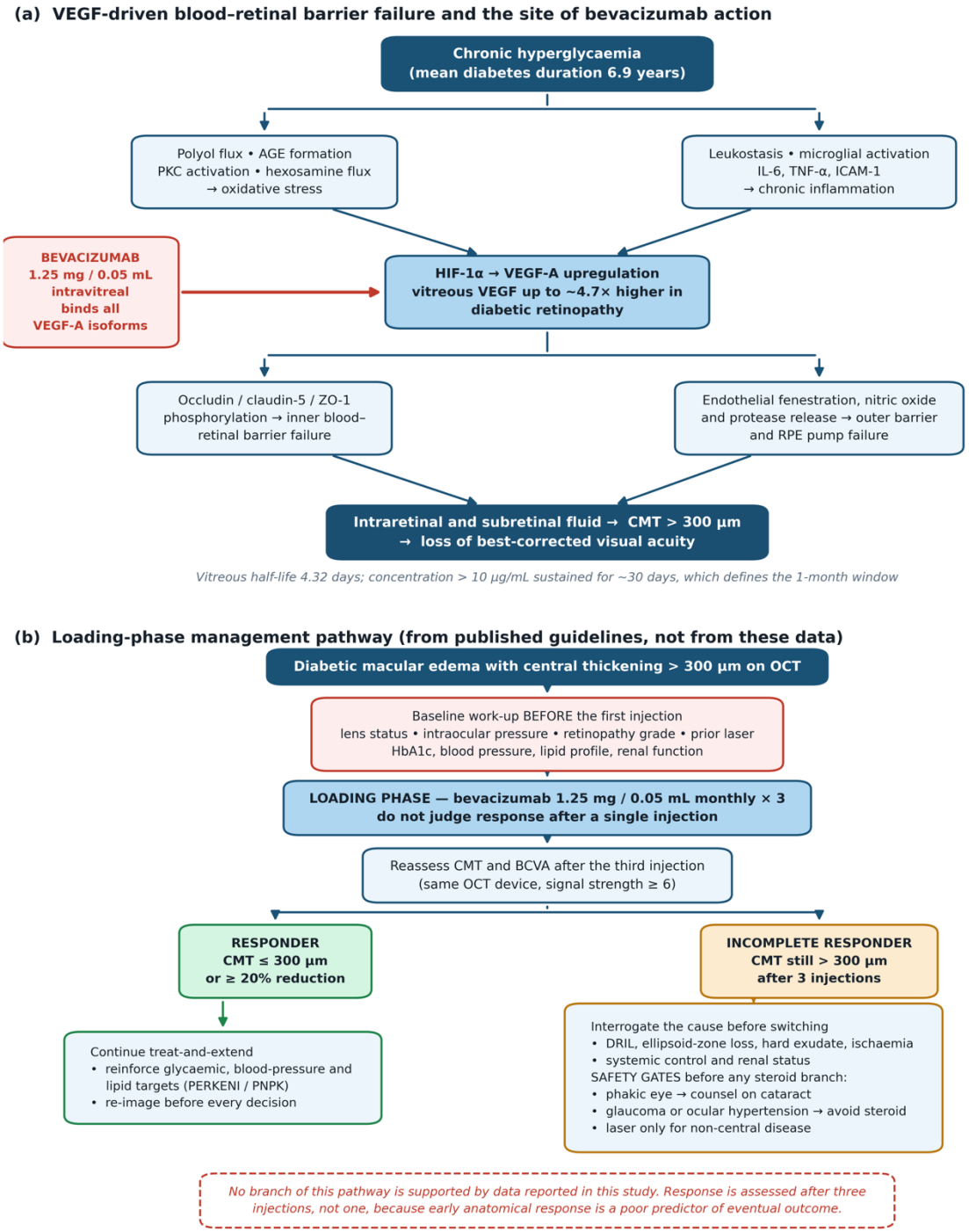

Figure 5. Mechanism and proposed pathway. (a) Chronic hyperglycaemia drives oxidative and inflammatory injury that converges on VEGF-A upregulation and tight-junction failure, producing the central thickening that bevacizumab is given to reverse; the vitreous pharmacokinetics of bevacizumab define the one-month assessment window. (b) A loading-phase management pathway drawn from the cited guidelines, not from the data in this paper. Response is assessed after three monthly injections rather than one, because early anatomical response is a poor predictor of eventual outcome; the safety nodes are explicit because a corticosteroid branch in a phakic eye carries avoidable cataract and a material risk of intraocular pressure rise. No branch of this pathway is supported by data reported in this study.

For Indonesian practice these findings carry two implications, and one non-implication. The non-implication first: 14 eyes cannot support a formulary decision. The case for bevacizumab as first-line therapy for DME within the Jaminan Kesehatan Nasional benefit package rests on the randomised and network-synthesised evidence and on the cost differential between agents[15,23], all of it external to this paper; our series is consistent with that case but adds nothing to it. What the series does show is, first, that a clinically meaningful anatomical response is achievable in this setting with the agent that is actually available, and second, that the binding constraint appears to be access rather than efficacy. Only 11 of 27 attending patients (40.7%) were treated, for reasons that were organisational rather than clinical, as Figure 1 records: intermittent drug supply, the need to accumulate a procedural quota before an injection list could run, and loss to follow-up among patients travelling from across West Sumatra — a pattern with documented consequences for diabetic eye disease elsewhere[26,27]. We are careful here: because the exclusion criteria were never applied to the 16 untreated patients, we cannot state that all of them were eligible. Third, the 54.5% of patients with HbA1c above the PERKENI target in Table 1 underline that ophthalmic and metabolic care must run together, since systemic

control governs the trajectory the injections are trying to modify[39].

Practically, four service-level changes would follow, and the first three cost nothing. Record laterality, per-eye HbA1c and per-eye outcome status on the clinical form, which would remove every non-identification documented in this paper at a stroke. Record retinopathy severity, DME morphological subtype, lens status, prior laser or intravitreal therapy and injection number — the variables listed in the last row of Table 1 — without which no response analysis can be interpreted. Replace Snellen categories with ETDRS letter scores[36,38]. Then, structurally: predictable injection scheduling that does not depend on quota assembly, and structured co-management with internal medicine so that HbA1c, blood pressure, lipid and renal values are documented at every injection visit[63,64]. Wider adoption of image-based or algorithmic screening, now demonstrated at national scale elsewhere in South-East Asia[65], would address the upstream detection gap, and Indonesian groups have already reported adjunctive treatment strategies that could be evaluated in the same framework[25].

This study has several strengths. It analyses a complete consecutive series with no eye excluded and no missing paired data, so selection within the treated group is absent. All inference is exact rather than asymptotic and every enumeration is exhaustive rather than simulated, which is the appropriate choice at n = 14 and removes Monte Carlo error from the reported figures entirely. Every estimate is accompanied by an interval, an effect size and an explicit robustness analysis covering both eye-in-patient clustering and reversal of individual observations, as Tables 2 and 3 set out. The unit of analysis is stated for every proportion in Table 1, which resolves a recurring ambiguity in ophthalmic reporting[30]. Quantities that are not identified are bounded rather than assumed away, including one — the functional transition table — where doing so materially weakens the paper's own conclusion. And the arithmetic consistency audit of the source tabulations in Table 6, including the joint mean-and-standard-deviation test[42], is reported in full with its assumptions stated.

The limitations are substantial and constrain interpretation accordingly. First, the sample of 14 eyes from 11 patients is small; the primary estimate carries a confidence interval 45 percentage points wide, as Figure 4d makes visible, and no analysis here can be more than descriptive. Second, the design is single-arm and uncontrolled, so regression to the mean, natural fluctuation of macular thickness and observer expectation cannot be separated from drug effect; with a single pre-injection scan and threshold-based selection, regression to the mean is not merely a caveat but an unquantified competing explanation. Third, the analysis was performed on aggregate tabulations rather than an eye-level file, which is the direct source of the non-identification reported in sections 3.3, 3.5 and 3.6; if the primary record can be re-abstracted, those bounds should be replaced by point estimates. Fourth, follow-up extends to one month, which characterises early response and says nothing about durability, re-treatment need or safety — with zero events in 14 eyes the upper 95% bound on any complication rate is 23.2%. Fifth, the number of injections per eye and whether the index injection was the first are not recorded, so we cannot describe the exposure as a single dose. Sixth, acuity was measured on a Snellen chart rather than an ETDRS chart, which compresses the outcome scale and precludes comparison in letters[36]. Seventh, no operational criterion for central involvement beyond CMT above 300 µm is documented, and the clinical variables listed in the final row of Table 1 are all unrecorded, so the eyes cannot be characterised clinically and the mechanisms invoked above remain literature-derived hypotheses. Finally, all 11 patients were treated at a single tertiary referral centre in West Sumatra, so the results describe a referred population and should not be generalised to community practice.

5. Conclusion

Intravitreal bevacizumab 1.25 mg produced a large reduction in central macular thickness one month later in eyes with diabetic macular edema at a West Sumatran tertiary referral centre: mean CMT fell by 209 µm (40%, Glass's Δ 1.37) and 64.3% of eyes (95% CI 38.8-83.7) reached 300 µm or less. That anatomical finding is identified by the available data and survives the most conservative assumption about correlation between fellow eyes (worst-case exact P = 0.031). The functional finding is suggestive but not established: its paired table is not identified, and across the admissible forms the exact P for the acuity category shift spans 0.008 to 0.057, widening to 0.227 once clustering is added. The series adds no new efficacy evidence to a question already settled by randomised trials, but it does show that a clinically meaningful response is achievable with the agent actually available in Indonesian public practice, and that the binding constraint is access rather than efficacy, since fewer than half of attending patients were treated at all. The study could not and did not establish any role for diabetes duration or glycaemic control in modifying that response: those associations are formally unidentified, and no between-stratum difference at all is detectable at 80% power in this design. Answering that question does not require a larger study so much as a better record: adding laterality, per-eye HbA1c and per-eye outcome status to the clinical form would remove every non-identification reported here at no cost, and a prospective series of at least 82 eyes with ETDRS acuity, a full loading phase and follow-up beyond one month would then answer it.

Declarations

Ethical approval

This study received ethical approval from the Ethics Committee of CMHC Research Center, Indonesia (Ref no. CMHC/VII/2022/108). The study was conducted in accordance with the Declaration of Helsinki. The analysis used aggregate tabulations derived from data already recorded during routine clinical care; no individually identifiable record was accessed.

Consent for publication

Not applicable; no individually identifiable participant data are presented.

Availability of data and materials

The de-identified aggregate data and the complete analysis script that generated every number, table and figure in this article are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Authors' contributions

YWE conceived the study, collected and curated the clinical data, performed the analysis and drafted the manuscript. WH supervised the vitreoretinal management, verified the clinical data and critically revised the manuscript. KR supervised the study, interpreted the findings and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The authors thank the staff of the Vitreoretina Subdivision, Department of Ophthalmology, Dr. M. Djamil Central General Hospital, Padang, for maintaining the clinical records on which this analysis depends.

6. References

1. Teo ZL, Tham YC, Yu M, et al. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045. Ophthalmology. 2021;128(11):1580-1591. doi:10.1016/j.ophtha.2021.04.027

2. Hashemi H, Rezvan F, Pakzad R, et al. Global and Regional Prevalence of Diabetic Retinopathy; A Comprehensive Systematic Review and Meta-analysis. Seminars in Ophthalmology. 2022;37(3):291-306. doi:10.1080/08820538.2021.1962920
3. Cai K, Liu Y, Wang D. Prevalence of diabetic retinopathy in patients with newly diagnosed type 2 diabetes: A systematic review and meta-analysis. Diabetes Metabolism Res. 2023;39(1). doi:10.1002/dmrr.3586
4. Wondmeneh TG, Mohammed JA. Prevalence of diabetic retinopathy and its associated risk factors among adults in Ethiopia: a systematic review and meta-analysis. Sci Rep. 2024;14(1). doi:10.1038/s41598-024-78596-9
5. Marques-Couto P, Afonso J, Coelho-Costa I, et al. Prevalence and Risk Factors of Diabetic Macular Edema in Type 1 Diabetes. Ophthalmology Retina. 2026;10(3):294-309. doi:10.1016/j.oret.2025.09.008
6. Varghese JS, Ravi Kumar V, Bartelt J, et al. The Role of Urban Residence, Race and Ethnicity, and Glycemic Control in Receiving Standards of Care and Progression to Vision-Threatening Diabetic Retinopathy. Diabetes Care. 2025;48(1):29-37. doi:10.2337/dci24-0024
7. Halim A, Syumarti S, Rini M, et al. Prevalence and Associated Factors of Diabetic Retinopathy in People with Type 2 Diabetes Attending Community Based Diabetic Retinopathy Screening in Greater Bandung, Indonesia. IJRetina. 2022;5(1):1. doi:10.35479/ijretina.2022.vol005.iss001.172
8. Widyaputri F, Rogers SL, Kandasamy R, et al. Global Estimates of Diabetic Retinopathy Prevalence and Progression in Pregnant Women With Preexisting Diabetes. JAMA Ophthalmol. 2022;140(5):486. doi:10.1001/jamaophthalmol.2022.0050
9. Inada M, Xu H, Takeuchi M, et al. Microglia increase tight-junction permeability in coordination with Müller cells under hypoxic condition in an in vitro model of inner blood-retinal barrier. Experimental Eye Research. 2021;205:108490. doi:10.1016/j.exer.2021.108490
10. Shu Y, Zhang C, Bi Y, et al. Hyperreflective foci and subretinal fluid predicts microglia activation involved in the breakdown of outer blood-retinal barrier in treatment-naïve patients with diabetic macular edema. Asia-Pacific Journal of Ophthalmology. 2025;14(5):100168. doi:10.1016/j.apjo.2025.100168
11. Qin YJ, Zhang YL, Zhang YQ, et al. Elevated level of uric acid, but not glucose, in aqueous humor as a risk factor for diabetic macular edema in patients with type 2 diabetes. Retina. 2022;42(6):1121-1129. doi:10.1097/iae.0000000000003424
12. Wang Y, Ding Y, Zhuang Q, et al.. Comparison of the cytokines levels in aqueous humor in vitrectomized eyes versus non-vitrectomized eyes with diabetic macular edema. Int Ophthalmol. 2024;44(1). doi:10.1007/s10792-024-03136-3
13. Pradhana D, Priya MN S, Surya J, et al.. Optical Coherence Tomography–Based Prevalence of Diabetic Macular Edema and its Associated Risk Factors in Urban South India: A Population–Based Study. Ophthalmic Epidemiology. 2022;29(2):149-155. doi:10.1080/09286586.2021.1907846
14. Virgili G, Curran K, Lucenteforte E, et al.. Anti-vascular endothelial growth factor for diabetic macular oedema: a network meta-analysis. Cochrane Database of Systematic Reviews. 2023;2023(6). doi:10.1002/14651858.cd007419.pub7
15. The Diabetic Retinopathy Clinical Research Network. Aflibercept, Bevacizumab, or Ranibizumab for Diabetic Macular Edema. N Engl J Med. 2015;372(13):1193-1203. doi:10.1056/nejmoa1414264
16. Rush RB, Rush SW. Faricimab for Treatment-Resistant Diabetic Macular Edema. OPTH. 2022;Volume 16:2797-2801. doi:10.2147/opth.s381503
17. Shimura M, Oh H, Ueda T, et al.. Efficacy, durability, and safety of faricimab with extended dosing up to every 16 weeks in diabetic macular edema: 2-year results from the Japan subgroup of the phase 3 YOSEMITE trial. Jpn J Ophthalmol. 2024;68(5):511-522. doi:10.1007/s10384-024-01078-y
18. Cheng Z, Liu X. Comparing the efficacy of glucocorticoids and anti-VEGF in treating diabetic macular edema: systematic review and comprehensive analysis. Front. Endocrinol.. 2024;15. doi:10.3389/fendo.2024.1342530
19. Chen L, Chen M, Huang W. Comparative efficacy and safety of anti-VEGF agents with and without Laser therapy for diabetic macular edema: A network meta-analysis. European Journal of Ophthalmology. 2025;35(6):2347-2360. doi:10.1177/11206721251350017
20. Grad J, Hatamnejad A, Dadak R, et al. Anti-VEGF Monotherapy vs Anti-VEGF and Steroid Combination Therapy for Diabetic Macular Edema: A Meta-analysis. Journal of VitreoRetinal Diseases. 2025;9(1):70-83. doi:10.1177/24741264241280597
21. Ciulla TA, Hussain RM, Taraborelli D, et al. Longer-Term Anti-VEGF Therapy Outcomes in Neovascular Age-Related Macular Degeneration, Diabetic Macular Edema, and Vein Occlusion-Related Macular Edema. Ophthalmology Retina. 2022;6(9):796-806. doi:10.1016/j.oret.2022.03.021
22. Ciulla TA, Kapik B, Grewal DS, et al.. Visual Acuity in Retinal Vein Occlusion, Diabetic, and Uveitic Macular Edema. Ophthalmology Retina. 2021;5(7):633-647. doi:10.1016/j.oret.2020.10.016
23. Boz AAE. Comparison of costs of bevacizumab, ranibizumab, and aflibercept in loading dose administration for diabetic macular edema treatment: Observational cost analysis. Eur Eye Res. 2024;:85-89. doi:10.14744/eer.2023.55265
24. Aldokhail L, Alhadlaq A, Alaradi L, et al.. Outcomes of Anti-VEGF Therapy in Eyes with Diabetic Macular Edema, Vein Occlusion-Related Macular Edema, and Neovascular Age-Related Macular Degeneration: A Systematic Review. OPTH. 2024;Volume 18:3837-3851. doi:10.2147/opth.s489114
25. Lubis PM, Prabaniswara MP, Victor AA. Comparison of micropulse subthreshold laser plus anti-VEGF versus anti-VEGF alone in diabetic macular edema: Systematic review. Indian Journal of Ophthalmology. 2023;71(11):3448-3453. doi:10.4103/ijo.ijo_519_23
26. Khurana RN, Wang JC, Zhang S, et al.. Loss to Follow up in Patients with Proliferative Diabetic Retinopathy Treated with Anti-VEGF Therapy and/or Panretinal Photocoagulation in the United States. Ophthalmology Retina. 2024;8(10):953-961. doi:10.1016/j.oret.2024.04.016
27. Subhi Y, Potapenko I, Hajari JN, et al.. Anti-VEGF Treatment of Diabetic Macular Edema in Denmark: Incidence, Burden of Therapy, and Forecasting Analyses. JPM. 2023;13(3):546. doi:10.3390/jpm13030546
28. Gilio A, Sanfilippo G. Compound conditionals, Fréchet-Hoeffding bounds, and Frank t-norms. International Journal of Approximate Reasoning. 2021;136:168-200. doi:10.1016/j.ijar.2021.06.006
29. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.ad
30. Ying GS, Maguire MG, Glynn R, et al.. Tutorial on Biostatistics: Statistical Analysis for Correlated Binary Eye Data. Ophthalmic Epidemiology. 2018;25(1):1-12. doi:10.1080/09286586.2017.1320413
31. Karimi S, Fakhri N, Ansari I, et al.. Incidence and management of acute endophthalmitis after intravitreal injection of bevacizumab. Int Ophthalmol. 2022;42(6):1827-1833. doi:10.1007/s10792-021-02180-7
32. Seong HJ, Park YM, Kim J, et al.. Incidence of Acute Endophthalmitis after Intravitreal Antivascular Endothelial Growth Factor Injection in Age-related Macular Degeneration. Korean J Ophthalmol. 2022;36(5):435-442. doi:10.3341/kjo.2022.0088
33. Vujosevic S, Alovisi C, Piccoli G, et al.. Severity of Disorganization of Retinal Layers and Visual Function Impairment in Diabetic Retinopathy. Ophthalmology Retina. 2024;8(9):880-888. doi:10.1016/j.oret.2024.04.005
34. Chaturvedi S, Paul A, Singh S, et al.. The Ellipsoid Zone Is a Structural Biomarker for Visual Outcomes in Diabetic Macular Edema and Macular Hole Management. Vision. 2025;9(1):4. doi:10.3390/vision9010004
35. Gu Z, Xi T, Zhang C, et al. Optical coherence tomography classifications of diabetic macular edema and response to aflibercept: One-year follow-up outcomes in a Chinese population. Medicine. 2023;102(4):e32815. doi:10.1097/md.0000000000032815
36. Baker CW, Josic K, Maguire MG, et al. Comparison of Snellen Visual Acuity Measurements in Retinal Clinical Practice to Electronic ETDRS Protocol Visual Acuity Assessment. Ophthalmology. 2023;130(5):533-541. doi:10.1016/j.ophtha.2022.12.008
37. Yu HJ, Kaiser PK, Zamora D, et al. Visual Acuity Variability: Comparing Discrepancies between Snellen and ETDRS Measurements among Subjects Entering Prospective Trials. Ophthalmology Retina. 2021;5(3):224-233. doi:10.1016/j.oret.2020.04.011
38. Salabati M, Huang C, Kamalipour A, et al.. Magnitude of Visual Acuity Change with ETDRS versus Snellen Testing in Clinical Trials. Ophthalmology Science. 2024;4(2):100372. doi:10.1016/j.xops.2023.100372
39. Kim HU, Park SP, Kim YK. Long-term HbA1c variability and the development and progression of diabetic retinopathy in subjects with type 2 diabetes. Sci Rep. 2021;11(1). doi:10.1038/s41598-021-84150-8
40. Fagerland MW, Lydersen S, Laake P. Recommended tests and confidence intervals for paired binomial proportions. Statist. Med.. 2014;33(16):2850-2875. doi:10.1002/sim.6148
41. Vargas M, Marra A, Buonanno P, et al.. Fragility Index and Fragility Quotient in Randomized Controlled Trials on Corticosteroids in ARDS Due to COVID-19 and Non-COVID-19 Etiology. JCM. 2021;10(22):5287. doi:10.3390/jcm10225287
42. Brown NJL, Heathers JAJ. The GRIM Test. Social Psychological and Personality Science. 2017;8(4):363-369. doi:10.1177/1948550616673876
43. Choi SW, Lee SH, Kim MS. Real-world outcomes of anti-VEGF and corticosteroid therapies for diabetic macular edema: up to 10 years

of follow-up in Korean patients. Jpn J Ophthalmol. 2026;70(3):579-588. doi:10.1007/s10384-025-01305-0
44. Mirshahi R, Falavarjani KG, Molaei S, et al.. Macular microvascular changes after intravitreal bevacizumab injection in diabetic macular edema. Canadian Journal of Ophthalmology. 2021;56(1):57-65. doi:10.1016/j.jcjo.2020.07.004
45. Sędziak-Marcinek B, Wylęgała A, Chełmecka E, et al. How to Achieve Near-Normal Visual Acuity with Bevacizumab in Diabetic Macular Edema Patients. JCM. 2021;10(16):3572. doi:10.3390/jcm10163572
46. Dweikat A, Jarrar A, Akkawi M, et al. Baseline Subfoveal Choroidal Thickness as a Predictor for Response to Short-Term Intravitreal Bevacizumab Injections in Diabetic Macular Edema. OPTH. 2021;Volume 15:4175-4180. doi:10.2147/opth.s325951
47. Usui-Ouchi A, Tamaki A, Sakanishi Y, et al. Factors Affecting a Short-Term Response to Anti-VEGF Therapy in Diabetic Macular Edema. Life. 2021;11(2):83. doi:10.3390/life11020083
48. Katić K, Katić J, Kumrić M, et al. The Predictors of Early Treatment Effectiveness of Intravitreal Bevacizumab Application in Patients with Diabetic Macular Edema. Diagnostics. 2024;14(10):992. doi:10.3390/diagnostics14100992
49. Zhou J, Ma H, Zhou X, et al. Two-Week Central Macular Thickness Reduction Rate >37% Predicts the Long-Term Efficacy of Anti-vascular Endothelial Growth Factor Treatment for Macular Edema Secondary to Retinal Vein Occlusion. Front. Med.. 2022;9. doi:10.3389/fmed.2022.851238
50. Wirkkala J, Kubin AM, Ohtonen P, et al. Real-world outcomes of early and deferred anti-VEGF treatment in diabetic macular oedema in patients with type 1 diabetes. Annals of Medicine. 2025;57(1). doi:10.1080/07853890.2025.2530218
51. Daldal H, Turkyilmaz M, Balikoglu Yilmaz M, et al. The Effect of Ranibizumab Loading Treatment on Vision-Related Quality of Life in Diabetic Macular Edema. Clinics and Practice. 2021;11(3):659-670. doi:10.3390/clinpract11030081
52. Wirkkala J, Kubin AM, Ohtonen P, et al. Visual outcomes of observation, macular laser and anti-VEGF in diabetic macular edema in type 1 diabetes: a real-world study. BMC Ophthalmol. 2022;22(1). doi:10.1186/s12886-022-02482-z
53. Hsieh TC, Deng GH, Chang YC, et al. A real-world study for timely assessing the diabetic macular edema refractory to intravitreal anti-VEGF treatment. Front. Endocrinol.. 2023;14. doi:10.3389/fendo.2023.1108097
54. Tripathi A, Gaur S, Agarwal R, et al. Disorganization of retinal inner layers as an optical coherence tomography biomarker in diabetic retinopathy: A review. Indian Journal of Ophthalmology. 2025;73(9):1245-1250. doi:10.4103/ijo.ijo_800_25
55. Alvarez-Guzman C, Bustamante-Arias A, Colorado-Zavala MF, et al. The impact of central foveal thickness and integrity of the outer retinal layers in the visual outcome of uveitic macular edema. Int J Retin Vitr. 2021;7(1). doi:10.1186/s40942-021-00306-8
56. Okudan S, Acar Duyan S, Erdem A, et al. Optical Coherence Tomography Biomarkers Predict the Long-Term Restorative Effect of Early Anti-VEGF Treatment on Diabetic Macular Edema. Life. 2025;15(2):269. doi:10.3390/life15020269
57. Yoshida M, Murakami T, Nishikawa K, et al. Severity Scale of Diabetic Macular Ischemia Based on the Distribution of Capillary Nonperfusion in OCT Angiography. Ophthalmology Science. 2025;5(1):100603. doi:10.1016/j.xops.2024.100603
58. Elnahry AG, Elnahry GA. Optical Coherence Tomography Angiography of Macular Perfusion Changes after Anti-VEGF Therapy for Diabetic Macular Edema: A Systematic Review. Journal of Diabetes Research. 2021;2021:1-14. doi:10.1155/2021/6634637
59. Massengill MT, Cubillos S, Sheth N, et al.. Response of Diabetic Macular Edema to Anti-VEGF Medications Correlates with Improvement in Macular Vessel Architecture Measured with OCT Angiography. Ophthalmology Science. 2024;4(4):100478. doi:10.1016/j.xops.2024.100478
60. Salimi A, Vila N, Modabber M, et al.. One-year outcomes of Aflibercept for refractory diabetic macular edema in Bevacizumab nonresponders. Indian Journal of Ophthalmology. 2021;69(2):360-367. doi:10.4103/ijo.ijo_459_20
61. Pessoa B, Malheiro L, Carneiro I, et al.. Intravitreal Ranibizumab or Aflibercept After Bevacizumab in Diabetic Macular Edema: Exploratory Retrospective Analysis. OPTH. 2021;Volume 15:253-260. doi:10.2147/opth.s280644
62. Alsaedi N, Alselaimy RM, Alshamrani AA, et al.. Aflibercept versus Ranibizumab as a Second Line Therapy After Bevacizumab for Diabetic Macular Edema. OPTH. 2021;Volume 15:2975-2980. doi:10.2147/opth.s316271
63. Suzuki Y, Kiyosawa M. Relationship between Diabetic Nephropathy and Development of Diabetic Macular Edema in Addition to Diabetic Retinopathy. Biomedicines. 2023;11(5):1502. doi:10.3390/biomedicines11051502
64. Chou YB, Chang JY, Chou YJ, et al. Association between renal function and the Treatment of Diabetic Macular Edema in Long-Term Cohort Study. Sci Rep. 2024;14(1). doi:10.1038/s41598-024-77530-3
65. Ruamviboonsuk P, Tiwari R, Sayres R, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. The Lancet Digital Health. 2022;4(4):e235-e244. doi:10.1016/s2589-7500(22)00017-6