

Computational Framework for Automated Quantification of Panretinal Photocoagulation Technique from Ultra-Widefield Fundus Images

Brendan Fox

bt.fox@buffalo.edu

University at Buffalo, State University of New York

Nicholas Kemmis

University at Buffalo, State University of New York

Ryan Andrews

Boston College

Brayden Duprey

Boston College

Research Article

Keywords: Photocoagulation, Diabetic retinopathy, Ultra-widefield imaging, Panretinal photocoagulation, Quality metrics, Standardization

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10809011/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Computational Framework for Automated Quantification of Panretinal Photocoagulation Technique from Ultra-Widefield Fundus Images

Brendan Fox^{a,1,*} Nicholas Kemmis^a Ryan Andrews^b Brayden Duprey^b

^aJacobs School of Medicine and Biomedical Sciences, University at Buffalo, 955 Main Street, Buffalo, NY 14203, USA

^bBoston College, Chestnut Hill, MA 02467, USA

*Corresponding author.

E-mail: btfox@buffalo.edu

N. Kemmis: nrkemma@buffalo.edu

¹Present address: 775 Main St., Buffalo, NY 14203, Unit 509

Acknowledgments

The authors thank the public dataset contributors whose imaging made this analysis possible.

Key messages

What is known

- Panretinal photocoagulation remains a standard treatment for proliferative diabetic retinopathy, but technique is documented descriptively rather than with image-verifiable measurements.

What is new

- A quantitative framework recovers coverage, spacing, safety distances, clustering, and quadrant balance from post-procedure ultra-widefield fundus images.
- In 218 images (93,322 mapped burns), burn count varied 74-fold (range 13–959) and median coverage was 15.3% of the treatable region.
- These metrics enable standardized quality monitoring and future correlation of PRP geometry with clinical outcomes.

Purpose Panretinal photocoagulation (PRP) lacks standardized, image-verifiable metrics for procedural quality. We developed a quantitative framework to measure PRP technique from ultra-widefield fundus images.

Methods We analyzed 218 ultra-widefield images with confirmed PRP (93,322 mapped burns) from three public datasets. Using AI-assisted annotation with manual review, we extracted burn centroids and computed geometric metrics operationalizing protocol dimensions: coverage index (photocoagulated area/treatable region), spacing uniformity (1–CV of nearest-neighbor distances), safety distances to fovea and optic disc (in disc diameters), clustering index (local density variation), and quadrant balance. Metrics were compared to literature-derived thresholds from ETDRS and contemporary guidelines.

Results Substantial variability was observed across all dimensions. Median burn count was 407.5 (IQR 207.75; range 13–959). Median coverage was 15.3% of the treatable region (IQR 6.8%; range 0–33.7%), corresponding to 108.8 mm² (IQR 50.9). Median spacing uniformity was 0.728 (IQR 0.056). Median 5th-percentile safety distances were 2.71 disc diameters from the optic disc (IQR 0.87) and 0.86 from the posterior pole (IQR 0.68). Quadrant-stratified burn-count attainment ranged from 24.7% to 41.8% of scaled ETDRS thresholds.

Conclusion PRP technique can be objectively quantified from post-procedure images. Real-world delivery shows substantial variability across coverage, spacing, and safety parameters. This framework enables standardized quality monitoring, comparative effectiveness research, and correlation of technique geometry with clinical outcomes.

Keywords: Photocoagulation; Diabetic retinopathy; Ultra-widefield imaging; Panretinal photocoagulation; Quality metrics; Standardization

1 Introduction

Panretinal photocoagulation (PRP) has been a foundational treatment for proliferative diabetic retinopathy (PDR) since the Early Treatment Diabetic Retinopathy Study (ETDRS) demonstrated a substantial reduction in severe vision loss through ablation of ischemic peripheral retina [1, 2]. DRCR.net Protocol S later confirmed that PRP remains noninferior to ranibizumab for visual acuity, with differences in peripheral visual field effects and treatment burden [3]. PRP therefore continues to serve both as a standard therapy and as a comparator in modern anti-VEGF-based care.

Despite this established role, PRP lacks a standardized, image-verifiable language for procedural quality. Clinical documentation typically records that PRP was performed, sometimes with an approximate burn count, but not how it was delivered. Even in trials, Protocol S specified burn-count ranges (1200–1600 conventional; 1800–2400 pattern scan) and general parameters (burns spaced one burn width apart, extending from the vascular arcades to beyond the equator) without a mechanism to verify that those parameters were achieved [3, 4]. ETDRS similarly specified spot size, duration, spacing, and safety margins as subjective assessments during delivery [2]. Guidelines still give verbal targets—spot counts, 0.5–1 burn-width spacing, and minimum distances from the fovea and optic disc [5]—that are not recovered from the post-procedure image.

That gap has measurable consequences. Ultra-widefield studies of completed PRP found a mean treated area of $681 \pm 254.4 \text{ mm}^2$ versus an ideal $900 \pm 267 \text{ mm}^2$, leaving about 25% of the target untreated [6], yet notes would still read “PRP completed.” For the same nominal spot count, lesion diameters vary by a factor of 1.61 and treated area by 2.59 across operators [7]. Anti-VEGF therapy is documented to the milligram, volume, and visit [3]. Without geometric measurement of PRP, outcome differences between modalities may reflect delivery heterogeneity rather than treatment class.

Delivery itself varies along several protocol dimensions: spot count and density; spacing from sparse to confluent, freehand to pattern-scan [8, 9, 10]; full versus focal or staged extent [6, 11]; safety margins relative to the posterior pole and disc [5]; burn size and intensity [7, 12]; and conventional versus pattern-scan or navigated platforms [13]. Ultra-widefield imaging has begun to document incomplete peripheral coverage and inter-physician differences in lesion characteristics [6, 7], but those observations have not been translated into a standard set of

81 image-computable metrics.

82 Each post-PRP ultra-widefield (UWF) photograph contains a persistent spatial record
83 of laser scars across the visible retina. Rather than treating those images only as markers of
84 disease status, they can be read as geometric maps of delivery. Extracting burn locations relative
85 to critical anatomy allows PRP technique to be quantified retrospectively and reproducibly. We
86 operationalize coverage, spacing, safety distance, clustering, and quadrant balance as metrics
87 on post-procedure UWF images and apply the framework to a multi-source cohort of PRP-
88 confirmed photographs. The aim is not to adjudicate clinical appropriateness of individual
89 cases, particularly staged treatment, but to provide standardized units for benchmarking, quality
90 monitoring, and future correlation of treatment geometry with clinical outcomes.

91 **2 Methods**

92 **2.1 Data Sources and Cohort Construction**

93 We analyzed $n = 218$ ultra-widefield fundus images with confirmed PRP treatment,
94 representing 93,322 mapped laser burns (mean 428.1 per image). Images were pooled from
95 three publicly available datasets: the Open ultrawidefield fundus image dataset [14], UWF-RHS
96 [15], and MSHF [16]. Inclusion required visible PRP burns and adequate image quality for
97 segmentation; all images were manually reviewed. The multi-source cohort incorporates diverse
98 acquisition settings. Summary characteristics are provided in Online Resource 1, Table S1.

99 **2.2 AI-assisted Burn Mapping Pipeline**

100 PRP burns were annotated as segmentation instances using a hybrid manual-AI workflow
101 designed for scalable, high-quality mapping. Interactive label-assist from the Segment Anything
102 Model (SAM; Meta AI, 2023) [17] and SAM 3 (Meta AI) accelerated annotation while all labels
103 remained under manual control. These models were used only for interactive annotation, not to
104 generate manuscript text or to alter figures outside the mapping pipeline. Primary annotation
105 was performed by two trained annotators (R.A. and B.D.); each image was initially annotated
106 by one annotator and subsequently reviewed by the second. As the dataset grew, preliminary
107 models and SAM 3-assisted segmentation reduced manual time; every image still received final
108 manual review and quality control. For iterative model training during workflow development,
109 a standard 70/20/10 train/validation/test split was used.

Annotations were exported in COCO format, and burn centroids were computed from the segmentation masks on a per-image basis. The unit of analysis is the case/image, represented by centroid coordinates (x_i, y_i) and associated metadata (e.g., burn radius). This creates a complete procedural record from post-procedure photographs without requiring intra-procedure recording.

2.3 Eligible Treatment Region and Measurement Units

The *go_zone* is the visible retinal field in the UWF image—the area available for potential treatment. Within it, three *no-go* regions were excluded:

- Posterior pole: central macula and fovea, typically ≥ 2 disc diameters from the foveal center, to preserve central vision [5].
- Optic disc: to avoid nerve-fiber injury and associated field defects.
- Major retinal vessels, including the vascular arcades. Standard protocols treat from the arcades to beyond the equator [5]; evidence on treating smaller peripheral vessels is mixed [18].

The *treatable region* is *go_zone* minus these no-go zones and is the denominator for coverage and the spatial domain for burn-placement evaluation. Figure 1 shows the zonal overlays; Figure 2 shows the resulting burn map.

Safety distances are reported in disc diameters (DD) to match ETDRS conventions [19, 20]. Per-image optic disc diameter was the equivalent diameter of the annotated disc mask,

$$d_{\text{DD}} = 2 \times \sqrt{\frac{A_{\text{OD}}}{\pi}}, \quad (1)$$

where A_{OD} is disc pixel area. Pixel distances were divided by this d_{DD} for comparison to guideline thresholds (e.g., ≥ 2 DD from the fovea; ≥ 1 –2 DD from the disc) [5, 21]. For area in mm^2 , we assumed $d_{\text{DD}} \approx 1.5$ mm [22], enabling comparison with literature reporting absolute treated area. Computational implementation, including alternative disc-diameter estimators, is detailed in Online Resource 1.

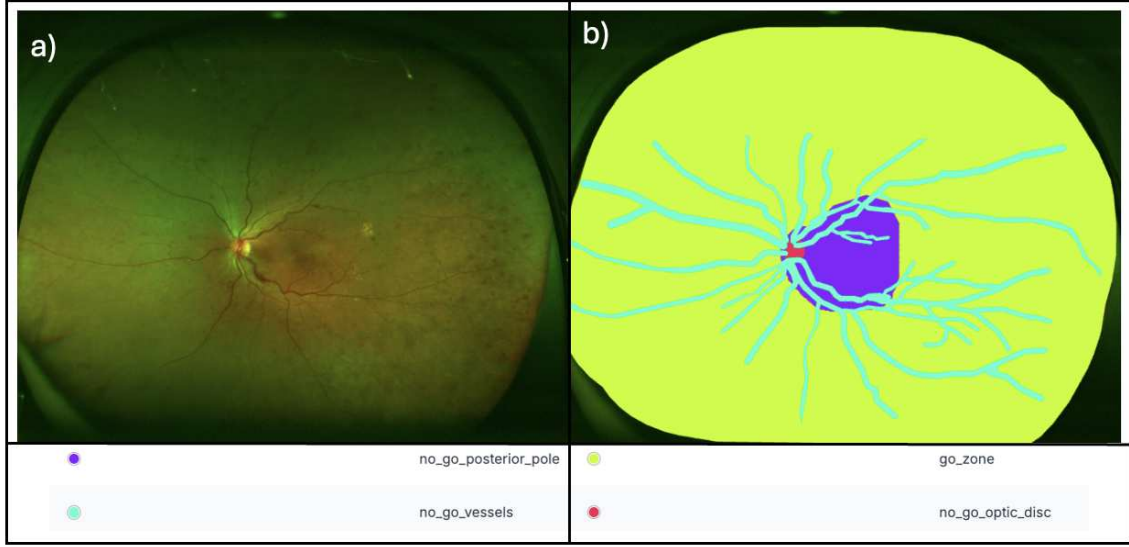


Figure 1: Eligible treatment region definition. Panel (a) shows an original ultra-widefield fundus image with visible PRP laser burns. Panel (b) displays the semantic segmentation map defining anatomical zones: *go_zone* (light green) represents the visible retinal field; *no_go_posterior_pole* (purple) excludes the central macula and fovea; *no_go_optic_disc* (red) excludes the optic disc; and *no_go_vessels* (light blue) excludes major retinal blood vessels. The treatable region for PRP is defined as *go_zone* minus all no-go zones.

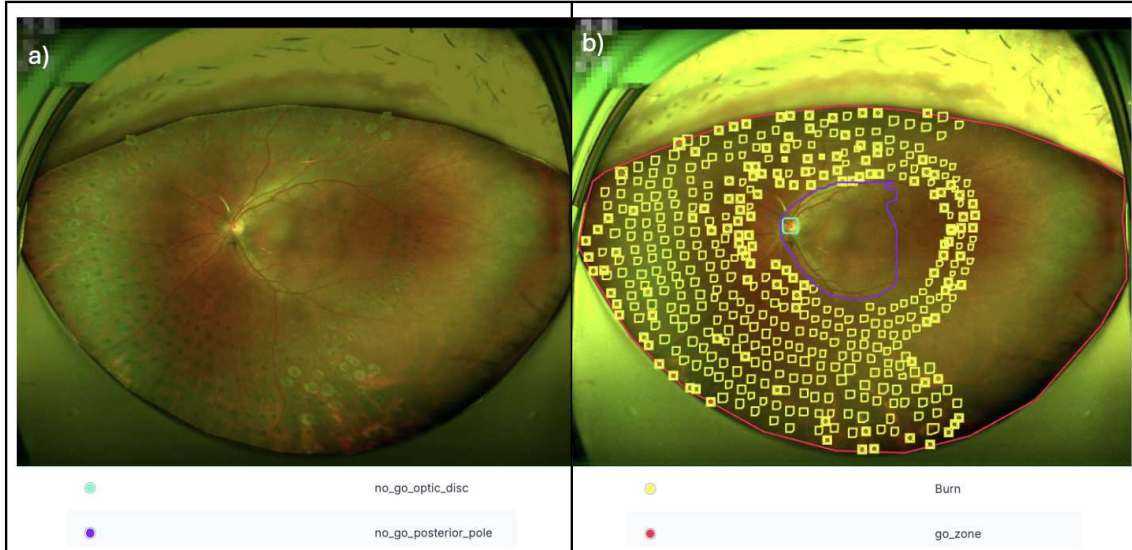


Figure 2: AI-assisted burn mapping visualization. Panel (a) shows a representative ultra-widefield fundus image. Panel (b) overlays the burn annotations (yellow boxes) within the visible retinal field, along with the *go_zone* boundary (pink) and the *no_go_posterior_pole* region (purple). This figure illustrates how post-procedure photocoagulation scars can be converted into a computable treatment map for geometric analysis.

2.4 Metric Definitions

We computed geometric metrics that operationalize PRP protocol dimensions. Table 1 lists definitions, clinical meaning, and literature anchors. Operational details follow.

Table 1: PRP quality metrics framework

Metric		Mathematical Definition	Clinical Meaning	Literature Anchor
Coverage Index (CI)		$ \text{burn} \cap \text{treatable} / \text{treatable} $	Quantifies geometric treatment extent within eligible retina. Restricted to eligible treatable region (excludes no-go zones).	ETDRS spot targets [5]; UWF adequacy study [6]
Spacing Uniformity Index (SUI)		$1 - \text{CV}(d_{\text{NN}})$	Measures consistency of inter-burn spacing relative to ETDRS heuristics	ETDRS spacing 0.5–1 burn width [5]; PI simulation [12]
Safety Index (OD/PP)	Distance	$p_{05}(\text{min distance to OD or PP})$	Conservative measure of minimum safety margin to fovea/disc	ETDRS ≥ 2 DD [5]; center-sparing [21]
Clustering Index (CLI)		$\text{CV}(\text{neighbor counts})$	Detects local density irregularity and cluster formation	Pattern-scan uniformity [23, 24]
Quadrant Balance Index (QBI)		$1 - \text{CV}(\text{quadrant counts})$	Assesses global spatial balance of PRP distribution	Full-retina intent [25, 4]
Photocoagulation Index (PI)		$ \text{burn union} / \text{go_zone} $	Accounts for burn size + overlap; area-based dose analog	Simulation range 17.9–83.1% [12]

Note: DD = disc diameter; CV = coefficient of variation; d_{NN} = nearest neighbor distance.

Coverage Index (CI). ETDRS full PRP specified 1200–1600 spots extending at least 2 DD from the fovea to the equator, and partial PRP is not recommended [5]. Ultra-widefield studies have found about 25% of target area untreated (mean 681 ± 254.4 vs ideal 900 ± 267 mm²) [6]. We define

$$\text{CI} = \frac{|\text{burn union} \cap \text{treatable region}|}{|\text{treatable region}|} \quad (2)$$

and also report coverage area (mm²) using the 1.5 mm disc reference:

$$\text{coverage_area_mm}^2 = |\text{burn union} \cap \text{treatable}| \times \left(\frac{1.5}{d_{\text{DD}}} \right)^2. \quad (3)$$

This metric quantifies geometric extent independent of clinical indication. Treatable area may be zero if *go_zone* is missing; coverage does not adjudicate clinical appropriateness (e.g., staged PRP).

Spacing Uniformity Index (SUI). ETDRS specified burns approximately one-half burn width apart [5]; photocoagulation index in simulations ranges 17.9%–83.1% with spacing [12]. Pattern-scan and navigated systems produce more regular arrays than conventional delivery [23, 24]. We define

$$\text{SUI} = 1 - \min \left(1, \frac{\sigma(d_{\text{NN}})}{\mu(d_{\text{NN}})} \right), \quad (4)$$

149 where d_{NN} is nearest-neighbor distance between centroids. Mean d_{NN} is also reported in DD
 150 for comparison with ETDRS 0.5–1 burn-width spacing. Higher SUI indicates more consistent
 151 inter-burn spacing; $\text{SUI} \geq 0.7$ is an exploratory threshold.

152 **Safety distances.** ETDRS specified at least 2 DD from the fovea [5]; center-sparing
 153 approaches used 3 DD from the fovea and 1 DD from the disc [21]. Ultra-widefield studies
 154 have reported mean distances of 6.6 ± 2.5 mm temporally and 4.2 ± 2.4 mm nasally [6]. We
 155 compute minimum, 5th-percentile, and median distances from burn centroids to the optic-disc
 156 and posterior-pole masks, converted to DD. Cohort summaries use the median and IQR of the
 157 per-image 5th percentile as a conservative margin.

158 **Clustering Index (CLI).** Local density variation reflects spacing strategy and delivery
 159 technology [23, 24, 4]. CLI is the coefficient of variation of neighbor counts within $r_{\text{cluster}} =$
 160 $8 \times r_{\text{burn}}$ (median burn radius; ≈ 4 burn diameters, capturing clusters of roughly 10–15 burns
 161 at ETDRS spacing). The multiplier $k = 8$ is a pragmatic starting point; it should be outcome-
 162 calibrated in future work. Lower CLI indicates more uniform local density regardless of k . Very
 163 sparse maps may yield degenerate counts.

164 **Quadrant Balance Index (QBI).** PRP is intended to place burns over the visible
 165 retina, sparing the central macula, from the vascular arcades to beyond the equator [25, 4].
 166 Staged regimens have used approximately 300 burns in each of the first two sittings [4]. QBI is
 167 $1 - \text{CV}$ of burn counts in four quadrants centered on the posterior-pole centroid when available
 168 (anatomical alignment), otherwise the image center. Values nearer 1 indicate more even spatial
 169 spread.

170 **Burn size.** Burn size is a major source of technique variation. Lesion diameter can
 171 vary 1.61-fold and treated area 2.59-fold for the same nominal spot count [7], with size directly
 172 affecting photocoagulation index and residual scarring [26]. Equivalent diameter is

$$d_{\text{burn}} = 2 \times \sqrt{\frac{A_{\text{burn}}}{\pi}}. \quad (5)$$

173 We report mean diameter, CV, and quartiles (q25, q50, q75) in pixels and DD.

174 **Photocoagulation Index (PI)** is the ratio of photocoagulated area to total visible
 175 retina,

$$\text{PI} = \frac{|\text{burn union}|}{|\text{go_zone}|}, \quad (6)$$

176 an area-based dose analog that accounts for burn size and overlap and can be compared with

simulation ranges of 17.9%–83.1% [12].

2.5 Field-of-View Adjustment and Staging Proxy

A single UWF image captures approximately 80–82% of the retinal surface, not the entire retina [27]. To bound the effect of this limit when scoring burn-count and area-coverage thresholds, we computed theoretical full-retina estimates by dividing observed values by 0.8 (coverage index capped at 1.0) and recalculating adherence. These values are an upper bound that assumes missing coverage is due to imaging, not incomplete treatment. The adjustment applies only to burn count and area coverage; spacing uniformity and safety margins are not adjusted because they describe properties within the visible field.

Because a single post-procedure UWF image does not encode whether PRP was delivered in one sitting or staged across sessions, we report a quadrant-based staging proxy. Each burn centroid is assigned to one of four retinal quadrants; an *active quadrant* contains at least 25 burns (to reduce sensitivity to sparse noise or edge artifacts). The number of active quadrants (1–4) indicates how spatially distributed the pattern is within a single capture.

To provide stage-consistent benchmarking, we report continuous burn-count attainment as the percentage of scaled thresholds achieved. Thresholds are proportional to the ETDRS minimum of 1200 burns: 300, 600, 900, and 1200 for 1–4 active quadrants. Attainment is $\min(\text{burn count}/\text{threshold}, 1) \times 100\%$ (e.g., 285 burns in a 1-quadrant case = 95% of 300). This scaling is consistent with staged regimens that distribute treatment over multiple sittings [4]. These thresholds are descriptive benchmarks and are not intended to determine clinical appropriateness.

2.6 Statistical Analysis

For each metric we report median, IQR, mean, SD, CV, and range. Where guidelines or prior studies provide quantitative targets (spacing heuristics, safety margins, protocol burn-count ranges), we computed cohort adherence, classified by evidence strength (strong: guidelines or consensus; moderate: reviews or multiple studies; weak: single studies; exploratory: no literature support). Source images are public; derived annotations and metrics are available from the corresponding author upon request. Analysis used Python 3 with NumPy, Pandas, OpenCV, and SciPy. Computational implementation, algorithm pseudocode, and software details are in Online Resource 1.

3 Results

3.1 Cohort Characteristics and Primary Findings

The cohort included $n = 218$ PRP-confirmed ultra-widefield fundus images, containing 93,322 total mapped laser burns. Analysis revealed substantial variability in human-delivered PRP across clinical practice: burn counts ranged from 13 to 959 (median 407.5; IQR 207.75), coverage ranged from 0% to 33.7% of the treatable region (median 15.3%; IQR 6.8%), and spacing uniformity ranged from 0.31 to 0.81 (median 0.728; IQR 0.056). All analyses reflect geometry within a single UWF capture. Summary cohort characteristics are provided in Online Resource 1, Table S1.

3.2 Coverage, Spacing, and Safety Metrics

Coverage and spacing metrics showed substantial dispersion across cases. Using burn-segmentation union within the treatable region, median coverage was 0.153 (IQR 0.068; range 0–0.337), indicating that most procedures covered less than one-fifth of the eligible treatment area. Coverage area in absolute terms (mm^2), computed using disc diameter as an internal reference, showed median 108.8 mm^2 (IQR 50.9) of treatable area covered, compared to the literature ideal of $900 \pm 267 \text{ mm}^2$ [6]. Spacing uniformity (defined as $1 - CV$ of nearest-neighbor distances) had median 0.728 (IQR 0.056; range 0.31–0.81), indicating moderate consistency in inter-burn spacing, though the wide range reveals substantial variation in distribution patterns. Mean nearest-neighbor distance, normalized to disc diameters, had median 0.543 DD (IQR 0.103).

Safety-distance summaries relative to the optic-disc and posterior-pole exclusion zones are reported in Table 2. Safety distances were computed in pixels and converted to disc diameters using the per-image optic disc diameter calculated from annotations. The 5th percentile distances provide a conservative measure of minimum safety margins: median 2.71 disc diameters (IQR 0.87) for optic disc and median 0.86 disc diameters (IQR 0.68) for posterior pole.

3.3 Variability Characterization

Online Resource 1, Table S2 presents variability measures for each metric. The coefficient of variation (CV) ranged from 0.088 (Spacing Uniformity Index) to 0.389 (Coverage Index), indicating substantial dispersion across cases. Burn count showed particularly high variability (CV 0.364), with a 74-fold range (13–959 burns). Coverage Index demonstrated the highest

Table 2: Metric summary statistics (n=218)

Metric	Median	IQR	Mean	SD	Range
Burn count	407.5	207.75	428.1	155.9	13–959
Coverage Index	0.153	0.068	0.150	0.058	0.000–0.337
Coverage area (mm ²)	108.8	50.9	108.8	41.1	–
Photocoagulation Index	0.150	0.061	0.153	0.054	–
Spacing Uniformity Index	0.728	0.056	0.716	0.063	0.310–0.814
Mean NN distance (DD)	0.543	0.103	0.542	0.082	–
Burn diameter mean (DD)	0.378	0.083	0.375	0.063	–
Burn diameter CV	0.148	0.050	0.175	0.124	–
Clustering Index	0.236	0.041	0.243	0.038	0.160–0.461
Quadrant Balance Index	0.894	0.118	0.856	0.118	0.250–0.980
Safety OD (5th percentile, DD)	2.71	0.87	–	–	–
Safety PP (5th percentile, DD)	0.86	0.68	–	–	–

Note: DD = disc diameter. CV = coefficient of variation. NN = nearest neighbor. Safety distance metrics are summarized as per-image 5th percentile distances; we report median and IQR (mean/SD not reported).

variability (CV 0.389), ranging from 0% to 33.7% of the treatable region.

3.4 Comparison to Literature Standards

Where literature supports specific quantitative thresholds, we computed strict adherence rates to these criteria (Online Resource 1, Table S3). Because historical thresholds (e.g., ETDRS-era burn-count targets) may not map cleanly to staged regimens or modern practice patterns within a single post-procedure image, we complement strict adherence with quadrant-stratified continuous attainment, computed as the percentage of scaled thresholds achieved (1 active quadrant: 300 spots; 2: 600 spots; 3: 900 spots; 4: 1200 spots). Quadrant-stratified continuous attainment showed median values of 24.7% for 1-quadrant cases (n=2), 41.8% for 2-quadrant cases (n=12), 33.2% for 3-quadrant cases (n=19), and 37.0% for 4-quadrant cases (n=184). Overall burn-count attainment across all cases was 34.0% of the ETDRS minimum (1200 spots). Active quadrants are defined as containing at least 25 burns, using quadrant boundaries centered on the posterior pole when available.

4 Discussion

4.1 A Reproducible Framework for Quantitative PRP Assessment

To our knowledge, this study establishes the first widely reproducible standard for quantitative review of PRP treatment quality from post-procedure fundus images. By operational-

izing clinical guidelines into image-computable geometric metrics, we provide a measurement approach that can be applied consistently across datasets, imaging systems, and institutions without intra-procedure recording. A motivating insight is that AI-assisted segmentation of photocoagulation scars makes it possible to reconstruct treatment records from ordinary fundus photographs, opening large-scale quality review of procedures performed in routine clinical care.

Our results reveal notable variability across all dimensions examined. Burn counts ranged from 13 to 959 (median 407.5), coverage from 0% to 33.7% (median 15.3%), and spacing uniformity from 0.31 to 0.81 (median 0.728)—findings consistent with prior reports of inter-physician variability [6, 7]. Observed variability alone does not indicate substandard care. Staged PRP protocols, patient tolerance, clinical indication, and imaging field limitations can each legitimately contribute to this range, and our metrics reflect what is measurable within a single visible fundus image, not a definitive adjudication of total retinal treatment.

4.2 From Binary Adherence to Continuous Attainment Benchmarks

Because a single capture may reflect staged regimens, a partial field, or practice that differs from legacy trial protocols, strict binary adherence to historical thresholds (e.g., ETDRS 1200–1600 spots; $\geq 75\%$ coverage) is informative as a comparator but misleading as a headline. Continuous attainment—overall burn-count attainment of 34.0% of the ETDRS minimum and quadrant-stratified medians of 24.7%–41.8% (Online Resource 1, Table S3)—shows how close typical images come to those minima under explicit, image-level assumptions.

Dividing observed counts and coverage by 0.8 is a bounded sensitivity analysis for the $\sim 80\%$ retinal surface visible on UWF imaging. It assumes uniform periphery and no multi-session accumulation, which a single image cannot verify. We therefore read these results as discordance between image-measured geometry and historical parameters, not as clinical failure. The contribution is a scalable measurement framework for later outcome-linked calibration of thresholds and prospective quality benchmarking.

4.3 Implications for Comparative Effectiveness Research

Existing documentation records that PRP occurred and, at most, an approximate burn count. It does not capture the spatial geometry that determines extent, safety, and consistency. A central feature of this approach is retrospective applicability: post-procedure ultra-widefield images contain a persistent spatial record of photocoagulation scars. Treating those scars as

geometric data rather than qualitative markers yields coverage, spacing uniformity, safety margins, and quadrant balance from photographs already acquired in routine care, without intra-procedure instrumentation.

That capacity matters for comparative effectiveness. Outcome differences between PRP and pharmacologic therapy are difficult to interpret when PRP geometry is unmeasured [27]. Quantitative studies have shown that procedures nominally identical in burn count can differ more than 2.5-fold in treated retinal area [7]. Without controlling for delivery geometry, technique heterogeneity cannot be separated from treatment-class effects. Image-derived metrics put PRP closer to the dose, volume, and interval documentation used for anti-VEGF therapy, enabling more rigorous analysis of modality-specific effects.

4.4 Toward Empirical Threshold Calibration and Outcomes Correlation

Guideline burn-count ranges, spacing heuristics, and safety margins were established under documentation constraints that precluded geometric verification [28, 5]. Whether those thresholds predict clinically meaningful outcomes—and whether they require refinement for specific disease stages, delivery technologies, or patient populations—are questions that can now be approached systematically. Prospective and retrospective studies linking these metrics to neovascularization regression, vitreous hemorrhage recurrence, and visual field preservation in imaged cohorts would provide an empirical basis for quality benchmarks [29, 30, 31, 32]. Whether spacing uniformity above a given threshold predicts less persistent neovascularization, or whether coverage correlates with time to PDR stabilization, can be tested in existing imaging cohorts with documented outcomes.

Prior work already suggests that adequacy is outcome-linked: regression of proliferative disease has been shown to correlate with cumulative laser burn count, with regressed eyes receiving substantially higher burn totals than non-regressed eyes [31]. Studies quantifying retinal ablation area found that progressively greater ablation was required for regression across increasing PDR severity—with severe PDR requiring ablation areas substantially higher than the median 108.8 mm² observed here [30]. These findings suggest that the geometric parameters measured by this framework may have direct clinical relevance for predicting treatment success, and that the attainment benchmarks reported here may inform future discussions of minimum adequate delivery.

313 4.5 Clinical Implementation and Quality Monitoring

314 At the institutional level, standardized geometric metrics could support quality improve-
315 ment by characterizing current practice and identifying systematic departures from target pa-
316 rameters, using existing image archives without changing clinical workflow. Retrospective appli-
317 cation to image banks would allow benchmarking against evidence-based thresholds once those
318 thresholds are empirically established. The same measures enable technology-specific evaluation:
319 pattern-scan and freehand delivery produce different geometries, which spacing uniformity and
320 clustering can quantify [23]. Delivery platform may influence outcomes through more than burn
321 count alone [29], which argues for a full geometric profile rather than a single parameter.

322 4.6 Broader Implications for Image-Guided Interventions

323 More broadly, this work shows that procedural quality in laser-based retinal treatment
324 can be expressed in standardized, image-derivable units. The transition from qualitative pro-
325 cedure documentation to quantifiable geometric characterization is a necessary precondition for
326 reproducible quality monitoring, inter-institutional benchmarking, and outcome-linked research
327 at scale. Photocoagulation scars, previously regarded as qualitative markers of completed treat-
328 ment, can now be analyzed as structured spatial records of delivery geometry—enabling ret-
329 rospective procedural measurement from imaging already archived in routine practice. The
330 framework presented here provides a methodological foundation for that transition, with po-
331 tential applicability to other image-guided interventions where post-procedure imaging contains
332 persistent spatial records of treatment delivery.

333 4.7 Limitations

334 Several limitations should be considered. This analysis is retrospective and observational;
335 causal relationships between technique parameters and clinical outcomes cannot be established.
336 Fundus images may not capture the full peripheral extent of PRP, potentially underestimating
337 coverage, and image quality variability may affect burn detection reliability. All images under-
338 went manual review, but some peripheral burns may remain undetected. Analysis is limited
339 to a single timepoint per image, precluding assessment of staged or multi-session treatment
340 progression. Safety distances are expressed in per-image disc diameters estimated from anno-
341 tations, which may vary with imaging modality and magnification. The cohort is pooled from
342 three publicly available datasets to maximize generalizability, which precludes source-specific

analyses and limits cohort diversity. Burn detection relies on a hybrid manual-AI workflow; systematic detection errors would propagate to metric computations and require evaluation in future validation studies. Burn intensity classification is deferred to future work. Additionally, fundus photography provides a 2D projection of a curved 3D retinal surface; all geometric metrics in this framework are therefore computed in image-plane coordinates. This introduces distortion in spatial measurements—particularly for coverage area and inter-burn distances in the periphery—that is not corrected in the current implementation. Computational work to project metrics onto a spherical retinal surface model is in progress. Finally, low coverage or burn counts may reflect appropriate clinical decisions (e.g., staged protocols, patient-specific contraindications) rather than suboptimal technique; interpretation requires clinical context not available in retrospective image analysis.

5 Conclusion

We present a quantitative framework for evaluating PRP laser technique using geometric metrics computed from fundus images. Analysis reveals substantial variability in human-delivered PRP across clinical practice.

Our findings demonstrate that PRP technique can be objectively measured using retrospective image analysis, that substantial variability exists in human-delivered PRP across clinical practice, and that literature-derived criteria provide a starting point for standardization, though many areas lack quantitative thresholds. By providing objective, image-computable metrics, we enable fair comparison between PRP and pharmacologic therapies, support quality monitoring and training, and create a foundation for evidence-based standardization.

The AI-assisted burn mapping pipeline we demonstrate—combining manual-reviewed annotations with SAM-based interactive label-assist—makes retrospective evaluation feasible at scale, enabling large-cohort analysis where every post-PRP fundus image provides a computable procedural record. This capability enables outcomes correlation, quality monitoring, and evidence-based standardization without requiring intra-procedure recording.

Future work should link technique metrics to clinical outcomes, develop automated quality monitoring tools to support clinical practice, and explore stratification by delivery technology (conventional vs pattern-scan vs navigated) to inform protocol optimization. The framework quantifies procedure geometry and consistency while explicitly not adjudicating clinical appro-

373 priateness, providing an honest and bounded approach to technique evaluation.

374 This work establishes that procedure quality can be objectively measured from post-
375 procedure images, creating a model for evaluating other image-guided interventions. Future
376 thresholds should be outcome-calibrated rather than historically inherited. As AI-assisted an-
377 notation tools become more sophisticated, retrospective procedural measurement may become
378 routine, enabling quality monitoring, outcomes correlation, and evidence-based standardization
379 across image-guided therapies.

Statements and Declarations

Funding

The authors did not receive support from any organization for the submitted work.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

This study used only publicly available, de-identified datasets. According to U.S. Department of Health and Human Services (HHS) and National Institutes of Health (NIH) guidance, the work does not meet the definition of human subjects research. Institutional Review Board (IRB) review and approval were therefore not required. No attempt was made to re-identify any individuals. No animal studies were performed.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

Data availability

Source ultra-widefield fundus images are publicly available from the cited datasets. Derived annotation data (COCO format) and computed metrics for the 218 images analyzed in this study are available from the corresponding author upon reasonable request, because a dedicated public repository for the derived annotation set has not yet been released. The burn segmentation module described in this manuscript will be made available at https://github.com/brendan12fox/PRP_Burn_Segmentation following additional validation.

Authors' contributions

Brendan Fox: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft, Writing – review and editing. **Nicholas Kemmis:** Conceptualization,

405 Methodology, Data curation, Formal analysis, Writing – review and editing. **Ryan Andrews:**
406 Data curation, Investigation, Writing – review and editing. **Brayden Duprey:** Data curation,
407 Investigation, Writing – review and editing.

References

- [1] Early Treatment Diabetic Retinopathy Study Research Group. Early photocoagulation for diabetic retinopathy. etdrs report number 9. *Ophthalmology*, 98(5):766–785, 1991. ETDRS report 9: PRP reduces severe vision loss by 50%.
- [2] Early Treatment Diabetic Retinopathy Study Research Group. Photocoagulation for diabetic macular edema. early treatment diabetic retinopathy study report number 1. *Archives of Ophthalmology*, 103(12):1796–1806, 1985. ETDRS report 1: PRP efficacy and protocol parameters.
- [3] Writing Committee for the Diabetic Retinopathy Clinical Research Network. Panretinal photocoagulation vs intravitreal ranibizumab for proliferative diabetic retinopathy: A randomized clinical trial. *JAMA*, 314(20):2137–2146, 2015. DRCR Protocol S: PRP non-inferiority to ranibizumab.
- [4] Diabetic Retinopathy Clinical Research Network, A. J. Brucker, H. Qin, et al. Observational study of the development of diabetic macular edema following panretinal (scatter) photocoagulation given in 1 or 4 sittings. *Archives of Ophthalmology*, 127(2):132–140, 2009. Staged treatment: 300 burns per sitting, spacing 1 burn width apart.
- [5] American Academy of Ophthalmology. Diabetic retinopathy preferred practice pattern. *Ophthalmology*, 2024. Evidence Log: ETDRS standards, 1200-1600 spots, 0.5 burn width spacing.
- [6] L S, Banerjee M, Venkatesh P, et al. Assessment of quality and adequacy of panretinal photocoagulation in proliferative diabetic retinopathy using ultra-widefield imaging. *Retina*, 43(11):1922–1927, 2023. PRP adequacy assessment, 25% untreated area, $681 \pm 254.4 \text{ mm}^2$ vs ideal $900 \pm 267 \text{ mm}^2$.
- [7] Saeger M, Heckmann J, Purtskhvanidze K, et al. Variability of panretinal photocoagulation lesions across physicians and patients. quantification of diameter and intensity variation. *Graefe’s Archive for Clinical and Experimental Ophthalmology*, 2016. Evidence Log: Variability study, 1.61-fold diameter variation, 2.59-fold area variation.
- [8] Mahiul M. K. Muqit, George R. Marcellino, David B. Henson, Lorna B. Young, Niall Patton, Stephen J. Charles, George S. Turner, and Paulo E. Stanga. Single-session vs

multiple-session pattern scanning laser panretinal photocoagulation in proliferative diabetic retinopathy: The manchester pascal study. *Archives of Ophthalmology*, 128(5):525–533, 2010. Manchester Pascal Study: pattern scan vs conventional PRP.

[9] A. V. Chappelow, K. Tan, N. K. Waheed, and P. K. Kaiser. Panretinal photocoagulation for proliferative diabetic retinopathy: pattern scan laser versus argon laser. *American Journal of Ophthalmology*, 153(2):351–357, 2012. Pattern scan vs argon laser PRP comparison.

[10] J. Nemcansky, A. Stepanov, S. Nemcanska, P. Masek, H. Langrova, and J. Studnicka. Single session of pattern scanning laser versus multiple sessions of conventional laser for panretinal photocoagulation in diabetic retinopathy: Efficacy, safety and painfulness. *PLoS One*, 14(7):e0219282, 2019. Pattern scan vs conventional: single vs multiple sessions.

[11] Russell JF, Al-Khersan H, and Shi Y. Retinal nonperfusion in proliferative diabetic retinopathy before and after panretinal photocoagulation assessed by widefield oct angiography. *American Journal of Ophthalmology*, 2020. Evidence Log: Nonperfusion assessment.

[12] Nishida K, Miura K, and Sakaguchi H. The impact of spot size, spacing, pattern, duration and intensity of burns on the photocoagulation index in a geometric simulation of pan-retinal laser photocoagulation. *Acta Ophthalmologica*, 97(4):e551–e558, 2019. Photocoagulation index simulation, range 17.9% to 83.1%.

[13] M. Kernt, E. Chee, A. Neubauer, C. Hirneiss, C. Haritoglou, M. W. Ulbig, and A. Kampik. Clinical efficacy of navigated panretinal photocoagulation in proliferative diabetic retinopathy. *Acta Ophthalmologica*, 93(6):e467–e473, 2015. Navigated laser PRP: clinical efficacy study.

[14] Shucheng He, Xin Ye, and Lijun Shen. Open ultrawidefield fundus image dataset with disease diagnosis and clinical image quality assessment [dataset], 2024. Open ultrawidefield fundus image dataset with disease diagnosis and clinical image quality assessment.

[15] Renkai Wu, Pengchen Liang, Yiqi Huang, Qing Chang, and Huiping Yao. Automatic segmentation of hemorrhages in the ultra-wide field retina: Multi-scale attention subtraction networks and an ultra-wide field retinal hemorrhage dataset. *IEEE Journal of Biomedical and Health Informatics*, 2024. UWF-RHS dataset: Ultra-wide field retinal hemorrhage segmentation dataset.

- [16] Kai Jin, Zhiyuan Gao, Xiaoyu Jiang, Yaqi Wang, Xiaoyu Ma, and Yunxiang Li. Mshf: A multi-source heterogeneous fundus (mshf) dataset for image quality assessment [dataset], 2022. MSHF dataset: Multi-source heterogeneous fundus image quality assessment.
- [17] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023. Segment Anything Model (SAM).
- [18] Moutray T, Evans JR, and Lois N. Different lasers and techniques for proliferative diabetic retinopathy. *Cochrane Database of Systematic Reviews*, 2018. Evidence Log: Laser techniques review.
- [19] Jonas JB and Papastathopoulos K. Ophthalmoscopic measurement of the optic disc. *Ophthalmology*, 102(7):1102–1106, 1995. Optic disc measurement methods.
- [20] Ishikawa A, Kusaka K, Nakanishi A, and Yokoyama Y. The performance of optic disc-macula distance to disc-diameter ratio in qualitative and quantitative assessment of the optic disc size. *Eye*, 39(5):969–976, 2025. DM/DD ratio method for disc diameter assessment.
- [21] G. P. Theodossiadis, A. Boudouri, G. Georgopoulos, and C. Koutsandrea. Central visual field changes after panretinal photocoagulation in proliferative diabetic retinopathy. *Ophthalmologica*, 201(2):71–78, 1990. Center-sparing approach: 3 DD from fovea, 1 DD from optic disc.
- [22] Long T, Xu Y, Zou H, et al. A generic pixel pitch calibration method for fundus camera via automated roi extraction. *Sensors*, 22(21):8565, 2022. Automated pixel pitch calibration without camera-specific parameters.
- [23] T. Alasil and N. K. Waheed. Pan retinal photocoagulation for proliferative diabetic retinopathy: Pattern scan laser versus argon laser. *Current Opinion in Ophthalmology*, 25(3):164–170, 2014. Pattern scan vs argon laser: well-aligned arrays.
- [24] J. Chhablani, A. Mathai, P. Rani, et al. Comparison of conventional pattern and novel navigated panretinal photocoagulation in proliferative diabetic retinopathy. *Investigative*

494 *Ophthalmology & Visual Science*, 55(6):3432–3438, 2014. Navigated vs pattern scan: uni-
495 form distribution, burn size variation.

496 [25] N. Cheung, P. Mitchell, and T. Y. Wong. Diabetic retinopathy. *Lancet*, 376(9735):124–136,
497 2010. Full PRP: entire retina, from vascular arcades to beyond equator.

498 [26] Lavinsky D, Cardillo JA, Mandel Y, et al. Restoration of retinal morphology and residual
499 scarring after photocoagulation. *Acta Ophthalmologica*, 91(4):e315–e323, 2013. Burn size
500 and residual scarring after photocoagulation.

501 [27] Jampol LM, Glassman AR, and Sun J. Evaluation and care of patients with diabetic
502 retinopathy. *New England Journal of Medicine*, 2020. Evidence Log: Diabetic retinopathy
503 evaluation.

504 [28] American Diabetes Association Professional Practice Committee. 12. retinopathy, neuropa-
505 thy, and foot care: Standards of care in diabetes-2026. *Diabetes Care*, 2025. Evidence Log:
506 ADA standards update.

507 [29] Perais JA, McCullough PG, McLaughlin GA, et al. Predictive factors associated with
508 anatomical and functional outcomes after panretinal photocoagulation in people with pro-
509 liferative diabetic retinopathy. *Retina*, 42(6):1081–1089, 2022. Predictive factors for PRP
510 outcomes; delivery technology effects on stabilization.

511 [30] Muqit MM, Marcellino GR, Henson DB, et al. Pascal panretinal laser ablation and re-
512 gression analysis in proliferative diabetic retinopathy: Manchester pascal study report 4.
513 *Eye*, 25(12):1560–1566, 2011. Manchester Pascal Study Report 4: ablation area and PDR
514 regression by severity.

515 [31] Cordeiro MF, Stanford MR, Phillips PM, and Shilling JS. Relationship of diabetic microvas-
516 cular complications to outcome in panretinal photocoagulation treatment of proliferative
517 diabetic retinopathy. *Eye*, 10(6):698–703, 1996. Burn count correlation with PDR regression
518 outcome.

519 [32] Reddy VM, Zamora RL, and Olk RJ. Quantitation of retinal ablation in proliferative
520 diabetic retinopathy. *American Journal of Ophthalmology*, 119(6):760–766, 1995. Retinal
521 ablation area quantification by PDR severity.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [OnlineResource1.pdf](#)
- [TableS2.pdf](#)
- [TableS1.pdf](#)
- [TableS3.pdf](#)