Supplementary Information

An agentic multidisciplinary framework for diabetic retinopathy treatment planning

Supplementary Note 1: Models

- **GPT-5**: the latest general-purpose LLM developed by OpenAI, widely regarded as one of the strongest models currently available and offering substantially improved reasoning over the GPT-4 series. We evaluated the version `gpt-5` through the official API.
- **Gemini 2.5 Flash**: a fast, cost-efficient general-purpose LLM developed by Google, designed to retain most of the reasoning ability of the larger Gemini 2.5 family at markedly lower latency. We evaluated the version `gemini-2.5-flash` through the official API.
- **DeepSeek-V3.1**: an open-source 671B-parameter mixture-of-experts LLM developed by DeepSeek. We used `deepseek-chat` through the official API, during the period in which this model string served DeepSeek-V3.1.
- **Qwen3.5-Flash**: a lightweight mixture-of-experts LLM (35B total parameters, 3B activated per token) developed by Alibaba, designed for low-latency inference. We used `qwen3.5-flash` through the official Alibaba Cloud Model Studio API.
- **HuatuoGPT-3**: an open-source medical-domain-specific LLM built on a Qwen3 backbone and adapted to medicine through an RL-only domain-adaptation procedure (SeedRL). It is included as a domain-specialised baseline. We used the 32B variant, with weights released on Hugging Face (`FreedomIntelligence/HuatuoGPT-3-32B`), deployed locally.
- **Claude Opus 4.8** (scoring model only): developed by Anthropic and used solely as the third-party scorer for the guideline-compliance evaluation. It is independent of all models under test and generated none of the plans it assessed. We used the version `claude-opus-4-8` through the official API.

Supplementary Note 2: Agent prompts

Prompts are listed in the order in which the agents are executed in the DRAgent (Fig. 2c).

Outpatient physician: clinical-record generation

Supplementary Prompt 1 — Outpatient physician (clinical-record generation)

I will provide you with fundus images of patients. The patient's disease type is provided. Your task is to generate a case report based on the provided information, primarily relying on the fundus image. The fundus image is only for a single eye. The terms "left" or "right" in the image filename indicate which eye is shown; describe only the affected eye. For other information not provided, generate appropriately according to the patient's situation.
- Different images come from different patients. The patient's name is the image filename; different names represent different patients. Therefore, ensure personalized case reports and avoid repetition! Ensure the reports are rich and varied, reflecting the unique situation of each patient.
Follow the fundus case report template below:
1. Basic Information

- **Name**: _____
- **Gender**: _____ (Male/Female)
- **Age**: _____ years

- **Date of Visit**: _____

2. Chief Complaint

- **Chief Complaint**: _____ (eye, symptoms, duration, reason for visit)

3. Present Illness

- **Present Illness**: _____

4. Past History (write "none" if not applicable)

- **Ophthalmic History**: _____
- **Past Ophthalmic Surgery**: _____
- **Other Eye Diseases**: _____ (if any)
- **Systemic Diseases**:
  - **Diabetes**: _____ (if any)
  - **Hypertension**: _____ (if any)
  - **Kidney Disease**: _____ (if any)
  - **Other Diseases**: _____ (if any)
- **Medication Control**: _____ (if any)

5. Ocular Examination (Left Eye/Right Eye)

- **Visual Acuity**: _____ (unaided/corrected)
- **Intraocular Pressure**: _____ mmHg
- **Anterior Segment**: _____
- **Vitreous**: _____
- **Optic Disc**: _____
- **Macula**: _____
- **Retina**: _____ (e.g., hemorrhages, neovascularization)

6. Diagnosis

- **Preliminary Diagnosis**: _____

Ophthalmologist: treatment plan

Supplementary Prompt 2 — Ophthalmologist (initial plan)

You are an ophthalmologist specialising in diabetic retinopathy. Based on the patient's disease type, the case report, the suggestions of the other agents, and the retrieved guideline content, formulate a precise and personalised treatment plan for the patient's current condition.
Follow the template below strictly, and do not repeat content across items:

1. Disease diagnosis

- Briefly state the patient's diagnosis, stage, and relevant medical history.

2. Treatment principles

- For example, controlling disease progression, improving vision, relieving symptoms, etc.

3. Treatment plan

- Select appropriate treatment modalities according to the disease type; these may include (but are not limited to):

  - **Medication** (e.g. anti-VEGF agents; specify the drug class or the medication strategy)
  - **Surgical treatment** (e.g. laser photocoagulation)

4. Follow-up plan

- Schedule regular follow-up and reassessment according to the nature of the disease and the treatment response.

5. Patient education

- Provide the patient with relevant education.

6. Precautions

- Points that require attention.

If the patient currently requires surgery, give a detailed photocoagulation plan under "Surgical treatment", including the photocoagulation type, laser wavelength (colour), spot size, power, exposure duration, treatment staging, avoidance zones, and total number of laser spots. If surgery is not required, write "None".
If any surgical detail conflicts with the opinion of the laser-surgeon agent, defer to the latter.
Output only the treatment plan generated according to the template above, and keep it concise. The plan must address the patient's current condition only; avoid over-treatment.

Laser surgeon

Supplementary Prompt 3 — Laser surgeon

You are a specialist in laser photocoagulation surgery. Based on the patient's case (which already includes the examination findings), the treatment plan proposed by the diabetic-retinopathy ophthalmologist, and the retrieved photocoagulation guideline content, answer the following question.
Do you fully agree with the treatment plan proposed by the ophthalmologist? Focus your review on the anti-VEGF medication, the laser surgery, and the surgical parameters (if any).

- If you **fully agree**, reply with "Yes" only.
- If you **do not fully agree**, reply "No", state your reasons, and give your alternative plan together with the guideline items on which it is based.

The plan must address the patient's current condition only; avoid over-treatment.

Endocrinologist

Supplementary Prompt 4 — Endocrinologist

You are an endocrinologist. Based on the patient's case, the treatment plan proposed by the diabetic-retinopathy ophthalmologist, and the retrieved diabetes-management guideline content, answer the following question with respect to the endocrine-related content only.
Do you fully agree with the endocrine-related part of this plan? State your judgement explicitly, with reference to the patient's specific condition.

- If you **fully agree**, reply "Yes" and give your reasons.
- If you **do not fully agree**, reply "No", state your reasons, and give your alternative plan, including its likely benefits and risks.

Pharmacist

Supplementary Prompt 5 — Pharmacist

You are a clinical pharmacist. Based on the patient's case, the treatment plan proposed by the diabetic-retinopathy ophthalmologist, and the retrieved clinical medication guideline content, answer the following question with respect to the medication-related content only.
Do you fully agree with the medication-related part of this plan? State your judgement explicitly, with reference to the patient's specific condition.

- If you **fully agree**, reply "Yes" and give your reasons.
- If you **do not fully agree**, reply "No", state your reasons, and give your alternative plan, including its likely benefits and risks.

Nurse: plan consolidation

Supplementary Prompt 7 — Nurse (consolidation)

You are a doctor. You need to consolidate the following two treatment plans:

1. the ophthalmic surgical treatment plan;
2. the endocrine treatment plan.

Consolidation requirements:

- Retain all key clinical information, in particular the anti-VEGF injection frequency, the surgical parameters, and the glycaemic targets.
- All ophthalmic treatment parameters must be written out in full. Vague cross-references such as "as given in the ophthalmic plan" are not permitted, and no parameter may be omitted.
- For ophthalmic content, follow the ophthalmic surgical plan; for endocrine content, follow the endocrine plan.
- Ensure the consolidated plan is internally consistent and free of conflicts.
- The output must read as a formal clinical plan; do not use hedging phrases such as "under the guidance of a physician".
- Consolidate the existing information only. Do not add new content, and do not add a case report — consolidate the treatment plans only.
- Organise the result into a clearly structured, comprehensive treatment plan following the original template.

The consolidation agent is assigned a doctor so that it is able to retain all key clinical information in full.

Standalone-LLM baseline

Supplementary Prompt 8 — Standalone LLM baseline

You are an ophthalmologist specialising in diabetic retinopathy. Your task is to formulate a precise and personalised treatment plan for the patient's current condition, based on the disease type and the case report.
If the patient currently requires surgery, design a detailed photocoagulation plan, including the photocoagulation type, laser wavelength (colour), spot size, power, exposure duration, treatment staging, avoidance zones, and total number of laser spots. If surgery is not required, write "None".
Follow the template below, and do not repeat content across items:
1. Disease diagnosis

- Briefly state the patient's diagnosis, stage, and relevant medical history.

2. Treatment principles

- For example, controlling disease progression, improving vision, relieving symptoms, etc.

3. Treatment plan

- Select appropriate treatment modalities according to the disease type; these may include (but are not limited to):
  - **Medication** (e.g. anti-VEGF agents, metabolic management; specify the drug class or the medication strategy)
  - **Surgical treatment** (e.g. laser photocoagulation)

4. Follow-up plan

- Schedule regular follow-up and reassessment according to the nature of the disease and the treatment response.

5. Patient education

- Provide the patient with relevant education.

6. Precautions

- Points that require attention.

Output only the treatment plan generated according to the template above, and keep it concise. The plan must address the patient's current condition only; avoid over-treatment.

Supplementary Note 3: Guideline-compliance assessment

A dynamic triggering mechanism activates only the items relevant to each patient's pathology; the Normalized Compliance Rate (NCR) is the ratio of the achieved score to the maximum score of the triggered items.

Supplementary Note 4: Objective similarity — distillation and weighting

Supplementary Prompt 10 — Plan distillation

Based on the content of the treatment plan provided to you, extract only the key information of each section according to the specified template. Keep the content concise.
The distillation must address the current treatment only and must not include hypothetical scenarios. Follow the template exactly: if an item does not appear, do not omit it — write "None" after it. Remove the parenthesised guidance when generating.
I. Disease diagnosis (brief)

- Disease name:
- Description of the condition:

II. Treatment principles
III. Treatment plan (write "None" if not mentioned)

- Is laser surgery currently required: Yes / No
- Surgery type:
- Laser type:
- Spot size:
- Power:

Supplementary Table 1: Ophthalmic guideline-compliance assessment criteria.

ID	Assessment point	Guideline requirement	Scoring
I. Normativity (existence; 0.5 point each)			
S1	Disease diagnosis present?	Brief diagnosis, staging, and history	0.5
S2	Treatment principles present?	At least one explicit treatment goal	0.5
S3	Treatment plan present?	At least one treatment modality	0.5
S4	Follow-up plan present?	Explicit follow-up frequency/content	0.5
S5	Patient education present?	At least one education item	0.5
S6	Precautions present?	At least one precaution	0.5
S7	Systemic disease control	Mentions glycaemic/blood-pressure control	0.5
II. Correctness (patient-driven; 3 points each)			
C1	Severe NPDR / high-risk PDR	Pan-retinal photocoagulation (PRP)	3
C2	Mild–moderate NPDR without DME	Observation, follow-up every 6–12 months	3
C3	Capillary aneurysm with hard exudates	Focal photocoagulation	3
C4	Retinal non-perfusion, IRMA, and capillary beds with diffuse leakage	Grid photocoagulation	3
C5	DME or neovascularization	Anti-VEGF	3
C6	Vitreous haemorrhage / tractional detachment	Pars plana vitrectomy (PPV)	3
III. Photocoagulation-specific (only when laser is recommended)			
P1	Surgery type specified	Surgery type stated explicitly	3
P2	Wavelength	Green / yellow / red laser	1
P3	Spot size	PRP 200–500 μm; grid/focal 100–200 μm	1
P4	Exposure time	0.2–0.5 s	1
P5	Laser intensity	Light grey-white (2–3 reaction)	3
P6	Treatment staging	Completed in 3–4 sessions	3
P7	Avoidance zone	Macular region explicitly spared	3
P8	Total number of laser spots	PRP: 1,600–2,000	2
IV. Treatment-selection logic (key decision points)			
D1	Cataract with DME	Anti-VEGF/steroid first, then cataract surgery	1
D2	Anti-VEGF regimen	Loading 4-weekly $\times$3, then $\times N$ (3+PRN or T&E)	3
D3	Poor anti-VEGF response	Switch to steroid therapy	3

Supplementary Table 2: Endocrine (diabetes) guideline-compliance assessment criteria.

ID	Assessment point	Guideline requirement	Scoring
I. Target setting (structural)			
A1	Individualized HbA1c target	Explicit target (e.g., <7.0%, relaxed/tightened as appropriate)	1
A2	Comprehensive control targets	Blood pressure (<130/80 mmHg), lipids (LDL-C <2.6/1.8)	1
II. Core-treatment correctness (patient-driven)			
X1	All T2DM without contraindication	Should include metformin	3
X2	T2DM with ASCVD	Prefer GLP-1RA or SGLT2i	3
X3	T2DM with heart failure / CKD	Prefer SGLT2i (HF/CKD) or GLP-1RA (CKD)	3
X4	Insulin initiation indication	Meets any: inadequate oral therapy (HbA1c$\geq$7.0%); hyperglycemic symptoms or HbA1c$\geq$9.0%; unexplained weight loss	3
III. Follow-up and monitoring			
U1	Glucose monitoring	Mentions self-monitoring or HbA1c testing frequency	1
U2	Complication monitoring	Mentions blood-pressure, lipid, and renal-function monitoring	1

- Exposure time:
- Treatment staging:
- Avoidance zone:
- Total number of laser spots:
- Other surgery:

IV. Medication (write "None" if not mentioned)

- Is anti-VEGF currently used: Yes / No
- Anti-VEGF regimen:
- Other drugs:

V. Glycaemic / blood-pressure control
VI. Follow-up plan

- Follow-up frequency:
- Follow-up items:

VII. Patient education
VIII. Precautions

Supplementary Table 3: Weights used to aggregate the per-field objective-similarity scores into a single plan-level score.

Distilled field	Section	Hybrid	Real
Is laser surgery currently required	III. Treatment plan	0.200	0.300
Surgery type	III. Treatment plan	0.100	0.200
Laser type	III. Treatment plan	0.025	—
Spot size	III. Treatment plan	0.025	—
Power	III. Treatment plan	0.025	—
Exposure time	III. Treatment plan	0.025	—
Treatment staging	III. Treatment plan	0.025	—
Avoidance zone	III. Treatment plan	0.025	—
Total number of laser spots	III. Treatment plan	0.025	—
Other surgery	III. Treatment plan	0.100	0.200
Is anti-VEGF currently used	IV. Medication	0.200	0.200
Anti-VEGF regimen	IV. Medication	0.050	—
Other drugs	IV. Medication	0.050	—
Glycaemic / blood-pressure control	V. Systemic control	0.025	0.100
Follow-up frequency	VI. Follow-up	0.050	—
Follow-up items	VI. Follow-up	0.050	—
Total		1.000	1.000

BERTScore-F1 was computed with `bert-base-chinese` as the backbone model.

Supplementary Note 5: Complete numerical results

Figures 1 and 2 of the main text present the comparative and ablation results graphically; this note reports the underlying values.

Supplementary Table 4: Guideline-compliance NCR (%) on the 588-case hybrid real–synthetic evaluation set, by backbone and DR grade. A = DRAgent; L = standalone LLM.

		Ophthalmic NCR (%)					Endocrine NCR (%)				
Model		DR-1	DR-2	DR-3	DR-4	Mean	DR-1	DR-2	DR-3	DR-4	Mean
GPT-5	A	**85.9**	**76.7**	**84.4**	**87.2**	**83.3**	**98.8**	**98.9**	**96.7**	**96.4**	**97.7**
	L	77.1	68.9	33.7	74.5	59.8	73.1	72.2	75.5	64.1	72.3
Gemini 2.5 Flash	A	**80.5**	**73.6**	**82.5**	**80.2**	**79.4**	**86.1**	**87.9**	**90.7**	**94.3**	**89.4**
	L	73.8	67.4	59.6	69.3	66.6	60.2	66.9	66.4	66.7	65.0
DeepSeek-V3.1	A	**81.8**	**71.7**	**79.0**	**77.6**	**77.6**	**85.8**	**90.5**	**94.8**	**97.4**	**91.8**
	L	74.1	55.5	65.1	71.6	66.0	52.3	60.1	53.0	49.3	54.0
Qwen3.5-flash	A	**80.9**	**74.6**	**83.2**	**84.5**	**80.6**	**87.7**	**94.2**	**98.3**	**97.1**	**94.4**
	L	80.2	67.9	43.8	67.5	62.7	50.6	54.6	68.7	60.9	59.4
HuatuoGPT-3†		75.5	65.8	61.4	69.1	67.2	50.6	52.4	48.9	47.2	50.0

Supplementary Table 5: Objective similarity on the 588-case hybrid real–synthetic evaluation set.

	BLEU-1		BLEU-2		BLEU-3		BLEU-4		ROUGE-L		METEOR		BERTScore-F1	
Model	L	A	L	A	L	A	L	A	L	A	L	A	L	A
GPT-5	0.603	**0.700**	0.587	**0.673**	0.578	**0.658**	0.573	**0.649**	0.611	**0.715**	0.282	**0.316**	0.831	**0.879**
Gemini 2.5 Flash	0.662	**0.725**	0.634	**0.699**	0.619	**0.681**	0.608	**0.668**	0.674	**0.744**	0.296	**0.326**	0.866	**0.896**
DeepSeek-V3.1	0.666	**0.689**	0.637	**0.662**	0.622	**0.647**	0.612	**0.637**	0.679	**0.703**	0.299	**0.312**	0.873	**0.880**
Qwen3.5-flash	0.647	**0.712**	0.625	**0.684**	0.614	**0.669**	0.605	**0.655**	0.657	**0.729**	0.296	**0.320**	0.859	**0.888**
HuatuoGPT-3†	0.626		0.603		0.591		0.583		0.634		0.286		0.842	

Supplementary Table 6: Agent-ablation results on the 588-case hybrid real–synthetic evaluation set. Ophthalmic NCR compares the full DRAgent (Final) with the variant lacking the laser surgeon (w/o Laser). Endocrine NCR compares the full system (DM Final) with variants lacking the pharmacist (w/o Pharma) or lacking both the pharmacist and the endocrinologist (w/o Both). All values are case-weighted means (%). These are the values plotted in Fig. 2.

	Ophthalmic NCR (%)		Endocrine NCR (%)		
Model	Final	w/o Laser	DM Final	w/o Pharma	w/o Both
GPT-5	83.3	71.5	97.7	97.6	80.2
Gemini 2.5 Flash	79.4	68.6	89.4	83.1	62.4
DeepSeek-V3.1	77.6	70.0	91.8	89.3	67.4
Qwen3.5-flash	80.6	71.5	94.4	92.3	62.6
Mean	**80.2**	70.4	**93.3**	90.6	68.2