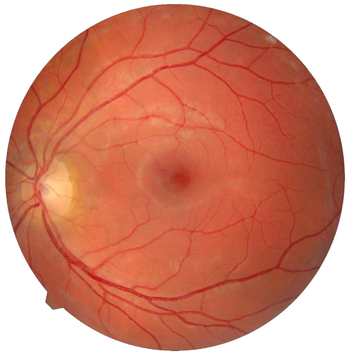
Prediction of Optical Coherence Tomography Retinal Layer Thickness from Colour Fundus Photography in the UK Biobank

James C Porter; Miguel O. Bernabeu; Baljean Dhillon

Abstract
429194

1. Background

- **Retinal thicknesses** such as RNFL are important biomarkers for neurodegenerative disease[1] but require expensive and large **Optical Coherence Tomography (OCT)** machines to image them.
- **Colour Fundus Photography (CFP)** is ubiquitous, cheap, simple and has historic data.
- **We investigate predicting retinal thicknesses using Ridge regression on deep CFP embeddings.** This could enable large-scale screening for neurodegenerative disease like Alzheimer's.



2. Data

We used paired CFP images and OCT-derived mRNFL, mGCIPL and mINL thicknesses from the **UK Biobank**. We retained **102,232 eyes** after quality control filtering CFPs via QuickQual[2] p(bad) ≥ 0.9 for CFP and OCTs via Topcon TABS quality metrics (Q-factor ≤ 45 & worst 20% across other).

Features were extracted **pre-trained RETfound** foundation model[3] embeddings and image statistics (e.g. RGB variance).

3. Methods

We trained separate **ridge regression** (10-CV) models to predict **mRNFL, mGCIPL and mINL** using subsets of the data to investigate their effects on prediction: eye laterality, sex, ethnicity, systemic-health.

`Healthcare system interaction score' (HS) is a simple proxy for systemic-health, with a higher value for worse health. Calculated via: ICD-10 count (D), medication count (M), baseline age (A_0), death age (A_d), prior cancer (C), Age factor (AF) & Death factor (DF).

$$\mathrm{AF} = 1 - \frac{A_0 - \min(A_0)}{\max(A_0) - \min(A_0)} \qquad \mathrm{DF} = 1 - \frac{(A_d - A_0) - \min(A_d - A_0)}{\max(A_d - A_0) - \min(A_d - A_0)}$$

$$\mathrm{HS} = AF(D + 0.5M + C) + DF$$

4. Results

Layer	Comparison	R^2	MAE [μm]	MAEΔ [μm]	p(ttest)	Cohen's d
RNFL	Left / Right	0.231 / 0.303	3.03 / 3.31	-0.28	<**0.01**	-0.06
	Female / Male	0.244 / 0.270	3.35 / 3.11	0.24	<**0.01**	0.06
	White / NonWhite	0.283 / 0.232	3.22 / 3.41	-0.19	0.054	-0.04
GCIPL	Left / Right	0.196 / 0.20	4.22 / 4.14	0.08	0.41	0.01
	Female / Male	0.223 / 0.215	4.14 / 4.29	-0.15	0.09	-0.02
	White / NonWhite	0.203 / 0.236	4.20 / 4.32	-0.12	0.29	-0.02
INL	Left / Right	0.090 / 0.105	1.94 / 1.86	0.09	**0.037**	0.03
	Female / Male	0.089 / 0.079	1.91 / 1.97	-0.06	0.12	-0.02
	White / NonWhite	0.085 / 0.083	1.95 / 1.86	0.10	**0.048**	0.03

All models had positive R^2 (0.08–0.30) and had MAE 1.9-4.3µm. This meant that the prediction errors (within 0.02-0.3µm) were **comparable to the natural variability in OCT-measured values**.

- **GCIPL** predictions were robust across all strata (no significant ΔMAE)
- **RNFL** predictions were best performing and **INL** the worst.
- **Sex** and **eye laterality** had a significant effect on RNFL predictions but only small effect sizes.
- All **effect sizes** were small (Cohen's d <0.1), therefore models were robust to strata and HS.

Layer	Subset	MAEΔ[μm]	p (t-test)	Cohen's d
GCIPL (left)	HS50	-0.02	0.41	-0.03
	HS60	-0.01	0.57	-0.02
	HS70	0.01	0.69	0.01
	HS80	0.00	0.95	0.00
GCIPL (right)	HS50	-0.03	0.35	-0.03
	HS60	-0.02	0.48	-0.02
	HS70	**-0.04**	**0.026**	-0.06
	HS80	0.02	0.32	0.03
INL (left)	HS50	0.00	0.97	0.00
	HS60	0.01	0.50	0.02
	HS70	0.00	0.82	0.01
	HS80	-0.01	0.25	-0.03
INL (right)	HS50	**-0.03**	**0.015**	-0.08
	HS60	-0.01	0.10	-0.05
	HS70	-0.01	0.24	-0.03
	HS80	0.01	0.29	0.03
RNFL (left)	HS50	**-0.10**	<**0.001**	-0.13
	HS60	-0.03	0.10	-0.05
	HS70	0.00	0.97	0.00
	HS80	-0.02	0.07	-0.04
RNFL (right)	HS50	**-0.06**	**0.015**	-0.08
	HS60	-0.01	0.51	-0.02
	HS70	-0.03	0.053	-0.05
	HS80	0.01	0.63	0.01

HS filtering did not significantly affect performance. Even at the strictest HS filter, removing the top 50% of highest scores only improved results by 0.03–0.10 µm MAE. Therefore, models were robust to variations in systemic health.

5. Conclusions

Limitations: single-centre (UKB); linear model may overlook nonlinear features; HS is crude.

Future: integrate deep neural networks; validate on non-UKB data; explore longitudinal CFP changes; extend to other retinal layers; improve HS to be more nuanced.

Take-Home Message:
Ridge regression on CFP achieves MAE comparable to OCT variability for mRNFL, mGCIPL and mINL, and is robust to eye laterality, sex, ethnicity and a systemic-health proxy.

THE UNIVERSITY of EDINBURGH | Usher institute | Precision Medicine
james.c.porter@ed.ac.uk

References:
1. Van der Heide et al. Associations of inner retinal layers. Alzheimer's & Dementia 2024
2. Engelmann J. et al. QuickQual, Ophthalmic Med Image Analysis 2023.
3. Zhou Y. et al. RETfound, Nature 622, 156 (2023).

Acknowledgments: Data from UK Biobank (app #91570).
Contact: James C. Porter, james.c.porter@ed.ac.uk