

Can Artificial Intelligence Platforms Diagnose Eye Disease?

A comparison between Chat GPT-4o mini and super-prompt agent Alan

Joanna Louis, Andrew Blaikie

School of Medicine, University of St Andrews, St Andrews, United Kingdom



University of
St Andrews

Introduction

Eye disease encompasses a range of conditions that can **affect** individuals' **quality of life** (1). **Artificial intelligence** (AI) platforms could provide **solutions** to address the eye care needs of **populations** that are **inadequately provisioned** (2). This study's objective was to **compare** two **AI platforms** (Alan vs Chat GPT-4o mini) on their outputs in **response** to clinical **eye disease cases**.

Aims

Compare each platform based on:

- 1.Differential diagnoses**
- 2.Referral urgency**
- 3.Language used**

When prompted with eye disease cases based on the World Health Organizations (WHO) Primary Eye Care Manual (PEC).

Method

Clinical **eye disease cases** made from the **WHO PEC manual** domains:

↓

Loss of vision

↓

Paediatrics

↓

Red eye

↓

Trauma

↓

Lumps and lashes

Section name - Case number

Age: x Sex: F/M

3 Presenting complaints:
A B C

3 variations of the case

↓

AB

↓

BC

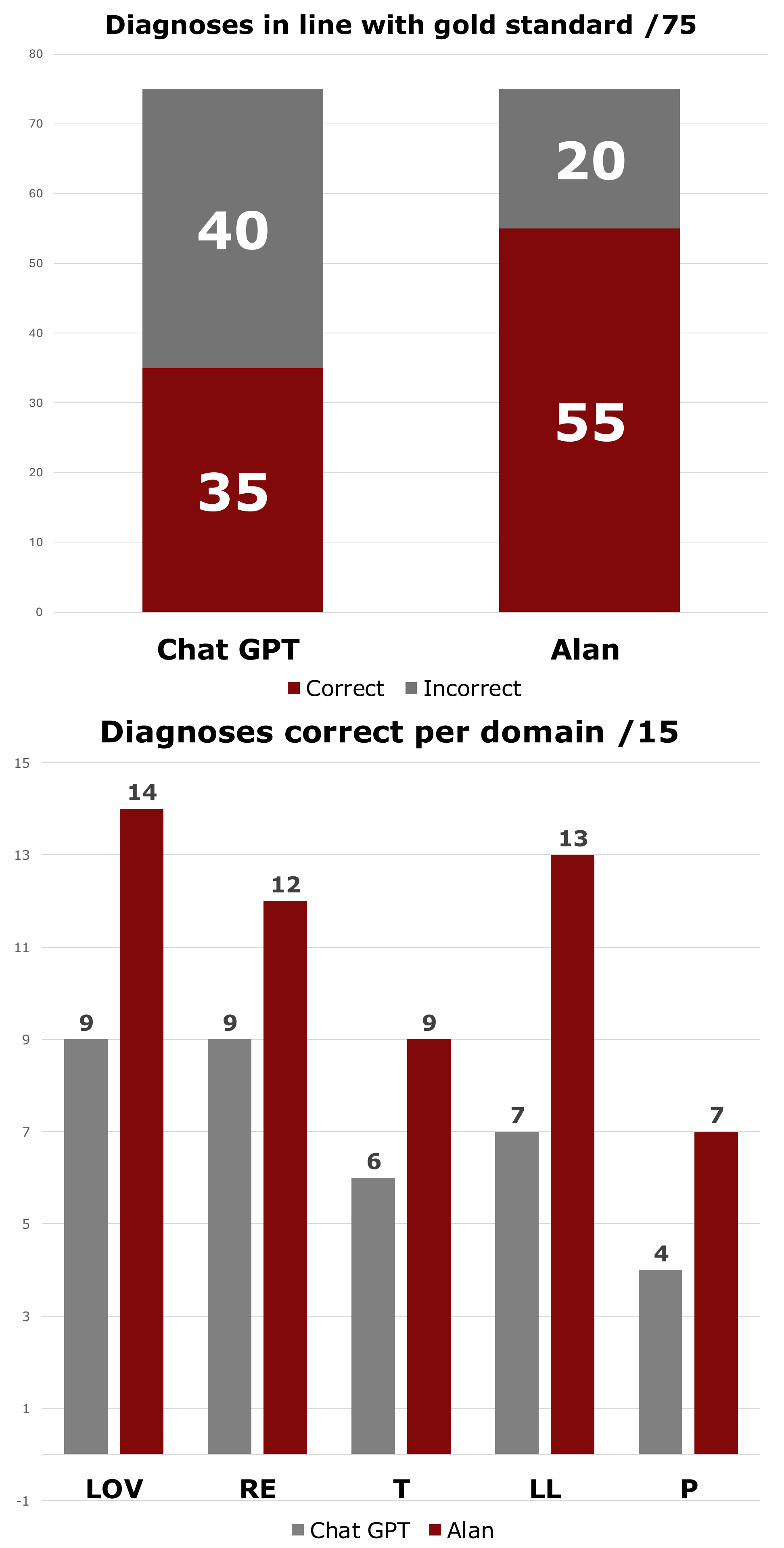
↓

AC

Analysis was performed based on the platforms suggested diagnosis, management, and referral outcomes in **comparison to gold standard answers** from a consultant ophthalmologist.

Additionally, **readability** of the transcripts was determined by performing a **Flesch-Kincaid (FK) analysis**; an online readability calculator.

Results



FK analysis section	Alan average	Alan median	ChatGPT average	ChatGPT median
Grade level	6.3	6.6	10.8	11
Reading ease	59.9	57.4	35.4	35.3
Words per sentence	6.1	6	10.2	10.1
Syllables per word	1.7	1.7	1.9	1.9
Sentences	36.8	37	140.7	127
Total words	222.7	223	1423.3	1332
Messages total	5.5	6	3.1	3
Messages till diagnosis	4.8	5	2.7	3

- Significant differences between** the diagnostic accuracies of the **platforms** were indicated through statistical testing.
- Referral suggestions** from the correctly diagnosed cases were correct in **39/55 for Alan** and **30/35 for ChatGPT-4o mini**, with **no statistical significance** found between referral suggestions.
- Paired t-tests showed **highly statistically significant linguistic differences between platforms**, most notably between the lower average reading ease and word count per transcript from Alan in comparison to ChatGPT-4o mini.

Conclusion

This study aims to build on existing research into the abilities of AI as a diagnostic tool and has shown more favourable outcomes to tailored platforms. Integration of specialised technologies into primary eye care and teleophthalmology could reduce the global burden of eye disease.

References

- World Health Organization (2023) *Blindness and visual impairment*. Available at: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- Ting DSJ, Foo VH, Yang LWY, et al Artificial intelligence for anterior segment diseases: Emerging applications in ophthalmology *British Journal of Ophthalmology* 2021;105:158-168.