

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort profile: AlzEye: longitudinal record-level linkage of ophthalmic imaging and hospital admissions of 353,157 patients in London, United Kingdom
AUTHORS	Wagner, Siegfried; Hughes, Fintan; Cortina Borja, Mario; Pontikos, Nikolas; Struyven, Robbert; Liu, Xiaoxuan; Montgomery, Hugh; Alexander, Daniel; Topol, Eric; Petersen, Steffen; Balaskas, Konstantinos; Hindley, Jack; Petzold, Axel; Rahi, Jugnoo; Denniston, Alastair; Keane, Pearse

VERSION 1 – REVIEW

REVIEWER	Kim, Jee Chung Ang University Hospital
REVIEW RETURNED	24-Nov-2021

GENERAL COMMENTS	<p>Cohort profile: AlzEye: longitudinal record-level linkage of ophthalmic imaging and hospital admissions of 353,157 patients in London, United Kingdom</p> <p>The authors have set the large retinal imaging database composed of 353,157 participants from Moorfields Eye Hospital NHS Foundation. It looks like great works. However, the author should address following issues.</p> <ol style="list-style-type: none">1.Representative images of eyes with cataract, glaucoma, AMD, and PDR, and systemic disease of dementia or cardiovascular disease would provide help the reader for understanding the dataset2.The manuscript should be shorten throughout the manuscript.
-------------------------	---

REVIEWER	Silverstein, Steven M University of Rochester Medical Center, Department of Psychiatry
REVIEW RETURNED	31-Jan-2022

GENERAL COMMENTS	<p>Wagner et al. describe an important and innovative initiative, AlzEye, that involves combining retinal image data collected at one site (with the world's largest retinal image dataset) with hospital record data on systemic diseases collected at multiple other sites. This has led to a data set of millions of images coming from thousands of patients. The resulting dataset is unique in several respects and will be able to answer pressing questions related to ocular biosignatures of systemic disease, especially (in the first phase of the study) in cardiovascular and neurodegenerative diseases. This type of integration of datasets is the first of its kind, and will likely serve as a model for future research efforts. As the authors note, while several large datasets are available to</p>
-------------------------	---

	<p>researchers, they tend to include data on healthy volunteers and so are not optimal for studying and predicting the development and course of disease states. In addition, an advantage of the AlzEye database is that contains multiple images, over time, from patients, thus allowing for important longitudinal analyses, including those focusing on prediction of disease course and/or treatment response.</p> <p>This is a well-written paper that is easy to follow. Nevertheless, it does an excellent job of describing all of the steps and technical complexities involved in setting up the database in a way that satisfies multiple regulatory requirements. The use of a third-party linkage source is an interesting solution to several of the issues. In addition, seeking feedback from the general public and including non-scientists as part of the working group is an important aspect of ensuring relevance to public health. From a statistical point of view, calculating a power analysis to ensure in advance that the project will be able to answer specific questions is another strength. In addition to describing the current project, the paper can serve as a “How-to” document that will be helpful for anyone hoping to do a structurally similar database linkage project, regardless of the nature of the data or the diseases studied. In short, the paper describes everything from the initial concept, through all the phases of development, to the final dataset, to future directions.</p> <p>The project described should generate much valuable research, but it is also clinically relevant. For example, as noted in the paper, people tend to seek out eye care when vision worsens, and thus the data from this study will be able to inform physicians and other health care providers about risk for disabling systemic diseases from routine eye examinations. This could lead to disease prevention and attenuation efforts that could improve the lives of many people who would not otherwise have sought out care for these diseases until only tertiary care would be possible.</p> <p>Finally, limitations and sources of bias in the data are described.</p> <p>This is a very useful paper that describes an exciting project that should influence big data studies in the emerging field of oculosomics, as well as in other areas of research.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1. The authors have set the large retinal imaging database composed of 353,157 participants from Moorfields Eye Hospital NHS Foundation. It looks like great works.

Authors' response: We thank the Reviewer for their kind words.

2. Representative images of eyes with cataract, glaucoma, AMD, and PDR, and systemic disease of dementia or cardiovascular disease would provide help the reader for understanding the dataset

Authors' response: We thank the Reviewer for this suggestion. We have indeed now created a new Figure, Figure 7, which is a composite of images pertaining to colour retinal photographs of individuals within the dataset with cataract, glaucoma, age-related macular degeneration, proliferative diabetic retinopathy, prevalent Alzheimer's disease, prevalent vascular dementia, incident ischaemic stroke and incident myocardial infarction.

3. The manuscript should be shorten throughout the manuscript.

Authors' response: We appreciate the Reviewer's comment about shortening the manuscript. Indeed, we sought to balance providing sufficient information for readers to conduct their own investigator-led linkage and details of the cohort while maintaining readability. In line with the Reviewer's suggestion, we have reduced the count by 500 words. All removals are visible through tracked changes.

Reviewer: 2

Authors' response: We thank Reviewer 2 for their positive feedback and for highlighting the aims, strengths and potential impact of the project.