PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prevalence, Risk Factors and Associated Ocular Diseases of Cerebral Stroke: The Population-based Beijing Eye Study
AUTHORS	wang, yaxing; Wei, Wenbin; Xu, L; Jonas, Jost

VERSION 1 - REVIEW

REVIEWER	Miki Uchino
	Keio University School of Medicine Japan
REVIEW RETURNED	01-Aug-2018

GENERAL COMMENTS	Cerebral Stroke: Its Prevalence, Risk Factors and Associated Ocular Diseases. The Beijing Eye Study.
	This study is very interesting and well written.
	However, the multivariate analysis needs to be re-done since it includes the parameters with collinearity. Please remove them and do re-analysis. This will change the result.
	In the result section, the multivariate analysis steps are not essential to state. I consider all the steps need to be deleted from the MS.

REVIEWER	Dr.Durairaj Sekar
	Dr.Durairaj Sekar Ph.D Department of Life Sciences Center for Research and PG studies Kristu Jayanti College (Autonomous) K Narayanapura, Kothanur P.O Bangalore - 560077, India. Phone: +91 7411464045
REVIEW RETURNED	08-Oct-2018

GENERAL COMMENTS	Dear Author
	1. Statistical validation is required and needs anyone validated biomarker for the conclusion for cerebral strokes
	2. the manuscript require native English speaker to rewrite it
	3. conclusions has to be well written
	thanks

REVIEWER	Xia Wang
	The George Institute for Global Health
REVIEW RETURNED	03-Jan-2019

GENERAL COMMENTS	The authors investigated the prevalence of stroke and its association with traditional risk factors and ocular disease. They found that diabetic retinopathy is of prognostic significance in addition to the traditional risk factors. The paper deals with an interesting topic, but there are some issues that should be paid attention to.
	1. Stroke is a dramatic event and therefore it is less likely to be under reported. However, transient ischemic attack was also included, most of which might be even not noticed by the patients. So it should be under reported.
	2. There should a rule stated to perform the statistical analysis with which model includes what variables. Results do no match with methods, please state clearly.
	3. 'As a first step, we examined the mean value of the main outcome parameter, i.e. the prevalence of stroke (presented as mean and 95% confidence intervals (CI))' the prevalence of stroke is presented as a proportion, what does 'mean' mean?
	4. This is a cross-sectional observational study, reverse causality may exist. Therefore, diabetic retinopathy might be the consequence of stroke. Please add this point into limitation.
	5. Diabetic retinopathy should only occur among patients with diabetes, if the two factors were put into the multivariable model together, correlation and collinearity should be explored.
	6. Terms need to be used correctly e.g. 'by a factor of', this should mean odds ratio.
	7. Writing style needs to be improved: e.g. in the introduction: 'YLLs' should be 'YLDs'; 'the population of China' is better to be phrased as Chinese population; 'the relationship between systematic factors and ocular disease', however the study is investigating their association with stroke.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1; Reviewer Name: Miki Uchino

6. Reviewer #1: This study is very interesting and well written. However, the multivariate analysis needs to be re-done since it includes the parameters with collinearity. Please remove them and do re-analysis. This will change the result.

Our response: As also pointed out in the reply to comment # 14 made by reviewer #3, the statistical analysis has been repeated for the revision of the manuscript and we have dropped all independent parameters which were collinear with other independent parameters: "As a third step, we conducted an extended multivariable binary analysis which included as independent parameters all those variables which were correlated (P<0.10) with stroke prevalence in the previous analysis. We then dropped step-by-step all those parameters which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of previous stroke. We first started with the systemic independent parameters, such as age and blood pressure. We calculated the odds ratio (OR) and its 95% CIs. All P-values were two-sided and

considered statistically significant, if the values were less than 0.05." (Page 7, line 177); and: "The multivariable analysis included the prevalence of previous stroke as dependent variable and as independent variables all those systemic parameters for which the P-value in the previous analysis was <0.10 (Table 2). We then dropped in step-by-step manner all independent parameters which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of a previous stroke. In the final model, a higher prevalence of previous stroke was correlated (Nagelkerke R2: 0.20) with older age (P<0.001), male gender (P<0.001), lower quality of life score (P<0.001), higher prevalence of arterial hypertension (P<0.001) and cardiovascular disease (P<0.001), and higher prevalence of diabetic retinopathy (P<0.001) (Table 3). If the parameter of prevalence of diabetic retinopathy was replaced by the diabetic retinopathy stage, the latter was associated with previous stroke (P<0.001; OR: 1.64; 95%CI: 1.26, 2.14)." (Page 9, line 224)

7. Reviewer #1: In the result section, the multivariate analysis steps are not essential to state. I consider all the steps need to be deleted from the MS.

Our response: In the revised manuscript, the details of the various steps of the multivariable analysis have been dropped from the manuscript.

Reviewer: 2; Reviewer Name: Dr.Durairaj Sekar

8. Reviewer #2: 1. Statistical validation is required and needs anyone validated biomarker for the conclusion for cerebral strokes

Our response: As also pointed out in the reply to comment # 14 made by reviewer #3, the statistical analysis has been repeated for the revision of the manuscript and we have dropped all independent parameters which were collinear with other independent parameters. A validated biomarker for previous cerebral strokes was unfortunately not available, due to the population-based character of the study in which a cerebral imaging method could not be applied due to a lack of practical feasibility.

9. Reviewer #2: 2. the manuscript require native English speaker to rewrite it 3. conclusions has to be well written

Our response: The whole manuscript including the Conclusions has been re-edited to improve its English in style and grammar.

Reviewer: 3 Reviewer Name: Xia Wang

10. Reviewer #3: The authors investigated the prevalence of stroke and its association with traditional risk factors and ocular disease. They found that diabetic retinopathy is of prognostic significance in addition to the traditional risk factors. The paper deals with an interesting topic, but there are some issues that should be paid attention to.

1. Stroke is a dramatic event and therefore it is less likely to be under reported. However, transient ischemic attack was also included, most of which might be even not noticed by the patients. So it should be under reported.

Our response: It has been added to the revised Discussion: "When discussing the results of our study, its limitations have to be taken into account. First, the data on the prevalence of a previous self-reported stroke depended on the information provided by the study participants in the face-to-face interviews. Since stroke is a dramatic event, it is unlikely to be under-reported. Transient ischemic attacks might occur unnoticed by the individuals so that transient ischemic attacks might be under-

reported in an interview of previous cerebral strokes. Our study was based however primarily on previous cerebral strokes which were defined as an occurrence of typical neurological symptoms for at least 24 hours. It may make it unlikely that unnoticed previous transient ischemic attacks might have markedly influenced the results of our study." (Page 11, line 284)

11. Reviewer #3: 2. There should a rule stated to perform the statistical analysis with which model includes what variables. Results do no match with methods, please state clearly.

Our response: In the revised manuscript, the method of the statistical analysis has been described in greater detail: "The statistical analysis was performed using a commercially available statistical software package (SPSS for Windows, version 25.0, IBM-SPSS, Chicago, IL, USA). As a first step, we assessed the prevalence of previous stroke (expressed as binary parameter as a proportion and the 95% confidence interval (CI)) and calculated the mean values of linear parameters such as ocular axial length (expressed mean ± standard deviation). We then assessed differences between the stroke group and the non-stroke group in age and gender. As second step, we performed a binary regression analysis with the prevalence of stroke as dependent parameter and with other measured parameters as independent variables, after adjusting for age and gender. As a third step, we conducted an extended multivariable binary analysis which included as independent parameters all those variables which were correlated (P<0.10) with stroke prevalence in the previous analysis. We then dropped step-by-step all those parameters which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of previous stroke. We first started with the systemic independent parameters, such as age and blood pressure. We calculated the odds ratio (OR) and its 95% CIs. All P-values were two-sided and considered statistically significant, if the values were less than 0.05." (Page 7, line 169)

12. Reviewer #3: 3. As a first step, we examined the mean value of the main outcome parameter, i.e. the prevalence of stroke (presented as mean and 95% confidence intervals (CI))' the prevalence of stroke is presented as a proportion, what does 'mean' mean?

Our response: As also pointed out in the reply to the previous comment, it has been stated in the revised Methods section that "The statistical analysis was performed using a commercially available statistical software package (SPSS for Windows, version 25.0, IBM-SPSS, Chicago, IL, USA). As a first step, we assessed the prevalence of previous stroke (expressed as binary parameter as a proportion and the 95% confidence interval (CI)) and calculated the mean values of linear parameters such as ocular axial length (expressed mean ± standard deviation)." (Page 7, line 169)

13. Reviewer #3: 4. This is a cross-sectional observational study, reverse causality may exist. Therefore, diabetic retinopathy might be the consequence of stroke. Please add this point into limitation.

Our response: It has been added to the revised Discussion: "When discussing the results of our study, its limitations have to be taken into account. First, ... Sixth, our investigation was a cross-sectional observational study, so that a reverse causality may have existed in the sense that diabetic retinopathy might have been the sequel of stroke." (Page 11, line 301)

14. Reviewer #3: 5. Diabetic retinopathy should only occur among patients with diabetes, if the two factors were put into the multivariable model together, correlation and collinearity should be explored.

Our response: The authors fully agree with the reviewer that there is a strong collinearity between the presence and duration of diabetes mellitus and the presence and stage of diabetic retinopathy. Accordingly, the analysis with adjusting for age and gender showed that a higher prevalence of previous stroke was associated with blood glucose concentration (P=0.01), glycosylated hemoglobin value (P=0.008), with the prevalence of diabetes mellitus (P<0.001), and prevalence (P<0.001) and score (P<0.001) of diabetic retinopathy. For the revision of the manuscript, we have repeated the

statistical analysis and dropped all independent parameters which were collinear with other independent parameters: "As a third step, we conducted an extended multivariable binary analysis which included as independent parameters all those variables which were correlated (P<0.10) with stroke prevalence in the previous analysis. We then dropped step-by-step all those parameters which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of previous stroke." (Page 7, line 177); and: "The multivariable analysis included the prevalence of previous stroke as dependent variable and as independent variables all those systemic parameters for which the P-value in the previous analysis was <0.10 (Table 2). We then dropped in step-by-step manner all independent parameters which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of a previous stroke. In the final model, a higher prevalence of previous stroke was correlated (Nagelkerke R2: 0.20) with older age (P<0.001), male gender (P<0.001), lower quality of life score (P<0.001), higher prevalence of arterial hypertension (P<0.001) and cardiovascular disease (P<0.001), and higher prevalence of diabetic retinopathy (P<0.001) (Table 3). If the parameter of prevalence of diabetic retinopathy was replaced by the diabetic retinopathy stage, the latter was associated with previous stroke (P<0.001; OR: 1.64; 95%CI: 1.26, 2.14)." (Page 9, line 224)

15. Reviewer #3: 6. Terms need to be used correctly e.g. 'by a factor of', this should mean odds ratio.

Our response: Throughout the manuscript, the term "by a factor of" has been replaced by the term "by an odds ratio of".

16. Reviewer #3: 7. Writing style needs to be improved: e.g. in the introduction: 'YLLs' should be 'YLDs'; 'the population of China' is better to be phrased as Chinese population; 'the relationship between systematic factors and ocular disease', however the study is investigating their association with stroke.

Our response:

- With respect to "YLLs", the sentence has been re-formulated to: "In particular China has witnessed a marked increase in the importance of stroke in the spectrum of diseases causing DALYS and years of life lost (YLLs). While in 1990, lower respiratory infections or preterm birth complications were the leading causes of YLLs in almost half of the provinces of China (16 out of 33), cerebrovascular disease were the leading cause in 27 of the 33 provinces in 2013.5,6" (Page 4, line 74)

- With full respect for the reviewer, the authors would like to suggest keeping the term "population of China", since the term "Chinese population" may mean any population of Chinese ethnicity.

- The last sentence of the Introduction has been re-worded: "Since comprehensive population-based studies on associations between stroke and ocular parameters have been scare so far and have not been conducted for the population of China, we investigated the prevalence of cerebral stroke and its potential associations with ocular diseases, after adjusting for systemic factors, in a population-based study performed in China." (Page 4, line 83)

VERSION 2 – REVIEW

REVIEWER	Durairaj sekar
	Saveetha Institutes of Medical and Technical Sciences Chennai India
REVIEW RETURNED	22-Jan-2019

GENERAL COMMENTS	1. statistical review is required

REVIEWER	Xia Wang Australia
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	1. The manscript is not well written and needs to be polished by a native English speaker 2. Comments #5 was not answered 3. This is an example to show how to improve the manuscrip using different professional wording in different sections. For eample, for comments #2,statistical language would be better, if this paragraph is rewritten like this, it should be better: Date were shown as mean (standard deviation), frequency (%, 95% confidence interval), or median (interquartile range) where appropriate. The difference of age and sex between stroke and non-stroke patients was assessed by xxxxx(whatever the method used). We tested associations between baseline characteristics and stroke with logistic regression adjusting for age and sex. Significant covariates from the step above (P <0.1) were included in multivariable models. We reduced the full model by successively removing non-significant covariates until all remaining predictors remained statistically significant (P <0.05).

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2: (Reviewer Name: Durairaj sekar; Institution and Country: Saveetha Institutes of Medical and Technical Sciences, Chennai, India; Please state any competing interests or state 'None declared': No competing Interested); Please leave your comments for the authors below:

1. Reviewer #2: 1. statistical review is required

Our response: The whole statistical analysis has been re-checked for flaws and limitations. The statistical method as such as been described more precisely in the re-revised manuscript as described in detail below (please see reply to comment #4 made by reviewer #3. If the reviewer still has specific points of the statistical analysis to be looked at, the authors should appreciate very much if the reviewer could briefly point them out.

Reviewer: 3 (Reviewer Name: Xia Wang; Institution and Country: Australia; Please state any competing interests or state 'None declared': None declared); Please leave your comments for the authors below

2. Reviewer #3: 1. The manuscript is not well written and needs to be polished by a native English speaker

Our response: The whole manuscript has been re-edited to improve its English in style and grammar.

3. Reviewer #3: 2. Comments #5 was not answered

Our response: Comment #5 of the previous review written by Prof. Xia Wang was: "5. Diabetic retinopathy should only occur among patients with diabetes, if the two factors were put into the multivariable model together, correlation and collinearity should be explored."

Our response: In full agreement with the reviewer and as pointed out more clearly in the re-revised manuscript, the parameter of the presence of diabetes was dropped in the course of the multivariable analysis due to its collinearity with diabetic retinopathy. It has been stated now that "We then dropped in step-by-step manner all independent parameters (such as the prevalence of diabetes mellitus) which either showed a collinearity with one of the other independent variables or which were no longer statistically significantly correlated with the prevalence of a previous stroke." (Page 9, line 223)

4. Reviewer #3: 3. This is an example to show how to improve the manuscript using different professional wording in different sections. For example, for comments #2,statistical language would be better, if this paragraph is rewritten like this, it should be better: Date were shown as mean (standard deviation), frequency (%, 95% confidence interval), or median (interquartile range) where appropriate. The difference of age and sex between stroke and non-stroke patients was assessed by xxxxx(whatever the method used). We tested associations between baseline characteristics and stroke with logistic regression adjusting for age and sex. Significant covariates from the step above (P <0.1) were included in multivariable models. We reduced the full model by successively removing non-significant covariates until all remaining predictors remained statistically significant (P <0.05).

Our response: Thanking the reviewer very much for the formulations, we have amended the description of the statistical methods as recommended: "The statistical analysis was performed using a commercially available statistical software package (SPSS for Windows, version 25.0, IBM-SPSS, Chicago, IL, USA). Date were shown as mean (standard deviation), frequency (%, 95% confidence interval [CI]), or median (interquartile range) where appropriate. The differences in parameters such as age and sex between participants with stroke and participants without stroke were assessed by the student t-test for unpaired samples or by the chi-square test. We tested associations between baseline characteristics and stroke with logistic regression adjusting for age and sex. Significant covariates from the step above (P <0.10) were included in multivariable models. We reduced the full model by successively removing non-significant covariates until all remaining predictors remained statistically significant (P <0.05). We calculated the odds ratio (OR). All P-values were two-sided and considered statistically significant, if the values were less than 0.05." (Page 7, line 170)

VERSION 3 - REVIEW

REVIEWER	Xia Wang The George Institute for Global Health
REVIEW RETURNED	07-Mar-2019

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.