Peer Review File

Article Information:

Reviewer A

It is surprising that the link between apolipoprotein A and hypertension has not been explored yet. Nevertheless, the authors have presented their hypothesis articulately. The study encompasses a substantial sample size of patients and control groups.

Reply: Thank you for your valuable feedback on our paper. We appreciate your recognition of the clarity with which we presented our hypothesis. We acknowledge that while many factors have been established as closely associated with the development of hypertension, the precise relationship between the rs662799 polymorphism of apolipoprotein A5 gene and hypertension remains unclear.This is the original intention of our experiment.

Reviewer B

This manuscript summarizes the potential correlation between APOA5 (rs662799) and EHT. Overall, the topic is fascinating, and the article is well-written and clearly structured. What may be probably improved are some details to further enhance the readability and repeatability of this study. The following are some suggestions:

Reply: Thank you for your valuable feedback. We have revised the manuscript to address your suggestions. We appreciate your guidance and believe that these revisions have improved the clarity and comprehensibility of the manuscript. To provide a more comprehensive description of the experiment, in addition to addressing the suggestions below, we have included the ROC curves and calibration curves for both the training and validation sets.

1. Abstract:

(1) Background:

It would be beneficial to briefly describe the relationship between APOA5 and rs66662799 for clarity.

Besides, please provide the aim of this study in the Background section, you could report "The aim of this study is to ..." (just for your reference).

Reply: In the revised version, we have included a brief description of the relationship between APOA5 and rs662799. Additionally, we have

provided the aim of this study in this section. (Line 30-32, 33-36)

(2) Methods:
a. Please kindly present the study's location and time span.
Reply: We have made the necessary revisions to include the location and time span in the manuscript. (Line 37-38)

b. The authors need to report the parameters and methods of measurement. For instance, "The serum concentrations of various lipid parameters including high-density lipoprotein cholesterol (HDL-C) ... were measured based on coulometric methods", "The genotyping of the SNP was carried out using the improved multiplex ligation detection reaction (iMLDR) method", etc.
Reply: We have made the necessary revisions to include a comprehensive description of the measurement methods in the revised manuscript. The detailed parameters and measurement methods are described in the Methods section. (Line 39-47)

c. "receiver operating characteristic (ROC) curve and calibration curve analyses were conducted to determine the clinical utility of the nomogram model", please be careful about using the word "clinical utility". AUC and calibration curve are statistical properties characterizing the performance of a nomogram model. And decision curve analysis is usually the utility measure to offer insight to the clinical consequences of using a prediction model. Therefore, please revise the "clinical utility".
Reply: We have made the necessary revisions to ensure the accurate representation of the statistical properties characterizing the performance and calibration of the nomogram model. (Line 46-47)

(3) Results:
The results of the Hosmer-Lemeshow test should be reported in the Abstract-Results.
Reply: We have made the necessary revisions to include the results of the Hosmer-Lemeshow test in the Abstract-Results section. (Line 57-59)

2. Methods:
(1) Line 82: "The study received ethical approval from..." Is the location of this study also here? Please clearly state it in the text.
Reply: We have made the necessary revisions. (Line 102)

(2) Please present the time span of this study.

Reply: We have made the necessary revisions. (Line 105)

(3) "Our study included 1400 participants" How was the data extracted? Did you have a trained professional team? Were automated tools utilized? Besides, were these patients sourced from outpatient clinics, case banks, or physical examination centers? More importantly, what were the sources of the control group and how were they selected? Did you make a match? What were the principles of matching?

Reply: Our team have all taken professional training and possess relevant qualifications. The data was obtained from the hospital's data system. The study firstly includes 954 patients diagnosed with EHT. The data were sourced from patients admitted to the First Affiliated Hospital of Xinjiang Medical University between Jan 2019 and Dec 2021. Detailed inclusion and exclusion criteria have been described in section 2.1. After applying the exclusion criteria, 700 patients with EHT remained and were selected as cases. Then a random sampling method was used to select 700 individuals who met the inclusion criteria to serve as the control group. In the revised version, the detailed process of data sources and data selection has been represented through a flow diagram (Figure 1). For the control group, we employed a non-matched control design. (Section 2.1, Figure 1)

(4) Lines 87-88: "as confirmed by at least 3 separate measurements taken on different days without the use of antihypertensive medication". Was this measured by a doctor? What are their qualifications like? How did you choose "different days"? Please kindly describe in detail for clarity.

Reply: We have made the necessary revisions to include these additional details, providing a more comprehensive explanation of the blood pressure measurement process in our study. The multiple blood pressure measurements of a patient are conducted by the same doctor. (Line 106-110, 112-120)

(5) When was the "Biochemical analysis" done? Did you only do it once?

Reply: All biochemical measurements were performed at the Biochemistry Laboratory of the First Affiliated Hospital of Xinjiang Medical University. To ensure data accuracy and minimize the influence of other factors, we use the analysis results of the first blood sample collected during the participants' hospital admission examination during Jan 2019 and Dec 2021. We have made the necessary revisions to the manuscript to explicitly state that the biochemical analysis time. This clarification will enhance the transparency and accuracy of our research findings. (Line

134-146)

(6) We suggest the authors provide the number of missing data, and perform corresponding data processing when necessary.
Reply: In the revised version, the detailed process of data sources and data selection has been represented through a flow diagram. (Figure 1)

(7) Lines 85-88: "EHT in our study was defined as a systolic blood pressure (SBP) ≥140 mmHg ... antihypertensive medication."
Lines 96-98: "Individuals with fasting plasma glucose (FPG) levels ≥7.0 mmol/L...type 2 diabetes mellitus (T2DM)."
Cite references to back up these evaluation criteria.
Reply: These two standards are sourced from the "Chinese Hypertension Prevention and Treatment Guidelines (2023)" and the "Chinese Type 2 Diabetes Prevention and Treatment Guidelines (2020)". We have made the requested modifications and the references have been included in the manuscript. (Line 109, 126)

(8) Lines 98-100: "All the participants completed a health and lifestyle questionnaire, which included an item on smoking and alcohol intake."
Would the authors submit the detailed information of the questionnaire as a supplement material?
Reply: Due to the electronic nature of the questionnaire survey conducted under the guidance of medical professionals, it is difficult to provide the original questionnaire materials. However, we can provide details in English regarding the questionnaires relevant to our study in this manuscript:

(1) Education:
A. Primary school or below
B. Middle school
C. High school
D. Bachelor's degree
E. Master's degree
F. Doctorate degree

(2) Alcohol Consumption:

A. Never
B. Occasionally (less than once a month)
C. Moderate (1-2 times a month)
D. Regular (1-2 times a week)
E. Heavy (more than 3 times a week)

(3) Smoking Status:
Non-smoker
Former smoker
Current smoker (number of cigarettes per day: _____)

(4) Occupation: _____.

(5) Physical Activity Level:
A. Sedentary (little to no physical activity)
B. Light (occasional walking or light exercise)
C. Moderate (regular exercise or sports 1-2 times a week)
D. Active (regular intense exercise or sports 3-5 times a week)
E. Very active (intense exercise or sports every day)

(9) Lines 135-137: "Our results showed that the APOA5 rs662799 G allele, age, BMI, smoking, diabetes, education, LDL-C, and TG were independent risk factors. Based on this result, we developed a risk evaluation model named GABSDTEL." The results should be presented in the Results section, not in the Methods.
Reply: We apologize for any confusion caused by the previous description. We have made the necessary revisions. (Line 179-196, 222-235)

(10) Statistical analysis:

The two paragraphs of the Statistical analysis are repeated and it is recommended that only one of them be retained.

Reply: We apologize for the repetition in the two paragraphs of the Statistical analysis section. We have made the necessary revisions to retain only one of them. (Line 170)

The methods of Hosmer-Lemeshow test should be described in the Methods section.

In addition, please present whether the P-value tests were one-sided or two-sided.

Reply: We have made the necessary revisions. We add the methods of Hosmer-Lemeshow test in Methods section. All of the P-value tests conducted in this study were two-sided and we have made a note in the manuscript. (Line 186-194)

3. Results:

(1) We hope the authors consider using a flow diagram to state the numbers of individuals in the screening and analysis stages of this study. Besides, give reasons for non-participation at each stage. You could refer to https://cdt.amegroups.org/article/view/117357/html (Figure 1)

Reply: We have made the necessary revisions to incorporate a flow diagram in our revised manuscript. This diagram will provide a visual representation of the participant selection process, including the number of individuals at each stage and the reasons for non-participation. (Figure 1)

(2) There seems to be a discrepancy in the numbers of males between the text and Table 1. Lines 170-172 in the text: the number of males with EHT is 454 (64.9%), and the number of males in the control group is 450 (64.3%). However, in Table 1, the data is the opposite. Please revise this issue for consistency.

Reply: Upon further inspection, we have confirmed that there is an error in the gender data in the text. The correct data should be as follows: in the EHT group, the number of males is 450, accounting for 64.3%, and in the control group, the number of males is 454, accounting for 64.9%. We will promptly revise the text to reflect these accurate figures. (Line 199-201)

(3) Lines 173-177: "Significant differences between ... and LDL-C (P<0.001)." Do not report only vague bounds such as "P<0.01" or "P<0.05",

but report the exact P value (e.g., "smoking (P=0.004)", "education level (P=0.003)", "FBG (P=0.026)", "HDL-C (P=0.004)"). Besides, in Table 1, LDL-C (P=0.001), not P<0.001.

Reply: We appreciate your attention to detail and will make the necessary revisions to address this concern. We will replace the vague p-value expressions such as "P<0.05" with the exact p-values in the manuscript. (Line 203-207)

(4) Lines 176-177: "No significant differences were observed in gender, drinking, TC , ApoA1, ApoB, or Lp(a)." According to Table 1, PR should also be involved in this sentence.

Reply: Upon further inspection, we have confirmed that PR was omitted in this sentence. We have carefully make the necessary adjustment to include PR in the sentence. (Line 206)

(5) Lines 179-183: "The results indicated that the rs662799 polymorphism ... the control group (P=0.006)." Except for the P-value results, the OR and 95% CI data should also be reported in the text.

Reply: We have made necessary revisions to ensure that the OR (odds ratio) and 95% CI (confidence interval) data are included in the text alongside the P-value results. (Line 208-215)

(6) Lines 184-189: "Furthermore, a multivariate logistic regression model was employed ... Table 2)." It should be Table 3.

Reply: We have made the necessary correction to accurately reference Table 3 instead of Table 2. (Line 217)

(7) Lines 191-193: "We further developed a diagnostic model of EHT using APOA5 rs662799, age, BMI, smoking, diabetes, LDL-C, TG, and education level as variables (GABSDTEL score)." Please clarify "APOA5 rs662799" as "G allele of rs662799". Besides, according to the results of logistic regression in Table 3, the P value of UA is 0.089, and as a result, UA is suggested to be removed from Figure 1.

Reply: We have made the necessary revisions to clarify "APOA5 rs662799" as "G allele of rs662799". Regarding the UA issue you mentioned, we have removed UA and re-generated the nomogram, calibration curve, ROC curve, and made the necessary modifications to the corresponding experimental results in the document. (Figure 2)

(8) Could you please add some variables in Table 3? Such as β, SE, and Wald χ^2.

Reply: We have made the revisions to incorporate additional variables into Table 3, including β, SE, and Walds. (Table 3)

(9) Lines 193-195: "The model had a AUC of 0.744 (95% CI: 0.69–0.80) with a sensitivity of 68.31% and a specificity of 73.19% (Fig. 2A)." Please report the P value (P<0.001) as well, which was reported in the Abstract.
Reply: We have made the revisions to report *P* value in this sentence. (Line 56-57)

(10) Lines 195-197: "The Hosmer-Lemeshow test indicated no significant difference between the predictive calibration curve and the ideal curve (P=0.989), suggesting that the model has an excellent calibration ability." Please cite Figure 2B.
Reply: We have made the revisions to cite Figure 3 (previously labeled as Figure 2B) in this sentence. (Line 229)

(11) Lines 269-270: "The established nomogram and GABSDTEL score in our study may serve as a new diagnostic method for EHT in the Chinese population." What is the best cut-off point? Please report it in the Results section.
Reply: We appreciate your attention to this detail. Referring to this model as a diagnostic method is inaccurate. This model can assist in assessing the risk of hypertension. Therefore, we have refined the expression here, calling it a risk assessment method. Besides, we have added a description of the model threshold and the related calculation methods. (Line 185-186, 236-240)

4. Please give the full name of the abbreviations when they were first mentioned in the text, e.g., BMI, EDTA, etc.
Reply: We have made the revisions to provide the full names of the abbreviations at their first mention in the text. (Line 66, 138)

5. Please mark in the main text where you want to cite fig 2B.
Reply: In the revised version, we have marked the location in the main text where we want to cite Figure 3 (previously labeled as Figure 2B). The citation is now included at the appropriate place, providing a clear reference to the corresponding figure. (Line 229)

6. Please indicate how the data were presented in the footnote of table 1.
Reply: In the revised version, we have included a footnote in Table 1 to indicate how the data were presented. The footnote provides a clear explanation of the format and units used for the data presented in the table. (Table 1)

7. ref 10 and 11 are duplicate.

Reply: We apologize for the oversight. Upon reviewing the manuscript, we have identified the duplicate references and made the necessary corrections in the revised version. (Line 363)