

Response Letter

Revision: (ID: 103679)

Journal: World Journal of Gastrointestinal Oncology

Article title: Blood-based machine learning classifiers for early diagnosis of gastric cancer via multiple miRNAs

January 16, 2025

Dear Editors and Reviewers:

Thanks for your kind efforts and constructive comments on our paper submitted to *World Journal of Gastrointestinal Oncology* (ID: 103679). All of us authors have carefully revised the manuscript in accordance with your kind advices and here are our replies to the remarks. A marked-up manuscript with the changes highlighted is going to be resubmitted simultaneously with the current response letter. The followings are point-by-point responses to the comments.

Reviewer #1

Comments to the authors:

Ma et al. have investigated the potential of using serum miRNAs as biomarkers for early detection of gastric cancer (GC). The authors have collected serum samples from multiple centers and used sRNA-seq and RT-qPCR to identify differentially expressed miRNAs in healthy controls, patients with early-stage GC, and patients with advanced-stage GC. A total 6 miRNAs (miR-452-5p, miR-5010-5p, miR-27b-5p, miR-5189-5p, miR-552-5p, and miR-199b-5p) were found. These results were validated using sequencing data from public databases, and a machine learning model was constructed to assess the efficacy of 6 miRNAs for GC detection. This manuscript reported high AUC for the GC classifier in both training cohort (AUC = 0.983) and two validation cohorts (AUCs = 0.995 and 0.979). The classifier has potential benefits for early detection and population screening of GC. Overall, the results are promising.

Response: Thanks for your positive comments. We have carefully revised this manuscript based on the reviewer's comments. We hope that the revision would

satisfactorily address the reviewer's concerns. Our point-to-point responses are as follows.

However, there are several questions and comments that need to be addressed:

1. The author's goal is to find miRNA biomarkers in blood for early gastric cancer detection. But the article mentions both serum and plasma samples, which are different. Please clarify which one was used in the study to avoid confusion.

Response: Thanks for your professional suggestion, and we are sorry for the confusion caused to you. It should be pointed out that totally 275 plasma samples from three clinical centers and 10026 serum samples from Gene Expression Omnibus (GEO) database were used in this study. We have corrected the confused descriptions in the revised manuscript and the details were listed as follows:

Line 12: the word "serum" was corrected to "plasma";

Line 26: the word "Serum" was corrected to "Plasma";

Line 46: the word "serum" was corrected to "plasma";

Line 111: the word "serum" was corrected to "plasma";

Line 117: the word "plasma" was corrected to "serum";

Line 158: the word "serum" was corrected to "plasma";

Line 193: the word "serum" was corrected to "plasma";

Line 252: the word "plasma" was corrected to "serum";

Line 258: the word "serum" was corrected to "plasma";

Line 263: the word "plasma" was corrected to "serum";

Line 289: the word "serum" was corrected to "plasma";

Line 397: the word "plasma" was corrected to "serum";

Line 409: the word "plasma" was corrected to "serum";

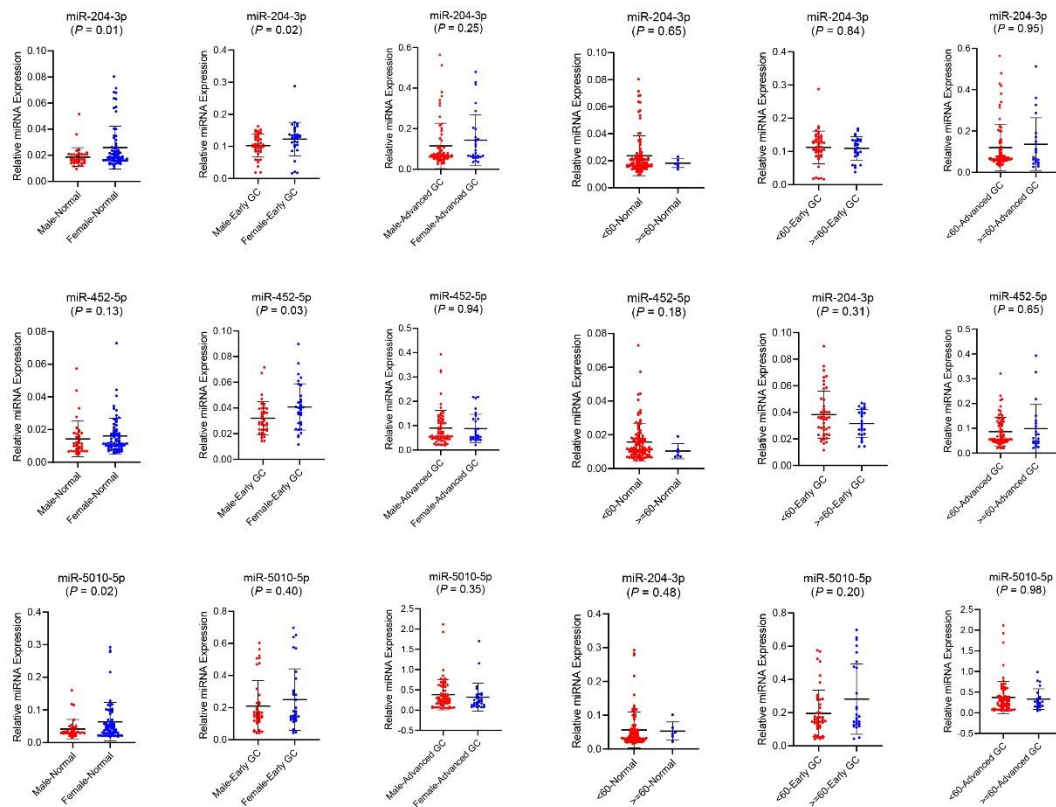
Line 419: the word "serum" was corrected to "plasma".

Thanks again for your advice, which avoids mistakes and makes the article more perfect.

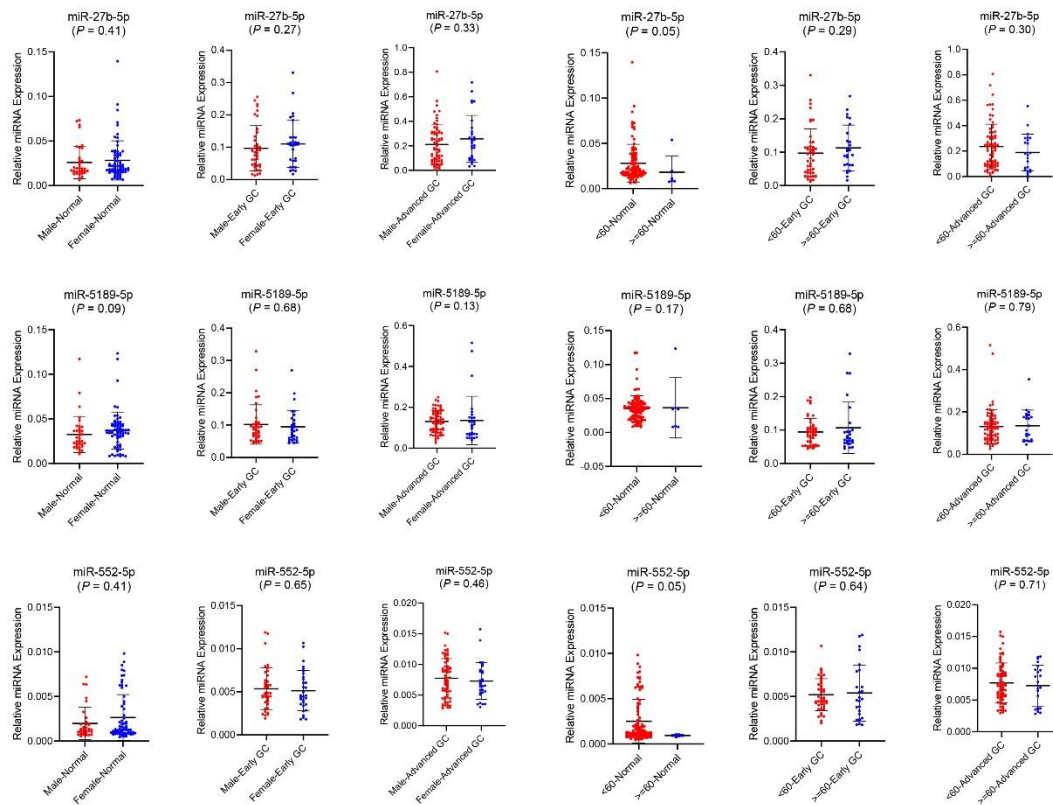
2. The authors noted differences in age and gender composition among the three

subject groups. However, could these variations in age and gender potentially impact the expression levels of miRNAs in serum?

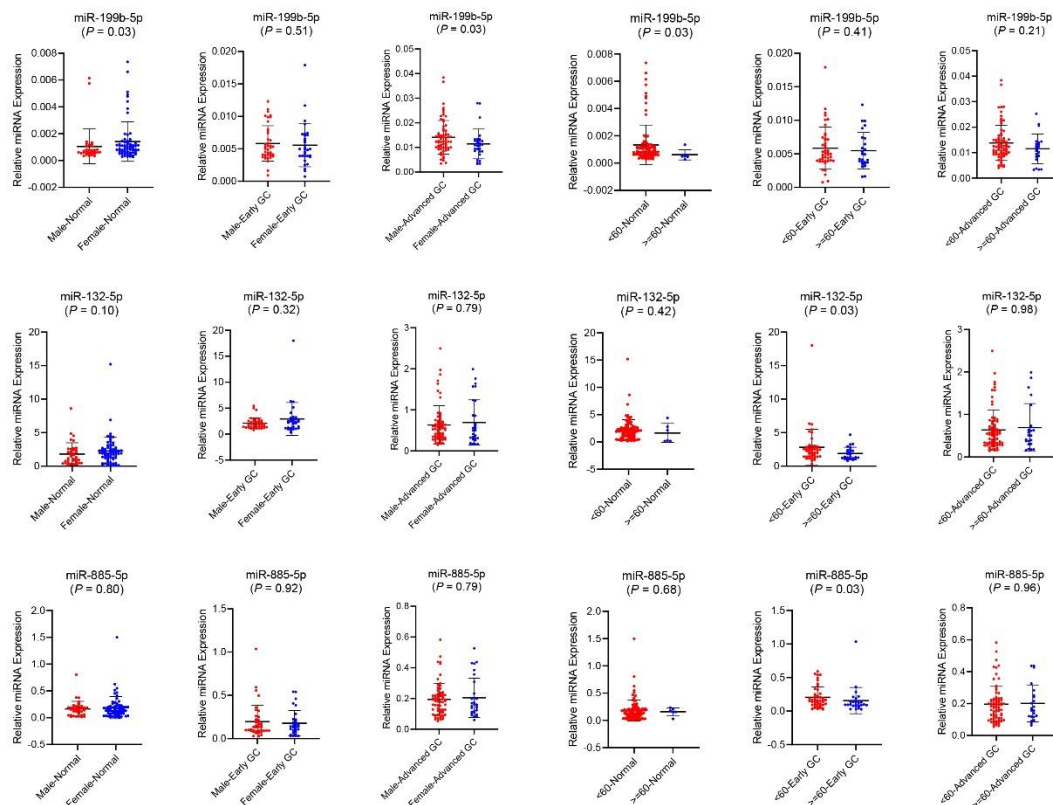
Response: Thanks for the question. The differences in age and gender composition would not impact the expression levels of miRNAs in plasma. To prove the perspective, we have analyzed the differences of miRNAs expression in age and gender utilized Mann-Whitney U test, respectively in the normal group, early gastric cancer group and the advanced gastric cancer group. As shown in Supplementary Figure 1-5, most of the miRNAs expression shown no significant differences in "Male" group and "Female" group, and in "<60" years group and ">=60" years group. The Supplementary Figure 1-5 were attached as follows:



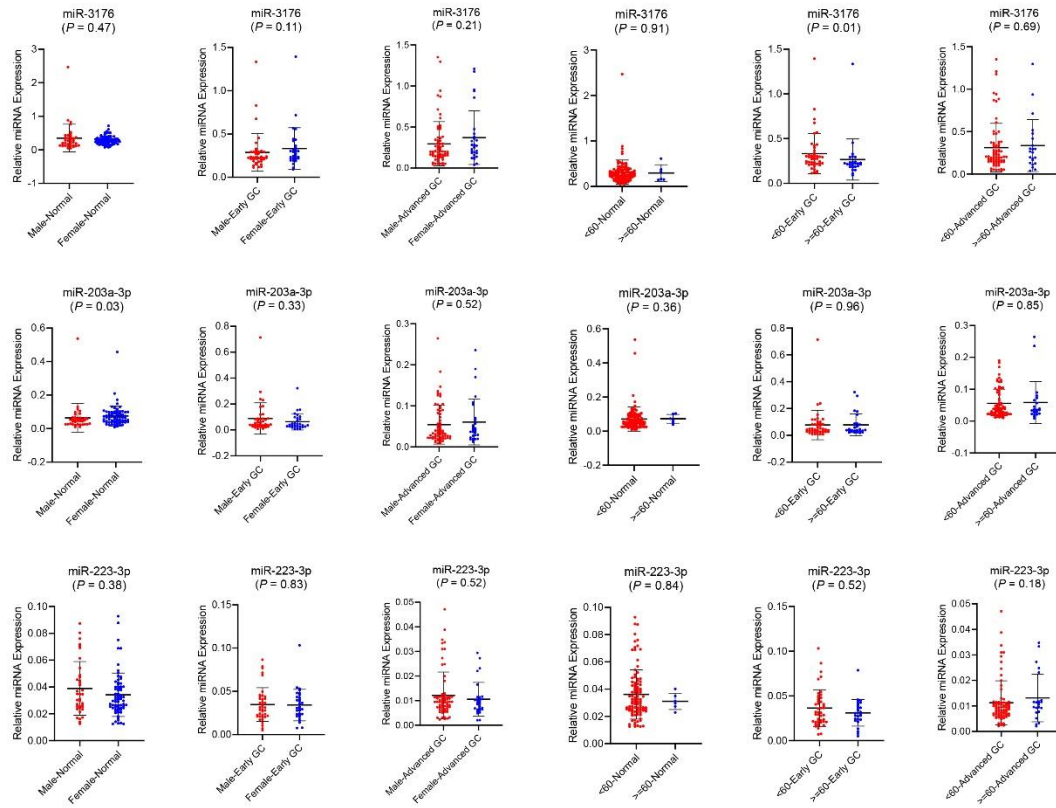
Supplementary Figure 1 The differences analyses of miR-204-3p, miR-452-5p and miR-5010-5p expression in gender and age by Mann-Whitney U test.



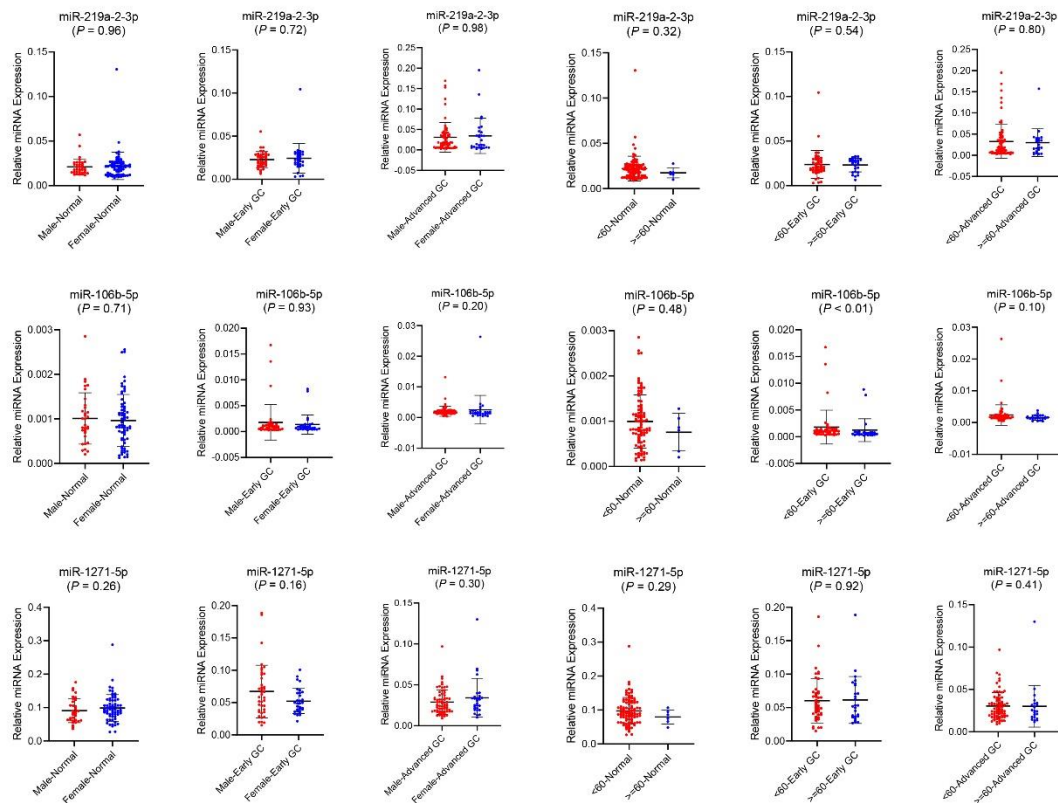
Supplementary Figure 2 The differences analyses of miR-27b-5p, miR-5189-5p and miR-552-5p expression in gender and age by Mann-Whitney U test.



Supplementary Figure 3 The differences analyses of miR-199b-5p, miR-132-5p and miR-885-5p expression in gender and age by Mann-Whitney U test.



Supplementary Figure 4 The differences analyses of miR-3176, miR-203a-3p and miR-223-3p expression in gender and age by Mann-Whitney U test.



Supplementary Figure 5 The differences analyses of miR-219a-2-3p, miR-106b-5p and miR-1271-5p expression in gender and age by Mann-Whitney U test.

The authors also added the statement in the revised manuscript, as " Nevertheless, no substantial difference in age was noted between the advanced GC group and control group. It was pointed out that the differences in age and gender composition would not impact the expression levels of miRNAs in plasma (Supplementary Figure 1-5)" (*Line 245 to 248*). We hope the revision can meet with your approval. Thanks again for your professional question, which makes our article more perfect.

3. The authors employed RNA-seq and RT-qPCR to identify and validate the differential expression of miRNAs in serum samples. Subsequently, the MLP-ANN model was constructed and validated using samples from three distinct databases, which were obtained through RNA-seq and miRNA microarray techniques, respectively. The intriguing questions arises during this process: Why does the author use the data obtained by different technical methods to construct the model? Are the

data obtained by different techniques comparable? How did that affect the results?

Response: Thanks for the questions. We will response from the following aspects.

Firstly, in view of miRNA can stably circulate in body fluid including plasma and serum, this study employed 275 plasma samples to identify and validate the critical miRNAs and employed 10,026 serum samples to construct the MLP-ANN model.

Secondly, for the technical methods, RNA-seq and RT-qPCR were used to identify the critical miRNAs. The technical method of three datasets used for constructing MLP-ANN model were miRNA microarray by a same experiment platform - GPL21263. The reason why the authors use a different technical method to construct is that the plasma sample size we collected is not enough for constructing a multi-miRNAs diagnostic model. To solve this deficiency, the authors have retrieved several public databases and ultimately screened the three datasets (GSE211692, GSE112264 and GSE106817), including totally 10,026 serum samples. Although the technical method of the three datasets is different than RNA-seq or RT-qPCR, the sample size of 10,026 serum samples can improve the robustness of diagnostic model. Additionally, the samples within the three datasets were analyzed by a same experiment platform and a same technical method, which has decreased the heterogeneity of the study as far as possible.

Thirdly, the purpose of our RNA-seq and RT-qPCR experiments is to find the critical miRNAs that express stably high in gastric cancer, which would be considered as the potential biomarkers. It is worth mentioned that the critical miRNAs were identified by the relative transcriptional expression levels and their ranks, which would not change with different detection methods. Thus, it is reasonable that we used the three datasets by miRNA microarray technique to construct a diagnostic model.

Thanks again for your professional review work on our article, and we hope the responses can meet with your approval.

4. The authors note that both early and advanced gastric cancers were diagnosed according to histological criteria, and that the surgical margins of postoperative

specimens were negative. Which type of TNM Stage IV patients were treated surgically?

Response: Thanks for your insightful question. I would like to clarify that the Stage IV gastric cancer patients in our study did not undergo surgical treatment. Their samples were obtained through endoscopic biopsy rather than surgical procedures. The statement "early and advanced GCs were definitively diagnosed on the basis of histological standards, with negative resection margins in postoperative specimens" has been corrected as "early and advanced GCs were definitively diagnosed on the basis of histological standards" (*Line 126-128*). Thanks again for your professional work, and we hope the revision can meet with your approval.

5. The authors mentioned that the differential expression of miRNAs was detected and verified in upper gastric cancer. Is there any difference between upper GC and other types of GC?

Response: Thanks for the question. We feel sorry for our careless language mistake. The statement of "Participants opportunistically screened for untreated early upper GC were selected from three centres between January 2021 and December 2023: Guangxi Medical University Cancer Hospital, Guilin People's Hospital and Youjiang Medical University Affiliated Hospital" was corrected as "Participants opportunistically screened for untreated early GC were selected from three centres between January 2021 and December 2023: Guangxi Medical University Cancer Hospital, Guilin People's Hospital and Youjiang Medical University Affiliated Hospital" (*Line 121 to 124*). Thank you again for your careful reading and professional work.

Minor comments:

1. In Figure 2-1C, the label of the Y-axis is incomplete. In Figure 2-1E, the vertical axis is incomplete. In Figure 2-1G, the statistical label is missing. In Figures 2-1 and 2-2, the size of the figures, the color of the error line, and the case of the statistical label P are inconsistent.

Response: Thanks for your comments. We have revised the Figure 2-1 and Figure 2-2 based on your comments. The details were listed below.

1) For Figure 2-1A and Figure 2-1B, the word "mRNA" in the label of the Y-axis was corrected to "miRNA".

2) For Figure 2-1C, the word "miR-5010" in the label of the Y-axis was corrected to "miR-5010-5p"; the word "mRNA" in the label of the Y-axis was corrected to "miRNA".

3) For Figure 2-1D, the word "mRNA" in the label of the Y-axis was corrected to "miRNA".

4) For Figure 2-1E, the word "mRNA" in the label of the Y-axis was corrected to "miRNA"; the vertical axis was revised to be complete.

5) For Figure 2-1F, the color of the error line was revised to be consistent with other figures.

6) For Figure 2-1G, the color of the error line was revised to be consistent with other figures, and the statistical label " $***P<0.01$ " was added.

7) For Figure 2-1 A-G, all the statistical label P were corrected to upper cases, and the size of the figures were adjusted.

8) For Figure 2-2D, the word "miR-203-5p" in the label of the Y-axis was corrected to "miR-203a-3p".

9) For Figure 2-2F, the word "miR-219a-3p" in the label of the Y-axis was corrected to "miR-219a-2-3p".

10) For Figure 2-2G, the color of the error line was revised to be consistent with other figures.

11) For Figure 2-2H, the word "miR-1271" in the label of the Y-axis was corrected to "miR-1271-5p", and the color of the error line was revised to be consistent with other figures.

12) For Figure 2-2 A-H, all the statistical label P were corrected to upper cases, and the size of the figures were adjusted.

The revised Figure 2-1 and Figure 2-2 have been inserted in the revised manuscript. We hope the revision can meet with your approval.

2. The punctuation usage in the article exhibits irregularities. Specifically, the article inappropriately utilizes Chinese punctuation and fails to adhere to the required spacing before certain symbols. It is recommended that the author make the necessary corrections to rectify these issues.

Response: Thanks for your careful reading and feedback. The authors have checked and corrected the Chinese punctuation throughout the whole manuscript, which were marked in the revised manuscript. We hope the revision can make the manuscript more perfect and meet with your approval.

Editorial office's comments to the authors:

Science Editor:

1. Scientific quality: The authors submitted a basic study of blood-based machine learning classifiers for the early diagnosis of gastric cancer via multiple miRNAs.

(1) Classification: Grade C;

Response: We sincerely thank the editors and reviewers for their professional work on our article. The authors have revised the manuscript based on the comments, and we hope the revision can meet with your approval.

(2) Summary of the Peer-Review Report: The classifier has potential benefits for early detection and population screening of GC. However, the following issues still need to be addressed.

(2.1) Please clarify which type of sample was used in this study – serum or plasma;

Response: Thanks for your suggestion, and we are sorry for the confusion caused to you. It should be pointed out that totally 275 plasma samples from three clinical centers and 10,026 serum samples from Gene Expression Omnibus (GEO) database were used in this study. We have corrected the confused descriptions in the revised manuscript. For the detailed list, please check the response of Reviewer #1. Thanks for the comment again and we hope the revision can be acceptable to you.

(2.2) They did not evaluate whether these differences in age and gender would affect the expression level of miRNAs in serum;

Response: Thanks for the question. The differences in age and gender composition would not impact the expression levels of miRNAs in plasma. To prove the perspective, we have analyzed the differences of miRNAs expression in age and gender utilized Mann-Whitney U test, respectively in the normal group, early gastric cancer group and the advanced gastric cancer group. For the details, please check the response of Reviewer #1.

(2.3) Do patients with TNM stage IV need to receive surgical treatment?

Response: Thanks for the question. In general, surgical intervention for patients with TNM stage IV gastric cancer is typically considered only in specific circumstances. While stage IV indicates that the disease has metastasized and is generally associated with a poor prognosis, there are cases where surgery may be beneficial. For instance, if the metastasis is limited and accessible, or if there is a particular complication that can be relieved by surgery (such as a blockage or stomach bleeding), then surgical options might be considered. However, the decision to proceed with surgery should be individualized, taking into account the patient's overall health, preferences, and the extent of the disease. Hence, a thorough assessment by a multidisciplinary team is crucial in guiding treatment decisions for these patients.

(2.4) Is there any difference between upper gastric cancer and other types of gastric cancer?

Response: Thanks for the question. We feel sorry for our careless language mistake. The statement "upper gastric cancer" should be corrected as "gastric cancer", which we have modified in the revised manuscript. For more details, please check the response of Reviewer #1.

(3) Manuscript Type: After verification, the manuscript type is "Basic Study".

Response: We feel great thanks for your professional work on our article. The authors

have revised the manuscript based on the comments, and we hope the revision can meet with your approval.

2. Specific comments

(1) Country/Territory of origin: China.

(2) The language classification is Grade C. Please provide the latest language certificate after Return the Manuscript to Author for Revision. Please visit the following website for the professional English language editing companies that we recommend: <https://www.wjgnet.com/bpg/gerinfo/240>.

Response: Thanks for your suggestion. The language of revised manuscript has been polished by a recommend professional English language editing company, MedE Editing Group (<http://meditorexpert.com>). The certificate is attached below.

EDITORIAL CERTIFICATE

(Ref. MEDESPGJY-MS2025011310A)

We herein certify that the following document has been edited for English language by a native English speaking medical editor at MedE Medical Editing Group. The edited paper has reached grade A in language evaluation for SCI journals.

Manuscript title

Blood-based machine learning classifiers for early diagnosis of gastric cancer via multiple miRNAs

Authors and affiliations

Not shown in the paper

Date issued

January 14, 2025

*We are **NOT** responsible for any error in the added content to our edited version after this date.

MedE Medical Editing Group Inc.
+86-10-82082089
<http://meditorexpert.com>
Email: inf@meditorexpert.com
File-code:MS2025011310-A



(3) Policy of allowing co-first authors and co-corresponding authors who made equal contribution to a manuscript (<https://www.wjgnet.com/bpg/GerInfo/310>).

(4) Audio Core Tip. In order to attract readers to read the full-text article, we request that the first author make an audio file describing the final core tip. This audio file will be published online, along with the article. The author can invite English language editing company to assist in resolving the language issues of Audio Core Tip.

Response: Thank you for your helpful comments regarding the Audio Core Tip. The first author has prepared the audio file, clearly describing the final core tip as requested. The language issues of Audio Core Tip have been polished by MedE language editing company (<http://meditorexpert.com>), and the audio file has been uploaded to the system, please check.

(5) There are issues with the references:

To ensure the accuracy of the references, please use "Edit References by Auto-Analyser"

(<https://www.f6publishing.com/Forms/main/ArticleReferenceTool.aspx>) to edit the references of the manuscript.

Response: Thanks for the advice. We have edited the references of the manuscript using " Edit References by Auto-Analyser" tool. The corrections of references have been marked in the revised manuscript, please check.

(6) Figures. Uniform presentation should be used for figures showing the same or similar contents; for example, "Figure 1 Pathological changes of atrophic gastritis after treatment. A: ...; B: ...; C: ...; D: ...; E: ...; F: ...; G: ...".

Abbreviations must be defined upon first appearance in the Figure Legends. Do not use non-standard abbreviations, unless they appear at least two times in the text preceding the first usage/definition.

Original figure documents. In the meantime, authors should provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure

that all graphs or arrows or text portions can be reprocessed by the editor, and upload it to the file destination of “Image File” in the F6Publishing system.

Response: Thanks for the comments. The authors have prepared and arranged the figures using PowerPoint, which have been uploaded to the F6Publishing system, please check.

(7) Notes in figures and tables.

Data with statistical significance in a figure or table should be denoted using superscripted alphabetical lettering (don't include symbols, such as *, #, †, §, ‡, ¥, @...), such that aP < 0.05 and bP < 0.01. If there are other series of P values, the alphabetical subscripted denotation format is continued, such that cP < 0.05 vs control, dP < 0.01 vs control, eP < 0.05 vs group A, and fP < 0.01 vs group B. Data that are not statistically significant should not be denoted, i.e. P > 0.05 is not an allowed denotation.

Please don't include any *, #, †, §, ‡, ¥, @...in your Figures or Tables, please use superscript numbers for illustration. For example, 1Computed tomography.

Response: Thanks for the comments. The authors have checked throughout the manuscript, and no special symbols is found.

(8) Please upload the approved grant application form(s) or funding agency copy(ies) of any approval document(s).

Response: Thanks for the reminder. The funding approval documents have been uploaded to the system, please check.

3. Recommendation: Conditional acceptance.

Language Quality: Grade C (A great deal of language polishing)

Scientific Quality: Grade C (Good)

Response: We sincerely thank you for your professional work on our article. The authors have revised the manuscript based on the comments, and we hope the revision can meet with your approval.

Once again, we sincerely thank the editors and reviewers for their valuable feedback that we have used to improve the quality of our manuscript. And we hope the revised manuscript can be suitable for publication in *World Journal of Gastrointestinal Oncology* this time. Looking forward to your reply.

Best regards,

Lian-Ying Ge

Department of Endoscopy, Guangxi Medical University Cancer Hospital, No. 71 Hedi Road, Nanning 530021, Guangxi Zhuang Autonomous Region, China.

Email: gxgly@hotmail.com