Peer Review File

Article Information: https://dx.doi.org/10.21037/gs-24-374

Reviewer A

We feel very grateful to receive your opinions and suggestions for our manuscript. We want to thank you for constructive and insightful advice. It brings us to review and reconsider our manuscript. In this new version of the manuscript, we have made careful modifications to the original manuscript. We hope the new manuscript will meet your standard. All changes made to the text are highlighted. Below you will find our point-by-point responses to your comments/ questions.

1.In the abstract, the authors need to explain why a MR approach could address the causal relationship in the background, the clinical samples in the datasets used including ethnic groups and how the cathepsins were measured and how major salivary gland neoplasms were diagnosed in the methods, and have a more detailed comment for the clinical implications of the findings.

Response: Thank you for your helpful review. We have revised the manuscript and 1) explained why the MR approach could address the causal relationship in the background,2) added the clinical samples in the datasets used, including ethnic groups, described how cathepsins were measured and how major salivary gland neoplasms were diagnosed in the methods, 3) further explained the clinical implications of the findings detailly.

1) Mendelian randomization (MR) is a significant genetic method that employs single nucleotide polymorphisms (SNPs) as instrumental variables (IVs). This approach reduces confounding effects, enabling the analysis of causal relationships between exposure traits and outcome diseases(Line 29-32).

2)In this research, we collected IVs associated with eleven different types of cathepsins (including cathepsins D, L1, B, E, F, G, H, O, S, L2, and Z) from the MRC integrative epidemiology unit (IEU) open genome-wide association studies (GWAS) database. Data for cathepsins D and L1 were sourced from the SCALLOP consortium, which included 21,758 Europeans identified via the Olink proximity extension assay (PEA). Cathepsins B, E, F, G, H, O, S, L2, and Z were obtained from the INTERVAL study involving 3,301 European participants using the SOMAscan assay. We also collected data on benign major salivary gland neoplasms (BMSGNs) from the FinnGen database, consisting of 3,353 cases and 450,380 controls, and information on major salivary gland carcinomas (MSGCs) from the UK Biobank, which included 105 cases and 456,243 controls. Diagnostic criteria for both BMSGNs and MSGCs followed the international statistical classification of diseases and related health problems 10th revision (ICD-10) classification. (Line 35-45).

3） Elevated CTSF, CTSO, and CTSL2 levels may serve as significant biomarkers for diagnosing BMSGNs definitively. Conversely, reduced levels of CTSL2 provide a novel foundation for diagnosing MSGCs and differentiating them from BMSGNs. Moreover, CTSF, CTSO, and CTSL2 represent potential new targets for therapeutic intervention in BMSGNs and MSGCs (Line 58-62).

2.In the introduction, please review the controversy regarding the causal relationship between cathepsins and major salivary gland neoplasms and analyze the potential reasons, as well as explain why MR approach could address the controversy.

Response: Thank you for your kind reminder. We have re-read all current studies on cathepsin and salivary gland neoplasms, revised the manuscript, and explained 1) the controversy regarding the causal relationship between cathepsins and major salivary gland neoplasms, and analyzed the potential reasons,2) described the advantages of the MR approach for addressing the controversy.

Various cathepsin-targeting inhibitors have been developed and utilized in treating clinical diseases in recent decades, showing promising therapeutic effects(17). As a result, exploring the relationship between cathepsins and MSGNs presents a significant opportunity for enhancing treatment strategies for these tumors. Current research on cathepsins in relation to MSGNs is limited, predominantly consisting of retrospective studies that examine the direct association between cathepsin expression levels in tissue samples and SGNs. A primary challenge in this field is that diagnosing salivary gland tumors largely relies on histopathological methods. Additionally, the assessment of cathepsin diversity and the measurement of circulating cathepsin levels are time-consuming and labor-intensive processes, further complicated by the small sample sizes often encountered in studies of salivary gland tumors. This challenge may be addressed through the MR method, which employs genetic IVs to investigate the causal relationship between cathepsins and MSGNs from a genetic standpoint. This approach can potentially provide valuable insights for future basic and clinical research while reducing the time and effort required for subsequent studies(Line 129-144).

3.In the methodology, please describe the clinical samples, the measurement of cathepsins and diagnoses of major salivary gland neoplasms.

Response: Thank you for your helpful suggestion. We have revised the manuscript and described in more detail the clinical sample , the measurement of cathepsins, and the diagnoses of major salivary gland neoplasms in the methods of data source. However, due to the restriction of the databases, individual information on age, gender, and disease progress cannot be obtained for further analysis.

The summary data for various cathepsins, including D, L1, B, E, F, G, H, O, S, L2, and Z, utilized in this study were sourced from two primary datasets: the Systematic and Combined Analysis of Olink Proteins (SCALLOP) consortium and the INTERVAL study. These datasets are accessible online through the MRC Integrative Epidemiology Unit (IEU) Open GWAS database (IEU OpenGWAS project (mrcieu.ac.uk)). Data for cathepsin D and L1 were obtained from the SCALLOP consortium, which comprised a European population of 21,758 individuals, measured using the Olink proximity extension assay (PEA) technique. In contrast, cathepsins B, E, F, G, H, O, S, L2, and Z were drawn from the INTERVAL study, which included a sample of 3,301 individuals from the European population and employed an expanded version of an aptamer-based multiplex protein assay known as SOMAscan(22,23)(Line 164-173)

The diagnostic criteria for Benign and Malignant Salivary Gland Neoplasms (BMSGNs) and Malignant Salivary Gland Carcinomas (MSGCs) adhered to the international statistical classification of diseases and related health problems, 10th revision (ICD-10), which encompasses both benign and malignant tumors of the parotid, submandibular, and sublingual salivary glands(Line 178-182).

4. Please also consider to cite several related papers : 1. Chang YH, Kuo C, Chang TH, Chen MK, Lin JC. Treatment outcome and prognostic factors analysis of carcinoma ex pleomorphic adenoma of major salivary glands. Ther Radiol Oncol 2023;7:14. 2. Asarkar AA, Chang BA. Editorial for nomograms-based prediction of overall and cancer-specific survivals for patients diagnosed with major salivary gland carcinoma. Ann Transl Med 2021;9(23):1709. doi: 10.21037/atm-2021-10. 3.Bou-Samra P, Chang A, Guo E, Azari F, Kennedy G, Santini JT Jr, Bensen ES, Jarrar D, Singhal S. Cathepsin detection to identify malignant cells during robotic pulmonary resection. Transl Lung Cancer Res 2023;12(12):2370-2380. doi: 10.21037/tlcr-23-370.

Response: Thank you for your valuable suggestion. We have thoroughly reviewed the related papers and revised the manuscript accordingly, citing them in the article.

A recent analysis of treatment efficacy and prognostic factors for Carcinoma ex Pleomorphic Adenoma (CXPA) of the major salivary glands revealed an overall treatment failure rate ranging from 33.3% to 53.0%. Furthermore, the 5-year overall survival rate is only between 30% and 76%. These results highlight the significant need for treatment effectiveness and patient prognosis improvement(6)(Line 79-84).

Many studies have explored the prediction of overall and tumor-specific survival in patients with major salivary gland carcinomas（MSGCs）through nomograms, which have proven effective in estimating patient outcomes. However, a notable scarcity of research on early diagnosis and treatment for these conditions remains(8)(Line 90-93)

Cathepsins, extensively studied in mammals, play a pivotal role in maintaining intracellular and extracellular balance and are linked to tissue differentiation, intracellular protein degradation, hormone maturation, antigen processing, immune responses, and the malignant metastasis of tumors (10,11)(Line 104-107)

Reviewer B

We feel very grateful to receive your opinions and suggestions for our manuscript. We want to thank you for constructive and insightful advice. It brings us to review and reconsider our manuscript. In this new version of the manuscript, we have made careful modifications to the original manuscript. We hope the new manuscript will meet your standard. All changes made to the text are highlighted. Below you will find our point-by-point responses to your comments/ questions.

1.In the abstract, the authors need to explain why a MR approach could address the causal relationship in the background, the clinical samples in the datasets used including ethnic groups and how the cathepsins were measured and how major salivary gland neoplasms were diagnosed in the methods, and have a more detailed comment for the clinical implications of the findings.

Response: Thank you for your helpful review. We have revised the manuscript and 1) explained why the MR approach could address the causal relationship in the background,2) added the clinical samples in the datasets used, including ethnic groups, described how cathepsins were measured and how major salivary gland neoplasms were diagnosed in the methods, 3) further explained the clinical implications of the findings detailly.

1) Mendelian randomization (MR) is a significant genetic method that employs single nucleotide polymorphisms (SNPs) as instrumental variables (IVs). This approach reduces confounding effects, enabling the analysis of causal relationships between exposure traits and outcome diseases(Line 29-32).

2)In this research, we collected IVs associated with eleven different types of cathepsins (including cathepsins D, L1, B, E, F, G, H, O, S, L2, and Z) from the MRC integrative epidemiology unit (IEU) open genome-wide association studies (GWAS) database. Data for cathepsins D and L1 were sourced from the SCALLOP consortium, which included 21,758 Europeans identified via the Olink proximity extension assay (PEA). Cathepsins B, E, F, G, H, O, S, L2, and Z were obtained from the INTERVAL study involving 3,301 European participants using the SOMAscan assay. We also collected data on benign major salivary gland neoplasms (BMSGNs) from the FinnGen database, consisting of 3,353 cases and 450,380 controls, and information on major salivary gland

carcinomas (MSGCs) from the UK Biobank, which included 105 cases and 456,243 controls. Diagnostic criteria for both BMSGNs and MSGCs followed the international statistical classification of diseases and related health problems 10th revision (ICD-10) classification. (Line 35-45).

3）Elevated CTSF, CTSO, and CTSL2 levels may serve as significant biomarkers for diagnosing BMSGNs definitively. Conversely, reduced levels of CTSL2 provide a novel foundation for diagnosing MSGCs and differentiating them from BMSGNs. Moreover, CTSF, CTSO, and CTSL2 represent potential new targets for therapeutic intervention in BMSGNs and MSGCs (Line 58-62).

2.In the introduction, please review the controversy regarding the causal relationship between cathepsins and major salivary gland neoplasms and analyze the potential reasons, as well as explain why MR approach could address the controversy.

Response: Thank you for your kind reminder. We have re-read all current studies on cathepsin and salivary gland neoplasms, revised the manuscript, and explained 1) the controversy regarding the causal relationship between cathepsins and major salivary gland neoplasms, and analyzed the potential reasons,2) described the advantages of the MR approach for addressing the controversy.

Various cathepsin-targeting inhibitors have been developed and utilized in treating clinical diseases in recent decades, showing promising therapeutic effects(17). As a result, exploring the relationship between cathepsins and MSGNs presents a significant opportunity for enhancing treatment strategies for these tumors. Current research on cathepsins in relation to MSGNs is limited, predominantly consisting of retrospective studies that examine the direct association between cathepsin expression levels in tissue samples and SGNs. A primary challenge in this field is that diagnosing salivary gland tumors largely relies on histopathological methods. Additionally, the assessment of cathepsin diversity and the measurement of circulating cathepsin levels are time-consuming and labor-intensive processes, further complicated by the small sample sizes often encountered in studies of salivary gland tumors. This challenge may be addressed through the MR method, which employs genetic IVs to investigate the causal relationship between cathepsins and MSGNs from a genetic standpoint. This approach can potentially provide valuable insights for future basic and clinical research while reducing the time and effort required for subsequent studies(Line 129-144).

3.In the methodology, please describe the clinical samples, the measurement of cathepsins and diagnoses of major salivary gland neoplasms.

Response: Thank you for your helpful suggestion. We have revised the manuscript and described in more detail the clinical sample， the measurement of cathepsins, and the diagnoses of major salivary gland neoplasms in the methods of data source. However, due to the restriction of the databases, individual information on age, gender, and disease progress cannot be obtained for further analysis.

The summary data for various cathepsins, including D, L1, B, E, F, G, H, O, S, L2, and Z, utilized in this study were sourced from two primary datasets: the Systematic and Combined Analysis of Olink Proteins (SCALLOP) consortium and the INTERVAL study. These datasets are accessible online through the MRC Integrative Epidemiology Unit (IEU) Open GWAS database (IEU OpenGWAS project (mrcieu.ac.uk)). Data for cathepsin D and L1 were obtained from the SCALLOP consortium, which comprised a European population of 21,758 individuals, measured using the Olink proximity extension assay (PEA) technique. In contrast, cathepsins B, E, F, G, H, O, S, L2, and Z were drawn from the INTERVAL study, which included a sample of 3,301 individuals from the European population and employed an expanded version of an aptamer-based multiplex protein assay known as SOMAscan(22,23)(Line 164-173)

The diagnostic criteria for Benign and Malignant Salivary Gland Neoplasms (BMSGNs) and Malignant Salivary Gland Carcinomas (MSGCs) adhered to the international statistical classification of diseases and related health

problems, 10th revision (ICD-10), which encompasses both benign and malignant tumors of the parotid, submandibular, and sublingual salivary glands(Line 178-182).

4. Please also consider to cite several related papers : 1. Chang YH, Kuo C, Chang TH, Chen MK, Lin JC. Treatment outcome and prognostic factors analysis of carcinoma ex pleomorphic adenoma of major salivary glands. Ther Radiol Oncol 2023;7:14. 2. Asarkar AA, Chang BA. Editorial for nomograms-based prediction of overall and cancer-specific survivals for patients diagnosed with major salivary gland carcinoma. Ann Transl Med 2021;9(23):1709. doi: 10.21037/atm-2021-10. 3.Bou-Samra P, Chang A, Guo E, Azari F, Kennedy G, Santini JT Jr, Bensen ES, Jarrar D, Singhal S. Cathepsin detection to identify malignant cells during robotic pulmonary resection. Transl Lung Cancer Res 2023;12(12):2370-2380. doi: 10.21037/tlcr-23-370.

Response: Thank you for your valuable suggestion. We have thoroughly reviewed the related papers and revised the manuscript accordingly, citing them in the article.

A recent analysis of treatment efficacy and prognostic factors for Carcinoma ex Pleomorphic Adenoma (CXPA) of the major salivary glands revealed an overall treatment failure rate ranging from 33.3% to 53.0%. Furthermore, the 5-year overall survival rate is only between 30% and 76%. These results highlight the significant need for treatment effectiveness and patient prognosis improvement(6)(Line 79-84).

Many studies have explored the prediction of overall and tumor-specific survival in patients with major salivary gland carcinomas（MSGCs）through nomograms, which have proven effective in estimating patient outcomes. However, a notable scarcity of research on early diagnosis and treatment for these conditions remains(8)(Line 90-93)

Cathepsins, extensively studied in mammals, play a pivotal role in maintaining intracellular and extracellular balance and are linked to tissue differentiation, intracellular protein degradation, hormone maturation, antigen processing, immune responses, and the malignant metastasis of tumors (10,11)(Line 104-107)