**Answering Reviewers**

**Manuscript Title:** Burnout and Endoscopy-Specific Workload among Gastroenterologists and Gastroenterology Fellows in Turkey: A National Cross-Sectional Survey

**Journal:** World Journal of Gastroenterology

Dear Editor-in-Chief and Reviewers,

We sincerely thank the Editorial Board and both reviewers for their careful reading of our manuscript and for their thoughtful and constructive comments. The reviewers' suggestions have substantially improved the rigour, transparency, and clarity of this work. We have carefully revised the manuscript in response to all comments. The main revisions include a more cautious interpretation of burnout proportions, clearer reporting of recruitment and sampling limitations, expanded description of the Endoscopy Workload Module, presentation of the study-specific workload items and analytic coding as Supplementary Material, more transparent multivariable regression reporting, additional sensitivity analyses, more cautious interpretation of the CBI threshold, revision of potentially causal language, correction of inequality-symbol formatting issues, and revision of tables and figure legends for clarity.

Below, we provide a point-by-point response to each reviewer comment. Section numbers refer to the revised manuscript.

Respectfully,

Yavuz Özden, MD
On behalf of all authors

**Response to Reviewer 1**

We thank Reviewer 1 for the detailed evaluation and constructive major revision recommendations. We have carefully addressed all major concerns and the additional formatting suggestions.

**Comment 1**

The recruitment strategy relied on professional networks and snowball sampling, and the eligible denominator could not be determined. Therefore, the reported burnout proportions should not be interpreted as unbiased national prevalence estimates. The authors appropriately acknowledge this limitation, but this point should be further emphasised in the Abstract, Discussion, and Conclusion.

**Response**

We agree and have strengthened the cautionary language throughout the manuscript.

In the Abstract, we now state that the reported figures describe burnout levels in a self-selected respondent sample and should not be interpreted as unbiased national prevalence estimates. In the Conclusion, we added that the findings support system-level interventions but should not be interpreted as causal effects or national prevalence estimates.

In the Core Tip, we added that because the sample was self-selected, the findings should be interpreted as respondent-level associations rather than national prevalence estimates.

In Methods Section 2.1, we clarified that the primary purpose of the study was to describe burnout levels among responding physicians and examine associations between burnout domains and endoscopy-specific workload exposures, and that the study was not designed to generate probability-based national prevalence estimates.

In Methods Section 2.6, we added that estimates should be regarded as descriptive survey estimates rather than probability-based population prevalence estimates.

In the opening paragraph of the Results, we added that all proportions should be interpreted as descriptive findings among respondents rather than probability-based national prevalence estimates. In Results Section 3.3, we also clarified that the proportions describe the frequency of high burnout scores in this self-selected respondent sample and are not intended as unbiased national prevalence estimates.

In the Discussion, we dedicated a paragraph to explaining why the reported proportions cannot be generalized to the entire Turkish gastroenterology workforce. We also added the possibility that physicians experiencing burnout may have been more motivated to complete the survey, which could have inflated descriptive burnout estimates and strengthened perceived workload associations.

**Comment 2**

The Endoscopy Workload Module is a central component of the study, but its development and validation require more detailed reporting. The authors should provide the complete module, preferably as supplementary material, and clarify how item thresholds were selected and how each variable was coded. If formal content validity indices were not calculated, this should be explicitly stated.

**Response**

We have addressed this comment comprehensively.

In Methods Section 2.3, we expanded the description of the Endoscopy Workload Module. We now describe that items were generated through a literature review and structured discussions among gastroenterologists experienced in high-volume endoscopy practice. We also added that the draft module was reviewed qualitatively for relevance, coverage, and clarity by gastroenterologists familiar with routine and high-volume endoscopy practice.

We explicitly state that formal quantitative content validity indices, such as item-level or scale-level content validity index calculations, were not performed.

We also added details of pilot testing: the module was pilot-tested with 10 gastroenterologists for clarity and face validity; pilot feedback led to minor wording refinements to improve clarity; no items were added or removed; and pilot responses were excluded from the analytic dataset.

The study-specific questionnaire components, including all Endoscopy Workload Module items, response options, and analytic coding, are now provided in Supplementary Appendix 1. The analytic thresholds are also reported in the Methods and Supplementary Material, including the coding of monthly on-call frequency, weekly endoscopy volume, weekly outpatient clinic burden, weekly endoscopy sessions, list overruns, emergency or unscheduled procedures, out-of-hours endoscopy duty, institutional psychological support, and Likert-type variables. The Copenhagen Burnout Inventory items are not reproduced in the supplement; the validated Turkish version was used and scored according to the original scoring instructions.

**Comment 3**

The multivariable logistic regression model requires greater transparency. The authors should report the full variable selection procedure, collinearity assessment, model fit indices, and the rationale for excluding or retaining candidate predictors. Given the high prevalence of the binary outcome, the authors should also consider whether odds ratios may overestimate the strength of associations, or provide sensitivity analyses using continuous burnout scores or alternative regression approaches.

**Response**

We have substantially expanded the reporting of the regression model and added sensitivity analyses.

In Methods Section 2.5, we now provide a detailed description of candidate variable selection. Candidate variables were selected from demographic, professional,

organisational, and endoscopy-specific variables considered clinically relevant before modelling. Variables were considered for the multivariable model if they were clinically relevant and/or showed univariable associations at $p < 0.10$.

We also added details of collinearity assessment. Collinearity among candidate variables was assessed before model entry using variance inflation factors and tolerance values. Variables with substantial conceptual overlap were not entered simultaneously into the same model. A variance inflation factor greater than 5 or tolerance lower than 0.20 was considered evidence of problematic collinearity.

In Results Section 3.5, we now report that no candidate variable showed problematic multicollinearity in the variance inflation factor assessment and that all variance inflation factors were below 2.5. We also report model fit indices: the Hosmer-Lemeshow goodness-of-fit test indicated acceptable calibration ($p = 0.67$), and the area under the receiver operating characteristic curve was 0.84, indicating good discrimination.

We clarified that institutional psychological support was not entered into the model because only six respondents reported its availability, precluding meaningful estimation.

Because the binary outcome was common, we now explicitly state that odds ratios should be interpreted as measures of association rather than risk ratios and may overestimate the magnitude of associations. To address this issue, we performed sensitivity analyses using continuous personal burnout scores as the outcome and modified Poisson regression with robust standard errors to estimate prevalence ratios. These analyses showed a consistent pattern of associations and are presented in Supplementary Table 1.

**Comment 4**

The use of the CBI ≥50 threshold should be discussed more cautiously, as the authors note that this cut-off has not been specifically validated among Turkish gastroenterologists. The phrase “clinically relevant burnout” should be used carefully, and continuous score analyses should be highlighted.

**Response**

We agree and have revised the manuscript accordingly.

In Methods Section 2.3, we now describe a CBI domain score of 50 or higher as a prespecified threshold used for comparability with previous CBI-based studies. We explicitly state that this cut-off has not been specifically validated among Turkish gastroenterologists and that threshold-based findings should therefore be interpreted cautiously.

Throughout the Abstract, Results, Discussion, tables, and figure legends, we use more descriptive wording such as “score of 50 or higher,” “high burnout scores,” and “50-or-higher threshold” where appropriate. The term “clinically relevant burnout” has been retained only where it refers to the prespecified threshold-based binary outcome and is interpreted cautiously.

We also emphasised continuous CBI score analyses, including median CBI scores, Spearman correlations, and sensitivity analyses using continuous personal burnout scores. In the limitations paragraph, we added that the CBI threshold was treated as a descriptive benchmark for comparability rather than as a diagnostic criterion for burnout.

**Comment 5**

Several formatting and typographical issues should be corrected, especially misplaced inequality symbols such as “6 ≥ monthly on-call shifts” and “score 50 ≥”. Tables and figure legends should also be carefully checked for readability and consistency.

**Response**

We have carefully revised the manuscript to avoid inequality-symbol formatting errors.

Throughout the Abstract, main text, tables, and figure legends, potentially problematic inequality symbols were replaced with written expressions. For example, we now use “six or more,” “fewer than six,” “50 or higher,” “more than 60,” “60 or fewer,” and “10% or more.”

Tables and figure legends were also revised for readability and consistency. Table notes were updated to clarify that threshold-based proportions should not be interpreted as unbiased national prevalence estimates.

---

**Response to Reviewer 2**

We thank Reviewer 2 for the positive evaluation of the manuscript and for the detailed and constructive minor revision suggestions. We have addressed each point as described below.

**Comment 1**

The Abstract is generally clear, but the prevalence figures should be framed even more cautiously. Because recruitment relied on professional networks and snowball

sampling, the results should be presented as burnout levels among respondents rather than national prevalence estimates.

**Response**

We agree and have revised the Abstract accordingly.

The Results section of the Abstract now states that the reported figures describe burnout levels in a self-selected respondent sample and should not be interpreted as unbiased national prevalence estimates. The Conclusion of the Abstract now states that the findings support system-level interventions but should not be interpreted as causal effects or national prevalence estimates.

The Aim was also revised to focus on burnout levels among responding gastroenterologists and gastroenterology fellows.

**Comment 2**

In the Introduction, the authors could better justify why endoscopy creates a distinct workload burden compared with general outpatient work. When discussing the cognitive and technical demands of lesion recognition, risk stratification, and therapeutic decision-making, a recent relevant source could be cited (PMID: 41399219). The Introduction would benefit from a clearer statement of the knowledge gap. The authors should explicitly distinguish between general physician burnout literature and the narrower question of procedure-specific workload, emergency endoscopy, list overruns, and unit staffing.

**Response**

We substantially revised the Introduction.

We now explain that, compared with general outpatient work, endoscopic practice imposes distinct cognitive, technical, and physical demands that are not adequately captured by total working hours alone. The revised text describes sustained concentration, fine motor precision, physical effort, immediate management of unexpected findings or complications, lesion recognition and characterization, risk assessment, therapeutic decision-making, and escalation to surgical or multidisciplinary management.

We added the suggested reference, PMID: 41399219, in the context of lesion recognition and endoscopic decision-making.

We also clarified the knowledge gap. The revised Introduction now distinguishes broader physician burnout literature, which identifies generic contributors such as excessive workload, work-life imbalance, administrative burden, and insufficient

organisational support, from the narrower question of how procedure-specific exposures shape burnout risk in gastroenterology. We specifically identify procedural volume, emergency or unscheduled endoscopy, list overruns, after-hours communication, documentation burden, and endoscopy unit staffing as distinct and potentially modifiable contributors.

**Comment 3**

The recruitment process should be described in more operational detail. Please clarify how many physicians were contacted through each channel, whether duplicate responses were prevented, and whether IP checks, email verification, or other safeguards were used.

**Response**

We expanded Methods Section 2.2.

We now state that initial invitations were sent to 835 physicians through overlapping professional channels, including available mailing lists, institutional contacts at gastroenterology units, and national gastroenterology networks. Because these channels overlapped and the survey link could be forwarded to eligible colleagues, snowball sampling was permitted. For this reason, channel-specific unduplicated denominators and the final number of eligible recipients could not be reconstructed, and a formal response rate could not be calculated.

We also clarified safeguards and duplicate handling. To preserve anonymity, no personal identifiers were collected; IP address tracking, cookies, and email verification were not used. Submitted records were reviewed before analysis for consent status, eligibility, professional status, completeness, and exact duplicate response patterns, and no duplicate eligible record was retained in the final analytic dataset.

**Comment 4**

The manuscript should clarify whether fellows and specialists differed in exposure to on-call duty, endoscopy volume, and public-sector obligations. These differences may partly explain burnout patterns and should be considered in subgroup interpretation.

**Response**

We added exploratory comparisons across professional status groups.

In Methods Section 2.5, we state that additional exploratory comparisons were performed to describe whether fellows, specialists completing obligatory service, specialists who had completed obligatory service, and academic faculty differed in selected workload exposures, including monthly on-call burden, weekly endoscopy volume, out-of-hours endoscopy duty, and public sector practice.

In Results Section 3.6, we report that monthly on-call frequency varied across professional status groups, with fellows showing a higher median on-call frequency than specialists who had completed obligatory service. Public-sector practice also differed across groups, with the highest proportions observed among fellows and specialists completing obligatory service. Specialists completing obligatory service had endoscopy volumes comparable to fellows, while inadequate staff support was numerically more frequent in this group.

These subgroup analyses are presented in Supplementary Table 2 and are explicitly described as exploratory and contextual rather than causal.

**Comment 5**

The purpose-designed Endoscopy Workload Module is useful, but its development should be reported more transparently. The authors should provide the full questionnaire as supplementary material and describe how item wording was refined after pilot testing.

**Response**

We expanded the description of the Endoscopy Workload Module in Methods Section 2.3 and provided the study-specific questionnaire components in Supplementary Appendix 1.

We now describe that the module was developed through literature review, structured gastroenterologist discussion, qualitative expert review, and pilot testing with 10 gastroenterologists. We explicitly state that pilot feedback led to minor wording refinements to improve clarity, that no items were added or removed, and that pilot responses were excluded from the analytic dataset.

The Supplementary Material provides all Endoscopy Workload Module items, response formats, and analytic coding. The Copenhagen Burnout Inventory items are not reproduced in the supplement; the validated Turkish version was used and scored according to the original scoring instructions.

**Comment 6**

The analysis would be strengthened by distinguishing diagnostic, therapeutic, and advanced endoscopic procedures where possible. Procedure volume alone may not capture complexity, duration, or physical/cognitive load; this point is especially relevant when interpreting workload burden in therapeutic endoscopy (PMID: 39467805).

**Response**

We agree and have added this dimension.

In Methods Section 2.3, we clarified that the Endoscopy Workload Module included items on the proportion of therapeutic or advanced procedures and the proportion of emergency or unscheduled procedures. These variables were analysed individually and were not combined with total procedure volume into a composite workload score.

In the Discussion, we added that procedure volume alone may not fully capture the physical and cognitive workload of endoscopy, and that therapeutic and advanced procedures often involve longer duration, greater complexity, higher clinical stakes, and potentially greater procedural fatigue. We added the suggested reference, PMID: 39467805.

**Comment 7**

The choice of personal burnout as the main binary outcome is reasonable, but the rationale should be expanded. Since work-related burnout is also central to occupational exposure, the authors should briefly explain whether similar models were explored for other CBI domains.

**Response**

We expanded the rationale in Methods Section 2.4.

We now state that clinically relevant personal burnout was selected as the main binary outcome because it captures general exhaustion and was considered most directly aligned with the workload exposures under investigation. This choice was prespecified because personal burnout reflects overall exhaustion without requiring respondents to attribute exhaustion specifically to work or patient care, thereby reducing conceptual overlap between occupational exposures and the outcome definition.

We also clarified that work-related and patient-related burnout were analysed as secondary outcomes using continuous scores, group comparisons, correlation analyses, and exploratory domain-specific sensitivity models. In Results Section 3.5,

we added that exploratory models for work-related and patient-related burnout showed similar directions of effect for the main workload variables.

**Comment 8**

The logistic regression model is clinically informative, but more detail is needed regarding variable selection, collinearity assessment, and model calibration. Reporting variance inflation factors or another collinearity diagnostic would improve transparency.

**Response**

We revised the Statistical Analysis and Results sections to provide these details.

In Methods Section 2.5, we now provide the candidate variable selection procedure, VIF/tolerance criteria, Hosmer-Lemeshow calibration assessment, and discrimination assessment using the area under the receiver operating characteristic curve.

In Results Section 3.5, we now report that no candidate variable showed problematic multicollinearity and that all variance inflation factors were below 2.5. We also report that the Hosmer-Lemeshow goodness-of-fit test indicated acceptable calibration ($p = 0.67$), and that the area under the receiver operating characteristic curve was 0.84, indicating good discrimination.

**Comment 9**

The authors should avoid causal language throughout the Results and Discussion. The cross-sectional design supports associations, not proof that specific workload factors caused burnout.

**Response**

We agree and revised potentially causal wording throughout the manuscript.

Causal language was carefully reviewed and qualified throughout the manuscript. We now consistently frame the findings as associations or correlations, and we explicitly state that the cross-sectional design precludes causal inference. The Results and Discussion use association-based language, including “associated with,” “correlated with,” “may be linked to,” “may contribute to,” and “should be interpreted as associations rather than causal effects.” The Conclusion also states that the findings should not be interpreted as causal effects. Exploratory subgroup analyses are explicitly labelled as hypothesis-generating.

**Comment 10**

The Discussion appropriately emphasizes modifiable organizational factors, but the proposed interventions could be made more concrete. For example, the authors could discuss protected recovery time after emergency endoscopy, caps on consecutive on-call duties, better nursing support, and protected administrative time.

**Response**

We expanded the practical implications in the Discussion and Conclusion.

The revised manuscript now discusses specific organisational interventions, including protected recovery time after emergency endoscopy, limits on consecutive on-call duties, formal out-of-hours endoscopy pathways, improved nursing and ancillary staffing, protected administrative time for documentation and follow-up, streamlined documentation processes, and confidential, easily accessible psychological support services.

**Comment 11**

When discussing fellows and early-career physicians, the authors could connect their findings to broader evidence that procedural exposure, administrative workload, and perceived free time shape training experience in procedure-based specialties (PMID: 41681108).

**Response**

We added a dedicated paragraph in the Discussion addressing fellows and early-career physicians.

We now state that fellows' higher on-call frequency and greater exposure to public-sector obligations provide useful context for interpreting burnout patterns by career stage. We also added that early-career gastroenterologists may experience a distinct combination of procedural exposure, emergency responsibilities, administrative burden, and limited recovery time. This interpretation is connected to broader evidence from procedure-based training environments showing that procedural exposure, administrative workload, and perceived free time shape training experience and professional well-being.

We added the suggested reference, PMID: 41681108.

**Comment 12**

The limitations section is balanced, but one additional limitation should be added: physicians experiencing burnout may have been more motivated to complete the survey, which may inflate descriptive burnout estimates and strengthen perceived workload associations.

**Response**

We added this limitation.

The revised limitations section now states that physicians experiencing burnout may have been more motivated to complete the survey, which may have inflated the descriptive frequency of high burnout scores and may also have strengthened perceived associations between workload and burnout. We also mention this issue in the Discussion paragraph addressing self-selection and the interpretation of burnout proportions.

**Additional Revisions**

In addition to the reviewer-requested changes, we made several further revisions to improve clarity, consistency, and accuracy:

1. We added the three reviewer-suggested references: PMID 41399219, PMID 39467805, and PMID 41681108.
2. We updated Reference 4 to Hodkinson et al. 2022 to cite a more recent systematic review and meta-analysis on physician burnout, career engagement, and quality of patient care.
3. We updated Reference 22 to Antonelli et al. 2024 to cite a current ESGE/ESGENA position statement on training in basic gastrointestinal endoscopic procedures.
4. We reviewed and revised the reference list for accuracy, relevance, and consistency with the revised manuscript.
5. We standardized inequality expressions throughout the manuscript to avoid symbol-ordering problems during formatting.
6. We revised the tables and figure legends for consistency and readability.
7. We added Supplementary Material containing the Endoscopy Workload Module items and coding, sensitivity analyses, and exploratory workload comparisons by professional status.

We again thank the Editor and reviewers for their valuable comments. We believe that the revised manuscript is substantially improved in methodological transparency, interpretive caution, reporting clarity, and clinical relevance.

Respectfully,

Yavuz Özden, MD
On behalf of all authors.