We gratefully thank the editorial board member and all reviewers for their time spent making their constructive remarks and useful suggestions, which has significantly raised the quality of the manuscript and has enabled us to improve the manuscript. We have studied the comments carefully and have made correction point by point which we hope to meet with approval. The main corrections in the manuscript and the responses to the comments are as follows:

Reviewer #1:

1.This study is a single-center, prospective cohort study designed to evaluate the incidence, severity, and associated factors of postoperative nausea and vomiting (PONV) after laparoscopic sleeve gastrectomy. A total of 1,054 patients were enrolled, and multidimensional symptom assessment was performed using the Chinese version of the Index of Nausea, Vomiting, and Retching (INVR). The incidence of PONV was found to be as high as 76.3%. The study identified the protective effect of dexmedetomidine and the high-risk predictive value of an Apfel score of 3–4. With a rigorous study design, large sample size, and appropriate statistical methods, this study provides valuable clinical reference.

Response: We sincerely thank you for the positive and encouraging evaluation of our work. We greatly appreciate the acknowledgment of the study's rigorous design and clinical value.

2.Although the study has a large sample size, it is a single-center investigation, and the results may be influenced by local variations in anesthesia, surgery, nursing, and other clinical practices. It is recommended that this limitation be more clearly stated in the discussion, and multicenter validation is encouraged.

Response: We thank you for this thoughtful and constructive comment. We fully agree that the single-center design may limit the generalizability of our findings due to

potential variations in local clinical practices. In accordance with your suggestion, we have revised the Discussion section to state this limitation more explicitly and to emphasize the need for multicenter validation in future research. The revised text is provided below.

“First, as a single-center investigation, the findings may be influenced by institution-specific practices in anesthesia, surgery, and perioperative care. This could limit the generalizability of the results to other populations or clinical settings. Therefore, multicenter studies are warranted to validate our findings and assess the consistency of the observed associations across diverse healthcare environments. External validation of our predictive model in independent cohorts is also needed.”

3.The use of dexmedetomidine was non-randomized and determined by the attending anesthesiologist, which may introduce selection bias. Although multivariable adjustment was performed, confounding factors cannot be completely excluded. It is suggested that this issue be supplemented in the discussion.

Response: We thank you for this insightful methodological observation. You are absolutely correct that the non-randomized administration of dexmedetomidine, based on the attending anesthesiologist's discretion, introduces potential selection bias. Although we employed LASSO regression for variable selection and multivariable logistic regression to adjust for measured confounders, we acknowledge that residual confounding may still exist. In accordance with your suggestion, we have now added this as a distinct limitation in the Discussion section. The revised text is provided below.

“Third, the use of dexmedetomidine was non-randomized and determined according to the attending anesthesiologist's clinical judgment, which may

introduce selection bias. Although we adjusted for a comprehensive set of perioperative variables, the possibility of residual confounding cannot be entirely excluded. Therefore, the observed protective effect of dexmedetomidine should be interpreted with caution, and future randomized controlled trials are warranted to confirm these findings."

4.There is an inherent correlation between the Apfel score and the use of antiemetic medications. Because prophylactic strategies were based on the Apfel score, collinearity exists between these two factors, which may affect the interpretation of the independent effect of the score. Further elaboration of this limitation is recommended in the Methods or Discussion section.

Response: We thank you for this important observation. You are correct that because our prophylactic antiemetic strategy was risk-stratified based on the Apfel score, there is an inherent correlation between these two variables. While we briefly acknowledged this limitation in the original manuscript, we agree with your suggestion that it warrants more detailed elaboration. In accordance with your recommendation, we have now expanded the Limitations section to more fully address this issue. The revised text is provided below.

"Fifth, because prophylactic antiemetic regimens were administered according to the Apfel score, an inherent correlation exists between these two variables. This collinearity complicates interpretation of the independent effect of the Apfel score, which not only reflects baseline risk but also determines the intensity of prophylaxis received. Consequently, the observed positive association between Apfel scores and PONV risk may be underestimated."

5.INVR assessment may be subject to recall bias. Although repeated assessments were conducted, patient self-reporting may still be influenced by subjective memory.

This point should be mentioned in the discussion.

Response: We thank you for this thoughtful observation. You are correct that patient-reported outcomes such as the INVR are inherently subject to recall bias, as patients may differentially remember or report symptoms based on their subjective experience. While we briefly acknowledged this limitation in the original manuscript, we agree with your suggestion that it warrants more detailed elaboration. In accordance with your recommendation, we have now expanded the Limitations section to more fully address this issue. The revised text is provided below.

"Finally, patient-reported outcomes such as the INVR are subject to recall bias, as patients may have difficulty accurately recalling the timing and severity of symptoms, particularly for milder episodes. Although we conducted prospective assessments at seven predefined time points over 3 days to minimize the recall window and reduce this bias, some patients may have over- or under-reported their symptoms based on subjective memory. More severe episodes of nausea or vomiting may be more readily recalled, potentially leading to overestimation of symptom severity. Future studies could consider incorporating real-time symptom diaries or digital tracking tools to further mitigate recall bias."

6.Some data were not sufficiently presented. For example, although the QoR-15 score and postoperative complications were listed, they were not subjected to stratified analysis according to PONV severity. Supplementary analysis or explanation is recommended.

Response: We thank you for this thoughtful suggestion. We agree that stratified analysis by PONV severity could potentially provide additional insights. However, after careful consideration, we would like to explain why we did not perform such analyses and why we believe the current presentation is appropriate for the primary

objectives of this study.

First, the primary aim of this study was to quantify the incidence and severity of PONV using the INVR scale and to identify factors independently associated with the occurrence of PONV (defined as INVR total score > 0). The secondary outcomes, including QoR-15 scores and postoperative complications, were presented to provide a descriptive overview of recovery profiles between patients with and without PONV, rather than to establish dose-response relationships across severity grades.

Second, while the INVR scale captures symptom severity continuously, categorizing patients into multiple severity subgroups (e.g., mild, moderate, severe) would result in uneven group sizes and potentially insufficient statistical power for meaningful comparisons, particularly for rare outcomes such as specific complications. For example, only 8 postoperative complication events were documented in the entire cohort, which precludes any reliable stratified analysis by severity.

Third, the relationship between PONV severity and quality of recovery is complex and may be confounded by numerous factors not adjusted for in this descriptive analysis. A formal investigation of this dose-response relationship would require a different analytical approach, such as ordinal regression or propensity score methods, which was beyond the scope of the current study.

Nevertheless, we fully acknowledge that exploring the association between PONV severity and recovery outcomes represents an important avenue for future research. We appreciate your suggestion and will consider this direction in our subsequent studies.

We hope this explanation clarifies our rationale for the current presentation of the data.

7. The English expression is generally fluent and the structure is clear. The figures and

tables are legible with complete captions. The references are formatted standardly and cited appropriately.

Response: We thank you for your positive feedback on the language, structure, and presentation of our manuscript. We are pleased that you found the figures, tables, and references clear and appropriately formatted.

Reviewer #2:

In general

1. The manuscript is generally well-written, but there are places where a more direct, active voice would improve readability: - Use Active Voice. - Ensure you consistently use either "postoperative nausea and vomiting" or the abbreviation "PONV" after it's first defined. - Ensure consistent tense (some sections shift between past and present).

Response: We thank you for your thoughtful and constructive comments on the language and style of our manuscript. We have carefully reviewed the entire manuscript and made the following revisions to address your concerns:

Active Voice: We have revised sentences throughout the manuscript to use a more direct, active voice where appropriate.

Consistent Use of Terminology: We have ensured that "postoperative nausea and vomiting" is fully spelled out at first mention in the Abstract and Introduction, and thereafter consistently abbreviated as "PONV" throughout the manuscript.

Consistent Tense: We have standardized the tense usage throughout the manuscript, using past tense for methods and results, and present tense for conclusions and established knowledge.

2. Abbreviations should be fully written at first appearance.

Response: We thank you for this important observation. We have carefully reviewed the entire manuscript to ensure that all abbreviations are fully spelled out at their first

appearance.

In Title

3. The current title is good but does not reflect the large sample size (n=1054), which is a major strength, does not highlight the use of INVR, which is methodologically important, and does not signal the predictive analysis component. Consider a more informative title, for example: "Incidence, Severity, and Predictors of Postoperative Nausea and Vomiting After Laparoscopic Sleeve Gastrectomy: A Prospective Cohort Study Using the INVR Scale".

Response: Thank you for this excellent suggestion. We agree that the proposed title would more accurately reflect the key strengths of our study, including the quantification of incidence and severity, the identification of predictors, the large sample size, and the methodological importance of using a validated scale. However, during the final preparation of the manuscript, we noted that the journal limits titles to a maximum of 18 words and does not permit abbreviations. To comply with these formatting requirements, we have revised the title to: “Postoperative nausea and vomiting after laparoscopic sleeve gastrectomy: A prospective cohort study.” We believe this version remains concise and informative while adhering to the journal’s guidelines.

In the Abstract section

4. The abstract is well written and informative, but it is relatively long and dense.

Response: We thank you for this valuable feedback. We agree that the abstract can be made more concise while preserving all essential information. We have revised the abstract to improve readability by condensing redundant phrases and streamlining the presentation of methods and results.

5. You mention using LASSO and multivariable regression, which is great. For a

broader audience, you might add a brief phrase like "...to identify independent factors associated with PONV." This just makes the statistical goal a tiny bit clearer for readers who aren't statisticians.

Response: We thank you for this thoughtful suggestion. We agree that adding a brief clarifying phrase will help non-statistician readers better understand the purpose of our statistical approach. We have revised the Methods section of the Abstract accordingly. The revised text is provided below.

“Least absolute shrinkage and selection operator regression followed by multivariable logistic regression was used to identify factors independently associated with PONV.”

6. PONV is defined as INVR > 0 at any timepoint. This results in a very high cumulative incidence (76.3%).

Response: We thank you for this important observation. You are correct that defining PONV as any INVR score > 0 at any of the seven postoperative timepoints yields a high cumulative incidence. It is important to clarify that this definition aligns with the established use of the INVR scale, where any score greater than 0 indicates the presence of nausea, vomiting, or retching. This threshold was intentionally chosen to capture the full spectrum of PONV symptoms, including mild or transient episodes that are often missed in studies using binary outcome measures or shorter follow-up periods. As you noted, this approach results in a higher reported incidence but provides a more comprehensive and accurate picture of the true burden of PONV in this high-risk population.

We have already addressed this point in the Discussion section (second paragraph), where we explain that the high incidence observed in our cohort can be attributed to several factors, including the use of the INVR scale to capture the full symptom

spectrum and the prospective assessments at seven timepoints over three postoperative days. We believe this explanation appropriately contextualizes the high cumulative incidence for readers.

We appreciate your attention to this methodological detail, which reinforces the value of using validated, multidimensional instruments in PONV research.

7. The abstract should clarify that this represents any symptom over 3 days, not clinically significant PONV.

Response: We thank you for this important clarification. You are absolutely correct that our definition of PONV (any INVR score > 0 at any timepoint) captures the full spectrum of symptoms, including mild or transient episodes, rather than only clinically significant PONV requiring intervention. We have revised the Abstract to make this distinction clear to readers. The revised text is provided below.

“The overall incidence of any PONV symptoms over the 3-day postoperative period was 76.3% (804/1054).”

8. Reporting “Non-PONV as the outcome” may confuse readers.

9. It is more intuitive to model PONV as the outcome and describe risk factors directly.

10. The way you present the odds ratios for the Apfel score is a bit confusing. Saying an Apfel score of 3 is associated with "lower odds of being in the Non-PONV group" is technically correct, but it forces the reader to do a double-take. It would be much clearer to frame it from the perspective of the outcome we care about: PONV. Instead of: "Apfel scores of 3... were associated with lower odds of being in the Non-PONV group, reflecting increased PONV risk." You could simply say: "Apfel scores of 3 (OR for PONV: 5.0; 95%CI: 1.3-10.0) and 4 (OR for PONV: 5.0; 95%CI: 1.4-10.0) were strongly associated with an increased risk of PONV." …….. (an example)

Response: We thank you for these excellent suggestions. We completely agree that modeling PONV directly as the outcome is more intuitive and reader-friendly. Following your recommendation, we have re-performed the multivariable logistic regression analysis with PONV as the outcome variable. The revised results and all relevant sections of the manuscript (Abstract, Results, Discussion, and Figure 4) have been updated accordingly. The revised text is provided below.

"Multivariable analysis with PONV as the outcome revealed that intraoperative dexmedetomidine was associated with a significantly lower risk of PONV [odds ratio (OR), 0.50; 95% confidence interval (CI), 0.36–0.70; $P < 0.001$]. Apfel scores of 3 (OR, 4.27; 95% CI, 1.33–13.55; $P = 0.013$) and 4 (OR, 4.93; 95% CI, 1.51–15.94; $P = 0.007$) were associated with an increased risk."

"Multivariable logistic regression analysis with PONV as the outcome was performed (Figure 4). Compared with patients having an Apfel score of 0, those with scores of 3 (OR, 4.27; 95% CI, 1.33–13.55; $P = 0.013$) and 4 (OR, 4.93; 95% CI, 1.51–15.94; $P = 0.007$) had significantly increased odds of PONV. Intraoperative use of dexmedetomidine was associated with significantly lower odds of PONV (OR, 0.50; 95% CI, 0.36–0.70; $P < 0.001$)."

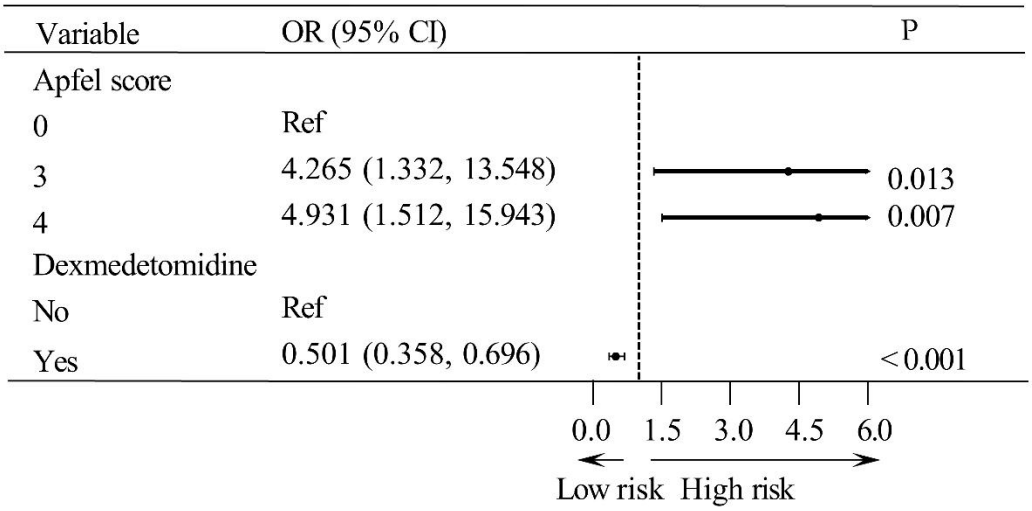

Figure 4 Forest plot of multivariable logistic regression analysis for factors associated with PONV. Multivariable logistic regression with PONV as the outcome. Apfel score (reference score = 0) and dexmedetomidine use (reference = no) were independently associated with PONV. Dots represent ORs; horizontal lines indicate 95% CIs. OR > 1 indicates increased risk of PONV; OR < 1 indicates decreased risk of PONV. CI: Confidence interval; OR: Odds ratio; PONV: Postoperative nausea and vomiting; Ref: Reference.

11. The conclusion states findings "support dexmedetomidine use." This is observational data. Causality cannot be claimed. Instead of: "These findings support dexmedetomidine use…" Use: "Dexmedetomidine use was independently associated with lower odds of PONV. Randomized trials are needed to confirm a causal protective effect."

Response: We thank you for this important methodological clarification. You are absolutely correct that causal claims cannot be made from observational data. We have revised the Conclusion section of the manuscript to reflect a more cautious interpretation, emphasizing the observed association and the need for randomized

trials to establish causality. The revised text is provided below.

“Intraoperative dexmedetomidine was independently associated with significantly lower odds of PONV, while Apfel scores of 3–4 predicted markedly higher risk even with guideline-directed prevention. These findings indicate that dexmedetomidine use is associated with a lower risk of PONV in the LSG population. However, given the observational nature of our study, randomized controlled trials are required to confirm a causal protective effect.”

In the Introduction section:

12. The introduction cites incidence ranges but does not clearly explain why LSG patients are uniquely high-risk (e.g., gastric manipulation, increased intra-abdominal pressure, hormonal changes). Add 2–3 mechanistic sentences to deepen the pathophysiology context.

Response: We thank you for this valuable suggestion. We agree that providing mechanistic context helps readers understand why LSG patients are particularly susceptible to PONV. We have added several sentences to the Introduction to elucidate the pathophysiological factors unique to this surgical population. The revised text is provided below.

“Patients undergoing LSG are uniquely predisposed to PONV due to several pathophysiological factors. The surgical procedure involves extensive gastric manipulation and resection, which may trigger vagal afferent stimulation and release of emetogenic mediators[3]. Additionally, removal of the gastric fundus and body reduces gastric compliance and increases intragastric pressure, promoting gastric distension and reflux that may trigger nausea and vomiting[3,4]. Furthermore, the rapid postoperative hormonal changes following gastrectomy, particularly alterations in ghrelin and other gut peptides, may

contribute to dysmotility and heightened emetic susceptibility[4].”

13. The gap is partially methodological (binary outcomes), but the predictive modeling aspect (LASSO + multivariable regression) is not emphasized enough.

Response: We thank you for this insightful observation. You are correct that while we highlighted the methodological limitation of binary outcomes, the predictive modeling component—specifically the use of LASSO for variable selection followed by multivariable regression—represents an important methodological strength of our study. We have revised the Introduction to better emphasize this aspect by incorporating it into the study objectives. The revised text is provided below.

“the application of least absolute shrinkage and selection operator (LASSO) regression followed by multivariable logistic regression to identify independent risk and protective factors”

14. Is your study novelty a large sample? Prospective design? Use of INVR? Or Machine-learning-assisted variable selection? This should be explicitly stated

Response: We thank you for this important comment. You are correct that the novelty of our study should be explicitly stated to help readers clearly understand its contributions. We have revised the end of the Introduction to concisely summarize the key novel aspects of our study. The revised text is provided below.

“To address this evidence gap, we conducted a large, prospective cohort study that incorporated the following key aspects: (1) the use of the validated INVR to provide a comprehensive, multidimensional assessment of PONV symptoms over the first 3 postoperative days; and (2) the application of least absolute shrinkage and selection operator (LASSO) regression followed by multivariable logistic regression to identify independent risk and protective factors, thereby providing procedure-specific, evidence-based insights to inform prophylactic strategies.”

In the “Epidemiological Evidence Section” section

16. No Sample Size Justification. Even in observational studies, a rationale should be provided. For example: expected PONV rate, number of events per variable for regression.

Response: We thank you for this important methodological comment. We acknowledge that a formal a priori sample size calculation was not performed, as this was an observational study employing consecutive sampling. However, we have now added a justification in the Statistical Analysis section demonstrating that the final sample size (n = 1,054) is adequate for the multivariable regression analysis. Specifically, the number of events (PONV cases, n = 804) and the number of candidate variables (9 variables entered into the final model) satisfy the commonly recommended criterion of at least 10 events per variable (EPV). With 804 events and 9 variables, the EPV ratio is approximately 89, which is well above the minimum threshold and ensures stable estimates and sufficient statistical power. The revised text is provided below.

“The final sample size of 1054 patients, with 804 PONV events, provided an events-per-variable ratio of approximately 89 for the nine variables entered into the multivariable logistic regression model. This exceeds the commonly recommended minimum of 10 events per variable[13], ensuring stable estimates and adequate statistical power.”

17. You mention a "standardized patient-controlled intravenous analgesia containing sufentanil and oxycodone." It would be helpful to add the drug doses or the PCA settings (e.g., bolus dose, lockout interval) for completeness. This allows for perfect replication.

Response: We thank you for this valuable suggestion. We agree that providing the full

details of the PCA regimen is important for study reproducibility. In response to your comment, we have revised the Perioperative Management section to include the complete composition of the PCA solution and the pump settings. The revised text is provided below. The revised text is provided below.

"The PCA solution consisted of sufentanil (100 μg), oxycodone (20 mg), and flurbiprofen axetil (300 mg) diluted to 100 mL with normal saline. The pump was programmed with a background infusion of 1 mL/h, an on-demand bolus of 3.5 mL, a lockout interval of 5 minutes, and a maximum volume of 10 mL/h."

18. A critical methodological issue concerns the definition of the primary outcome. The current definition classifies any instance of postoperative nausea, vomiting, or retching (PONV) as an INVR total score greater than zero at any postoperative assessment. This threshold means even minimal, clinically trivial nausea is counted as an event, which inflates the incidence rate and hinders comparison with other studies. To strengthen the interpretation of their results, the authors should justify why a score greater than zero was selected and provide additional analyses for moderate-to-severe PONV and vomiting-only outcomes.

Response: We thank you for this critical methodological comment. We fully agree that clarifying the rationale for our outcome definition and providing additional analyses to assess the robustness of our findings are essential for proper interpretation of the results. Regarding the definition, we chose INVR > 0 because this aligns with the validated use of the INVR scale, where any score greater than 0 indicates the presence of nausea, vomiting, or retching. Additionally, PONV is inherently a binary outcome (present vs. absent) in clinical practice and in major guidelines, and this definition ensures consistency with the majority of published studies, allowing for meaningful comparison. We deliberately selected this threshold to capture the full spectrum of

symptoms, including mild episodes that may still affect patient recovery but are often missed in studies using binary outcome measures.

In response to your suggestion, we have: (1) added this justification for the INVR > 0 definition in the Methods section; (2) performed additional multivariable logistic regression analyses for moderate-to-severe PONV and vomiting-only outcomes; and (3) incorporated these results into the manuscript as supplementary analyses. The revised text is provided below.

"PONV was defined as a total INVR score > 0 at any of the seven postoperative assessments, consistent with the validated use of the INVR where any score > 0 indicates the presence of nausea, vomiting, or retching[10,11]. This definition aligns with the standard clinical definition of PONV (presence *vs* absence) used in major guidelines and the majority of published studies[2-4,12]."

"Two sensitivity analyses were performed using alternative outcome definitions. The first analysis defined moderate-to-severe PONV as an INVR total score > 8, consistent with previous studies that used this threshold to identify clinically significant symptoms requiring intervention[14,15]. The second analysis defined vomiting only as a score > 0 on the vomiting dimension of the INVR at any of the seven postoperative assessments. The same LASSO regression variable selection process and multivariable logistic regression models were applied for these outcomes."

"Sensitivity analyses using moderate-to-severe PONV (INVR > 8) and vomiting only as outcomes confirmed that dexmedetomidine was associated with a lower risk of PONV (OR, 0.69; 95% CI, 0.52–0.90; *P* = 0.006 for moderate-to-severe PONV; OR, 0.68; 95% CI, 0.52–0.88; *P* = 0.004 for vomiting only). Detailed results are presented in Supplementary Tables 4 and 5."

19. The variable definitions are generally clear. However, the classification of dexmedetomidine exposure as “any administration” is problematic. It is unclear if the dose was standardized, if a maintenance infusion was used, or if the protocol was provider-dependent. This lack of specificity matters because it raises the possibility of confounding by indication.

Response: We thank you for this important methodological observation. You are absolutely correct that the definition of dexmedetomidine use requires greater specificity to avoid ambiguity and to address potential confounding by indication. We acknowledge that our original description was insufficient. In response to your comment, we have revised both the Variable Definitions and Perioperative Management sections to clarify the standardized dexmedetomidine protocol.

As described in the Perioperative Management section, dexmedetomidine was administered according to a standardized protocol: a loading dose of 0.5 μg/kg intravenously over 10 minutes upon arrival in the operating room, with no maintenance infusion. This protocol was uniformly applied to all patients who received dexmedetomidine. We have now explicitly stated this in the Variable Definitions section to ensure clarity.

In addition, we fully acknowledge that the decision to administer dexmedetomidine was based on the attending anesthesiologist's clinical judgment, which may introduce confounding by indication. This limitation has been explicitly addressed in the Limitation section, where we note that the observed protective effect of dexmedetomidine should be interpreted with caution and that future randomized controlled trials are warranted to confirm these findings. The revised text is provided below.

“Dexmedetomidine use was defined as administration of a standardized loading

dose of 0.5 μg/kg (no maintenance infusion), as documented in the anesthetic record."

"For patients receiving dexmedetomidine, a standardized loading dose of 0.5 μg/kg was administered intravenously over 10 minutes upon arrival in the operating room; no maintenance infusion was given. The decision to administer dexmedetomidine was at the discretion of the attending anesthesiologist."

"Third, the use of dexmedetomidine was non-randomized and determined according to the attending anesthesiologist's clinical judgment, which may introduce selection bias. Although we adjusted for a comprehensive set of perioperative variables, the possibility of residual confounding cannot be entirely excluded. Therefore, the observed protective effect of dexmedetomidine should be interpreted with caution, and future randomized controlled trials are warranted to confirm these findings."

20. A clear description of the antiemetic protocol is a strength of the perioperative management section. However, because antiemetics were administered based on the Apfel score, a structural collinearity exists between these two variables. While the manuscript acknowledges this issue, it should also describe how the collinearity was assessed, such as by reporting the variance inflation factor.

Response: We thank you for this important methodological suggestion. You are correct that while we acknowledged the collinearity between Apfel score and antiemetic use in the Limitations section, we did not describe how this collinearity was assessed. In accordance with your recommendation, we have now added the variance inflation factor (VIF) results for all variables included in the final multivariable logistic regression model. The results demonstrate that all VIF values were below 5, indicating no severe multicollinearity among the predictors. We have

added this information to both the Statistical Analysis and Results sections as described below.

"Multicollinearity was assessed using the variance inflation factor (VIF), with VIF < 5 considered acceptable."

"All VIF values were below 5 (range: 1.01–1.92), indicating no significant multicollinearity among the predictors."

In the "Results" section

22. The reported 76.3% incidence rate appears extremely high because it captures any nausea across multiple time points, a cumulative burden that should be emphasized.

Response: We thank you for this important observation. You are absolutely correct that the 76.3% incidence reflects the cumulative burden of any PONV symptoms captured across seven timepoints over three postoperative days, rather than the incidence at a single timepoint. To make this distinction clearer to readers, we have revised the Results section to explicitly emphasize the cumulative nature of this incidence estimate. The revised text is provided below.

"The cumulative incidence of PONV over the first 3 postoperative days, defined as any INVR score > 0 at any of the seven assessment timepoints, was 76.3% (804/1054)."

23. A summary table quantifying the distribution of PONV severity over the three days would clarify how many patients experienced clinically meaningful symptoms.

Response: We thank you for this helpful suggestion. As you noted, quantifying the distribution of PONV severity over the three postoperative days is important for understanding the proportion of patients who experienced clinically meaningful symptoms. We would like to note that this information has already been provided in the Supplementary Materials. Specifically, Supplementary Tables 1-3 detail the

distribution of severity grades (mild, moderate, severe) for nausea, vomiting, and retching across each assessment timepoint. These tables allow readers to identify not only the cumulative incidence but also the proportion of patients who experienced moderate-to-severe symptoms at each timepoint.

For clarity, we have added a brief reference to these tables in the Results section to highlight their relevance to the assessment of clinically meaningful symptoms. The revised text is provided below.

"The distribution of symptom severity grades over time is detailed in the Supplementary Tables 1–3, which provide a comprehensive overview of the proportion of patients experiencing mild, moderate, and severe symptoms at each assessment timepoint."

24. While QoR-15 scores differed between groups, the manuscript should interpret whether this difference meets the minimal clinically important difference to determine its relevance for recovery.

Response: We thank you for this thoughtful suggestion. We agree that interpreting the observed differences in QoR-15 scores in the context of the minimal clinically important difference (MCID) adds valuable clinical context to our findings. As you suggested, we have added a brief discussion of this point in the Discussion section. The MCID for the QoR-15 scale has been reported to be approximately 6 points. In our study, the median difference in QoR-15 scores between the PONV and Non-PONV groups was 9 points at 24 hours postoperatively, which exceeds this threshold, suggesting that the difference may be clinically meaningful. However, we note that this difference decreased to 3 points at 48 hours. Given that QoR-15 was a secondary outcome and formal statistical comparisons were not performed between severity subgroups, these findings should be interpreted with caution and warrant

confirmation in future studies. The revised text is provided below.

"The QoR-15 scores differed between patients with and without PONV, with a median difference of 9 and 3 points at 24 and 48 hours postoperatively, respectively. The minimal clinically important difference for the QoR-15 scale is approximately 6 points[28]. The 24-hour difference exceeding this threshold suggests that PONV may have a clinically meaningful impact on early postoperative recovery. However, the QoR-15 score was a secondary outcome, and formal statistical comparisons were not performed across PONV severity subgroups. Therefore, these findings should be interpreted with caution and warrant confirmation in future studies."

25. The analysis of complications is underpowered with only eight events, and any suggested relationship with PONV should be avoided.

Response: We thank you for this important methodological observation. You are absolutely correct that with only eight postoperative complication events, the analysis of complications is underpowered, and any attempt to infer a relationship with PONV would be inappropriate. In response to your comment, we have revised the Results section to simply describe the events without suggesting any association. The revised text is provided below.

"Postoperative recovery outcomes are summarized in Table 2. Among the 1054 patients, 8 postoperative complication events were documented: 3 cases of suspected intra-abdominal infection, 4 cases of respiratory infection, and 1 reoperation."

In the "Discussion" section

26. Discussion attributes it to a multidimensional scale. Should also admit that >0 threshold may overestimate clinically relevant PONV.

Response: We thank you for this balanced and constructive comment. You are correct that while we highlighted the comprehensive nature of the INVR scale as a strength, we should also acknowledge the potential limitation that using a >0 threshold may overestimate clinically relevant PONV. In response to your suggestion, we have added a brief statement in the Discussion to address this point, providing a more balanced interpretation of the high incidence observed in our study. The revised text is provided below.

"Our definition of PONV (INVR score > 0) follows the validated use of the scale[10,11] and was deliberately chosen to avoid underreporting mild but clinically relevant symptoms that may affect patient recovery, while also ensuring consistency with the standard clinical definition of PONV[12]. However, we acknowledge that this threshold may overestimate clinically significant PONV, as even minimal symptoms are counted as events."

27. Apfel Score Section, explanation of collinearity is Good. However, emphasize that risk-adapted prophylaxis did not eliminate high-risk burden — this is clinically important.

Response: We thank you for this valuable suggestion. We agree that the clinical significance of our finding—that risk-adapted prophylaxis did not eliminate the high-risk burden—deserves greater emphasis. In response to your comment, we have revised the Discussion section to more explicitly highlight this clinically important point. The revised text is provided below.

"Clinically, this finding underscores that even guideline-directed, risk-adapted antiemetic prophylaxis does not eliminate the high-risk burden in this population. This highlights the need for continued optimization of antiemetic strategies in this high-risk subgroup, possibly through the addition of other pharmacological

agents or the use of non-pharmacological interventions, such as transcutaneous electrical acupoint stimulation, among others[24,25]."

28. Limitations Section Needs Expansion, Add: No external validation, No sensitivity analysis with different PONV thresholds, Potential residual confounding, and Lack of anxiety/pain control metrics.

Response: We thank you for this thorough and constructive suggestion. We agree that expanding the Limitations section to include these additional points will provide readers with a more complete understanding of the study's methodological considerations. In response to your comment, we have added the following limitations to the Limitation section. The revised text is provided below.

"This study has several limitations. First, as a single-center investigation, the findings may be influenced by institution-specific practices in anesthesia, surgery, and perioperative care. This could limit the generalizability of the results to other populations or clinical settings. Therefore, multicenter studies are warranted to validate our findings and assess the consistency of the observed associations across diverse healthcare environments. External validation of our predictive model in independent cohorts is also needed. Second, despite LASSO regression and multivariable adjustment, unmeasured confounders (*e.g.*, genetic polymorphisms and anxiety levels) may have influenced the results. Residual confounding may still be present, as we did not collect detailed preoperative anxiety scores or intraoperative pain control metrics, both of which could affect PONV risk."

"Sixth, while we performed sensitivity analyses using moderate-to-severe PONV and vomiting-only outcomes to assess robustness, we did not systematically explore all possible PONV thresholds."

We would like to once again express our sincere gratitude to the editorial board member and the reviewers for their continued engagement with our manuscript. We greatly appreciate the additional time and expertise that have been devoted to providing further constructive comments and suggestions during this second round of review. These thoughtful observations have helped us to refine the manuscript and strengthen the clarity and rigor of our work.

We have carefully studied each of the new points raised by both reviewers and have made corresponding revisions to the manuscript. All changes are detailed in the point-by-point responses below. We hope that these further improvements will meet with your approval.

Reviewer 1:

1. The manuscript is generally clear and well written. A final light edit to simplify a few longer or more formal sentences could further improve readability.

Response: We sincerely thank you for your encouraging comment and for recognizing the overall clarity of our manuscript. We appreciate your suggestion to further improve readability by simplifying a few longer or more formal sentences. As noted, the manuscript had already been professionally edited by a language polishing service recommended by the journal prior to this submission. However, in response to your specific guidance, we have carefully re-reviewed the entire text and identified several sentences that could be revised to enhance conciseness and flow without altering scientific content. Representative modifications are detailed below. We hope these final refinements meet with your approval.

Revised text:

“Laparoscopic sleeve gastrectomy (LSG) is among the most common metabolic and bariatric surgeries worldwide[1]. Yet postoperative nausea and

vomiting (PONV) remains a major barrier to recovery, affecting up to two-thirds of patients despite routine multimodal prophylaxis[2]."

"Smoking history was defined as any regular tobacco use documented in the medical record. PONV history and motion sickness history were self-reported during the preoperative interview, defined respectively as prior postoperative nausea/vomiting and prior travel-related dizziness, nausea, or vomiting."

"The strong association between high Apfel scores (3 or 4) and PONV risk is striking, given that these patients already received intensified prophylaxis (three or four antiemetics). Their persistently elevated risk suggests that current protocols do not fully overcome inherent patient-related factors."

2. Model performance reporting: the manuscript still lacks information on model performance (e.g., discrimination such as AUC/C-statistic, and calibration). Including these metrics would help readers better understand the predictive value and robustness of the model.

Response: We sincerely appreciate your attention to the methodological details of our analysis. We fully understand your suggestion regarding model performance metrics such as AUC/C-statistic and calibration. These are indeed essential elements when reporting the development and validation of a prediction model.

However, we would like to respectfully clarify that the primary analytical goal of this study was not to develop a clinical prediction model, but rather to identify factors independently associated with PONV in this specific surgical cohort. In the context of an etiological or associational study, the focus is on the magnitude and direction of associations (odds ratios and confidence intervals) rather than on the predictive accuracy or discriminative ability of the multivariable model. We employed LASSO regression for variable selection and subsequent multivariable logistic regression for

inference, not for prediction.

Because we did not perform internal validation, bootstrap resampling, or develop a nomogram/scoring system intended for future clinical use, reporting metrics such as the C-statistic or calibration plot would be methodologically inappropriate and potentially misleading. Such metrics evaluate how well a model predicts outcomes in new data; our analysis was designed to explain relationships within the observed data, controlling for confounders.

Nonetheless, we recognize that the term "model" in the "Statistical analysis" section may have inadvertently created this expectation. To prevent future misunderstanding, we have revised the wording in the Methods, Results, and Discussion sections to emphasize the associational nature of our analysis, replacing "model" with more precise terminology.

Revised text:

“The final sample size of 1054 patients, with 804 PONV events, provided an events-per-variable ratio of approximately 89 for the nine variables entered into the multivariable logistic regression analysis.”

“All candidate variables listed in Table 1 were entered into the LASSO regression”

“The same LASSO regression variable selection process and multivariable logistic regression analyses were applied for these outcomes.”

“nine candidate variables were identified and subsequently entered into the multivariable logistic regression analysis”

“Traditional risk factors such as female sex, prior PONV history, and motion sickness history were not independent predictors in our final multivariable analysis”

“When the Apfel score itself is included in the analysis”

“This interpretation is supported by the fact that the number of antiemetic types administered was also considered in the LASSO selection but did not remain significant in the final multivariable regression”

“External validation of these associations in independent cohorts is also needed.”

3.Handling of missing data: The manuscript mentions exclusion of cases with missing essential data, but it would be helpful to report:

- the number of excluded patients due to missing data, and
- whether any assessment of missingness was performed.

A brief clarification would improve transparency.

Response: We thank you for this suggestion to enhance the transparency of our data handling procedures. We agree that providing explicit details on missing data, even when the amount is minimal, improves the completeness of the report.

As shown in Figure 1, a total of 1181 patients were screened for eligibility. Among the exclusions, 5 patients (0.4%) were excluded due to missing data. Given the very low proportion of missing data (< 1%) and the large overall sample size (n = 1054), we did not perform formal statistical assessments of missingness patterns or imputation. In such circumstances, a complete-case analysis is widely accepted and is unlikely to introduce meaningful bias.

In accordance with your recommendation, we have added a brief statement in the Results section to explicitly report the number of patients excluded for missing data and to note the negligible proportion.

Revised text:

“Among the 127 excluded patients, 5 (0.4% of those screened) were excluded due to missing data; given the negligible proportion, no formal imputation or

missingness assessment was performed. ”

4.The authors have provided a reasonable justification for using INVR > 0 and have strengthened the manuscript by adding sensitivity analyses. However, the statement that this definition aligns with “standard clinical definitions" may be somewhat overstated. A slightly more cautious wording would improve accuracy and avoid potential confusion with studies focusing on clinically significant PONV.

Response: We thank you for this thoughtful observation. You are absolutely correct that our original phrasing—"aligns with the standard clinical definition of PONV"—could be interpreted as overstating the equivalence between an INVR score > 0 and the binary clinical assessment typically used in large outcome studies. While both approaches aim to capture the presence vs absence of PONV, the INVR is inherently more sensitive to mild symptoms. To avoid any potential confusion and to reflect a more cautious and accurate stance, we have revised the relevant sentence in the Methods section. In the Discussion section, we have removed a redundant restatement of this point, as the subsequent sentence already acknowledges the potential overestimation of clinically significant PONV.

Revised text:

“This definition follows the validated use of the INVR scale and is consistent with the binary framework (presence *vs* absence) commonly employed in clinical PONV research, although the INVR may detect milder symptoms than those captured in some traditional assessments[2-4,12].”

“Our definition of PONV (INVR score > 0) follows the validated use of the scale[10,11] and was deliberately chosen to avoid underreporting mild but clinically relevant symptoms that may affect patient recovery. However, we acknowledge that this threshold may overestimate clinically significant PONV, as even minimal

symptoms are counted as events."

5.The inclusion of detailed PONV severitydata in the supplementary materials is valuable. However, adding a brief summary (e.g., proportion of patients with moderate-to-severe symptoms) in the main text would enhance readability and clinical interpretability.

Response: We thank you for this constructive suggestion to improve the accessibility of our severity data. We agree that while the detailed tables in the supplementary material provide a comprehensive record, a concise summary in the main text would help readers quickly grasp the clinical burden of moderate-to-severe symptoms.

In response, we have added a brief summary to the Results section (under "Secondary outcomes"). This addition reports the peak proportions of patients experiencing grade III–IV (moderate-to-severe) nausea, vomiting, and retching across the study period.

Revised text:

"The distribution of symptom severity grades over time is detailed in the Supplementary Tables 1–3. In summary, the peak proportions of patients experiencing moderate-to-severe (grade III – IV) symptoms were: nausea 16.9%, vomiting 5.9%, and retching 5.9% — all occurring on the morning of postoperative day 1 (POD 1a). The daily incidence of each symptom dimension is visualized in Figure 3."

Reviewer 2:

1. The definition of the primary outcome may overestimate clinical significance. PONV was defined as any assessment with an INVR > 0, meaning even mild nausea was counted as an event, resulting in an incidence as high as 76.3%. The authors have acknowledged in the Discussion that this threshold may overestimate clinical significance and used INVR > 8 as the threshold for moderate-to-severe PONV in a sensitivity analysis. It is recommended to further clarify the incidence rates of "any

symptom" versus "clinically significant symptoms" in the Abstract or Results.

Response: We thank you for this valuable recommendation. We agree that explicitly reporting both the incidence of any PONV symptoms (INVR > 0) and that of moderate-to-severe PONV (INVR > 8) in the main text will improve clarity and clinical interpretability. In response, we have added the incidence of moderate-to-severe PONV to both the Abstract and the Results section.

Revised text:

“A total of 1054 patients were included. The overall incidence of any PONV symptoms over the 3-day postoperative period was 76.3% (804/1054); the incidence of moderate-to-severe PONV (INVR > 8) was 34.3% (362/1054).”

“The cumulative incidence of PONV over the first 3 postoperative days, defined as any INVR score > 0 at any of the seven assessment timepoints, was 76.3% (804/1054). The incidence of moderate-to-severe PONV, defined as an INVR score > 8, was 34.3% (362/1054).”

2. The use of dexmedetomidine was non-randomized, introducing selection bias. The use of dexmedetomidine was determined by anesthesiologists and not randomized. The authors have clearly listed this limitation in the Discussion and suggested future RCTs. This is appropriate. It is recommended to re-emphasize the "associative rather than causal nature" in the Conclusion.

Response: We thank you for this helpful suggestion. You are correct that, given the observational design and non-randomized use of dexmedetomidine, our findings reflect an association rather than a causal relationship. While we already acknowledged this limitation in the Discussion and called for future randomized controlled trials, we agree that the Conclusion should more explicitly reinforce the associative nature of the observed relationship.

In response, we have revised the Conclusion section to explicitly state the non-causal nature of the association and to avoid any language that might inadvertently imply causality, even when discussing future research directions.

Revised text:

“Intraoperative dexmedetomidine was independently associated with significantly lower odds of PONV, while Apfel scores of 3–4 predicted markedly higher risk even with guideline-directed prevention. It is important to emphasize that the observed relationship between dexmedetomidine use and reduced PONV risk is associative, not causal. Given the observational nature of our study, randomized controlled trials are required to determine whether this association reflects a true protective effect.”

3. Collinearity exists between the Apfel score and the antiemetic regimen. The antiemetic regimen was implemented in a stratified manner based on the Apfel score, leading to a high correlation between the two. The authors have reported a VIF < 5, indicating no serious collinearity. However, it is recommended to add a statement in the Methods that even if collinearity is not severe, it may still underestimate the independent effect of the Apfel score.

Response: We thank you for this insightful comment. We fully agree with your assessment: despite a VIF < 5 indicating acceptable collinearity for model stability, the inherent correlation between Apfel score and the number of prophylactic antiemetics (due to risk-stratified administration) may still attenuate the estimated effect size of the Apfel score.

We have therefore added a clarifying statement to address this nuance. Recognizing that this point pertains to the interpretation of findings and study limitations rather than the statistical methods per se, we have placed this addition in the Discussion section where the collinearity issue is already being addressed.

Revised text:

“Fifth, because prophylactic antiemetic regimens were administered according to the Apfel score, an inherent correlation exists between these two variables. This collinearity complicates interpretation of the independent effect of the Apfel score, which not only reflects baseline risk but also determines the intensity of prophylaxis received. Although the variance inflation factor was below 5, indicating no severe collinearity that would destabilize the model, this correlation may still attenuate—and thus lead to underestimation of—the observed association between Apfel score and PONV risk.”

4. Interpretation of the clinical significance of QoR-15 differences requires more caution. The difference in QoR-15 at 24 hours was 9 points, exceeding the MCID (6 points), but was only 3 points at 48 hours. The authors have noted in the Discussion that this is a secondary outcome, no formal statistical comparison was performed, and the interpretation is cautious. It is recommended to clarify whether statistical testing was performed for the 48-hour difference; if not, this should be explicitly stated.

Response: We thank you for this important clarification request. You are correct that the manuscript did not explicitly state whether formal statistical testing was performed for the QoR-15 differences at either 24 or 48 hours. To address this, we have revised the relevant passage in the Discussion to clearly state that no formal statistical comparisons were conducted for the QoR-15 scores at any time point. The 9-point and 3-point differences are presented solely as descriptive observations, and we emphasize that these should be interpreted with caution.

Revised text:

“The QoR-15 scores differed between patients with and without PONV, with a median difference of 9 and 3 points at 24 and 48 hours postoperatively, respectively. No

formal statistical testing was performed for these comparisons. The minimal clinically important difference for the QoR-15 scale is approximately 6 points[28]. The 24-hour difference exceeding this threshold suggests that PONV may have a clinically meaningful impact on early postoperative recovery, whereas the 3-point difference at 48 hours is below the MCID. Given that QoR-15 was a secondary outcome, these descriptive findings should be interpreted with caution and warrant confirmation in future studies."

5. Add the incidence of moderate-to-severe PONV in the Abstract or Results.

Response: We thank you for this suggestion. This comment aligns with your earlier recommendation in Comment 1. As detailed in our response to Comment 1, we have added the incidence of moderate-to-severe PONV (INVR > 8) to both the Abstract and the Results section. Specifically, we now report that moderate-to-severe PONV occurred in 34.3% (362/1054) of patients. We appreciate your attention to consistency and thank you again for your thoughtful review.

Revised text:

"A total of 1054 patients were included. The overall incidence of any PONV symptoms over the 3-day postoperative period was 76.3% (804/1054); the incidence of moderate-to-severe PONV (INVR > 8) was 34.3% (362/1054)."

"The cumulative incidence of PONV over the first 3 postoperative days, defined as any INVR score > 0 at any of the seven assessment timepoints, was 76.3% (804/1054). The incidence of moderate-to-severe PONV, defined as an INVR score > 8, was 34.3% (362/1054)."