Response to Reviewer 1

This manuscript presents a comprehensive, well-structured, and highly relevant review of the current landscape and future directions of multimodal artificial intelligence (AI) in the precision diagnosis and treatment of gastrointestinal and hepatobiliary diseases. The authors successfully integrate technical depth with clinical applicability, offering a holistic perspective on how multimodal AI—including endoscopy, radiology, pathology, genomics, proteomics, and natural language processing (NLP)—can transform clinical workflows and improve patient outcomes. The manuscript's strengths are numerous. It demonstrates exceptional breadth, covering a wide range of AI modalities and clinical domains, while also maintaining clarity and depth in its explanations. Each section is thoughtfully developed and clearly written, with well-chosen examples from recent literature that effectively illustrate the state-of-the-art. The inclusion of specific model types (e.g., Vision Transformers, graph neural networks, LSTM), application scenarios (e.g., immunotherapy prediction, liver transplantation, recurrence monitoring), and integration strategies (e.g., federated learning, digital twins) shows the authors' in-depth understanding of both AI and medical practice. Moreover, the authors go beyond merely summarizing existing studies. They provide a critical synthesis that highlights not only the technical progress but also real-world barriers—such as data heterogeneity, model interpretability, and ethical or regulatory issues. The proposed solutions, including the development of cross-institutional data-sharing platforms, interpretable AI frameworks, and dynamic clinical validation systems, are both timely and pragmatic. The manuscript is also notable for its multidisciplinary accessibility. It avoids excessive technical jargon while maintaining academic rigor, making it suitable for readers ranging from medical professionals to data scientists. The figures are useful and clearly complement the text. The writing is fluent, concise, and well-polished, with minimal if any language issues. I did not identify any major or minor issues requiring revision. This article makes a valuable contribution to the field and will serve as a helpful reference for researchers and clinicians seeking to understand and implement AI solutions in gastroenterology and hepatology. Recommendation: Accept without revision.

Response: We sincerely appreciate your thorough and insightful review! Your detailed recognition of our manuscript's structure, content depth, and interdisciplinary value is incredibly encouraging. We're thrilled to hear that our work on multimodal AI in gastrointestinal and hepatobiliary disease diagnosis and treatment resonates with your expertise.Your positive feedback on the technical-clinical integration, real-world barrier analysis, and accessible writing style motivates us to keep pushing boundaries. We'll remain committed to refining our research, addressing challenges like data heterogeneity and model interpretability, and bridging AI with clinical practice—efforts we hope will contribute meaningfully to the field.Thank you again for your time, support, and constructive insights. We're honored by your recommendation and look forward to continuing contributions that aid researchers and clinicians alike.

Response to Reviewer 2

1. "...lacks critical assessment of the limitations within cited studies... inherent biases and limitations associated with such data sources... overfitting risks and lack of ethnic diversity... constrain the generalizability..." "...reads like a promotional document... high optimism... real-world implementation and limitations are often understated... insufficient discussion of false positive rates, patient anxiety, workflow burden, or the costs of implementation... A more balanced view...".

Response:We sincerely thank the reviewer for their insightful and constructive comments, which have helped us strengthen the critical assessment within our manuscript.

Added Text:

1."Cui et al. (2022) developed a multi-center deep learning-based radiomics nomogram (DLRN) that in their study integrated CT imaging features and clinical parameters to predict the response of locally advanced gastric cancer

(LAGC) patients to neoadjuvant chemotherapy (NACT) (AUC 0.829 in internal validation) and correlated with disease-free survival (DFS). This approach suggests a potential basis for individualized treatment decisions. However, this was a retrospective study, and its performance in external cohorts (AUCs 0.804, 0.827), while still good, highlights the challenge of generalizability beyond the specific populations and imaging protocols used"(Setion:2.1.2).

2."However, this meta-analysis identified significant heterogeneity among the radiomics studies and noted a scarcity of deep learning evidence at the time. It also highlighted the risk of bias and methodological limitations (e.g., low Radiomics Quality Scores) in the underlying studies, emphasizing the need for cautious interpretation of these promising results."(Setion:2.1.2).

3."Similar to gastric cancer studies, these CRC investigations are predominantly retrospective and face challenges regarding generalizability across different populations and healthcare settings. The performance observed in validation cohorts (e.g., AUC 0.76 in Wang et al.) can be lower than in training sets, underscoring overfitting risks and the need for rigorous external validation."(Setion:2.1.2).

---

2. "...frequently uses highly technical jargon and acronyms... without adequate explanation... Definitions and brief explanations upon first use would enhance clarity."

Response:We thank the reviewer for pointing out the need for greater clarity regarding technical terms and acronyms. In response, we have created a comprehensive Glossary of Key Terms.

Added Table:

Abbreviation	Full Term	Definition
AI	Artificial Intelligence	Computer systems performing tasks requiring

		human-like intelligence (e.g., medical image analysis).
CCIM	Colorectal Cancer Immune Module	Transcriptomic signature predicting immune activity and prognosis in colorectal cancer.
CEUS	Contrast-Enhanced Ultrasound	Ultrasound technique using microbubble contrast agents to visualize tissue perfusion dynamics.
CNN	Convolutional Neural Network	Deep learning architecture specialized in image recognition and feature extraction.
CT	Computed Tomography	Cross-sectional X-ray imaging for structural assessment of abdominal organs.
ctDNA	Circulating Tumor DNA	Tumor-derived cell-free DNA in blood; biomarker for early cancer detection and monitoring.
DFS	Disease-Free Survival	Duration after treatment without clinical or radiological evidence of disease recurrence.
DLRN	Learning-based Radiomics Nomogram	Prognostic model integrating deep learning radiomic features with clinical variables.
DSA	Donor-Specific Antibodies	Antibodies targeting donor HLA antigens; biomarker for transplant rejection risk.
EGC	Early Gastric Cancer	Gastric carcinoma limited to mucosa/submucosa without lymph node involvement.
FL	Federated Learning	Privacy-preserving ML framework training models across decentralized data sources.
GAN	Generative Adversarial Networks	AI system generating synthetic data via adversarial training of generator/discriminator networks.
HLA	Human Leukocyte Antigen	Major histocompatibility complex (MHC) proteins governing immune recognition.
LAGC	Locally Advanced Gastric Cancer	Gastric tumors invading beyond submucosa (T2-T4) or with regional lymph node metastasis.
LNM	Lymph Node Metastasis	Metastatic dissemination of cancer cells to regional lymph nodes.
EGC	Early Gastric Cancer	Gastric carcinoma limited to mucosa/submucosa without lymph node involvement.
FL	Federated Learning	Privacy-preserving ML framework training models across decentralized data sources.
GAN	Generative Adversarial Networks	AI system generating synthetic data via adversarial training of generator/discriminator networks.
HLA	Human Leukocyte	Major histocompatibility complex (MHC)

	Antigen	proteins governing immune recognition.
LAGC	Locally Advanced Gastric Cancer	Gastric tumors invading beyond submucosa (T2-T4) or with regional lymph node metastasis.
LNM	Lymph Node Metastasis	Metastatic dissemination of cancer cells to regional lymph nodes.
MAF	Mutation Allele Frequency	Proportion of DNA fragments harboring a somatic mutation in liquid biopsies.
MRI	Magnetic Resonance Imaging	Non-ionizing imaging using magnetic fields for high-resolution soft-tissue characterization (e.g., liver fibrosis).
MSI	Microsatellite Instability	Hypermutation phenotype from defective DNA mismatch repair; predictor of immunotherapy response.
NACT	Neoadjuvant Chemotherapy	Cytotoxic therapy administered before curative-intent surgery.
NER	Neutrophil-to-Eosinophil Ratio	Hematologic ratio reflecting inflammatory status in gastrointestinal disorders.
NLP	Natural Language Processing	Computational analysis of unstructured clinical text (e.g., pathology reports).
NPV	Negative Predictive Value	Probability that a negative diagnostic test result indicates true absence of disease.
LAGC	Locally Advanced Gastric Cancer	Gastric tumors invading beyond submucosa (T2-T4) or with regional lymph node metastasis.
LNM	Lymph Node Metastasis	Metastatic dissemination of cancer cells to regional lymph nodes.
MAF	Mutation Allele Frequency	Proportion of DNA fragments harboring a somatic mutation in liquid biopsies.
MRI	Magnetic Resonance Imaging	Non-ionizing imaging using magnetic fields for high-resolution soft-tissue characterization (e.g., liver fibrosis).
MSI	Microsatellite Instability	Hypermutation phenotype from defective DNA mismatch repair; predictor of immunotherapy response.
NACT	Neoadjuvant Chemotherapy	Cytotoxic therapy administered before curative-intent surgery.
NER	Neutrophil-to-Eosinophil Ratio	Hematologic ratio reflecting inflammatory status in gastrointestinal disorders.
NLP	Natural Language Processing	Computational analysis of unstructured clinical text (e.g., pathology reports).
NPV	Negative Predictive Value	Probability that a negative diagnostic test result indicates true absence of disease.
NRG	Non-apoptosis	Alternative cell death pathways (e.g.,

	Regulated Cell Death	necroptosis) implicated in therapy resistance.
SWE	Shear Wave Elastography	Quantitative ultrasound technique measuring tissue stiffness via shear wave propagation.
LNM	Lymph Node Metastasis	Metastatic dissemination of cancer cells to regional lymph nodes.
MAF	Mutation Allele Frequency	Proportion of DNA fragments harboring a somatic mutation in liquid biopsies.
MRI	Magnetic Resonance Imaging	Non-ionizing imaging using magnetic fields for high-resolution soft-tissue characterization (e.g., liver fibrosis).
MSI	Microsatellite Instability	Hypermutation phenotype from defective DNA mismatch repair; predictor of immunotherapy response.
NACT	Neoadjuvant Chemotherapy	Cytotoxic therapy administered before curative-intent surgery.
NER	Neutrophil-to-Eosinophil Ratio	Hematologic ratio reflecting inflammatory status in gastrointestinal disorders.
NLP	Natural Language Processing	Computational analysis of unstructured clinical text (e.g., pathology reports).
NPV	Negative Predictive Value	Probability that a negative diagnostic test result indicates true absence of disease.
NRG	Non-apoptosis Regulated Cell Death	Alternative cell death pathways (e.g., necroptosis) implicated in therapy resistance.
SWE	Shear Wave Elastography	Quantitative ultrasound technique measuring tissue stiffness via shear wave propagation.
TACS	Tumor-Associated Collagen Signatures	Peritumoral collagen alignment patterns correlating with invasive potential.
TLS	Tertiary Lymphoid Structures	Ectopic lymph node-like structures in tumors; surrogate for active anti-tumor immunity.
TTB	Total Tumor Burden	Volumetric summation of all tumor lesions in serial imaging assessments.
WSI	Whole-Section Images	Digitized high-resolution scans of complete histopathological tissue sections.

3.“...does not sufficiently address regulatory and ethical implications... lacks depth regarding data privacy laws, liability... frameworks for explainability... omission is significant.”.

Response:Thank you for highlighting these critical gaps. We have comprehensively revised Section 4.4.1 to integrate regulatory/ethical dimensions, with the following specific enhancements:

1.Data Privacy Laws Depth Added:

➠ Explicitly analyzed divergent legal frameworks:

"GDPR Article 9's explicit consent vs. HIPAA's 'safe harbor' de-identification increase cross-institutional sharing costs by 30–50%"

➠ Implemented "GDPR-aligned differential privacy (ε=0.1) with IRB-approved governance" in GI-FedNet.

2.Liability Frameworks Defined:

➠ Diagnosed current ambiguity:

"No clear frameworks for assigning responsibility when AI errors occur"

➠ Proposed concrete solution:

"Liability distribution frameworks defining clinician-vendor responsibilities".

3.Explainability Standards Integrated:

➠ Addressed technical deficits (e.g., "Batch variations reducing DL sensitivity by 20%") with regulatory-compliant tools:

"ISO/IEC TR 24028-compliant uncertainty quantification to meet EU AI Act transparency standards"

➠ Deployed "SHAP-based explainability modules for clinical decisions" and "blockchain-based provenance tracking for GDPR 'right to explanation'".

---

5. "...over-reliance on quantitative metrics... Clinical utility and integration are not thoroughly discussed... include case studies... highlighting barriers such as clinician acceptance, integration into electronic medical records, or procedural delays."

Response:Thank you for your valuable feedback on our manuscript, "[Your Manuscript Title]". We appreciate your guidance on enhancing the discussion of clinical utility, integration challenges, and barriers to adoption. Below is a summary of key revisions addressing these points:

Reduced Emphasis on Pure Quantitative Metrics & Enhanced Discussion of Clinical Utility/Integration:

1.We have significantly restructured and expanded the section discussing the transition from research to clinical practice (formerly potentially overemphasizing technical metrics). This is now a dedicated, in-depth core section: Section [Number/Title, e.g., 4. "Key Barriers to Clinical Translation and Integration of AI Diagnostic Tools"].

2.Within this new section, we systematically analyze critical barriers hindering the practical integration and utility of AI tools in routine clinical workflows. Our discussion moves beyond performance numbers to explicitly address how these barriers impact real-world clinical applicability, adoption, and value, directly tackling the concerns about clinical utility and integration.

6.“...although the article makes multiple references to multimodal integration, it often discusses individual modalities in isolation... without delving into how these modalities can be meaningfully combined... Providing concrete examples of multimodal fusion... would enhance the manuscript's coherence and innovation narrative.”

Response:We sincerely thank the reviewer for this insightful observation. We fully agree that concrete examples of multimodal fusion are essential to strengthen the manuscript’s innovation narrative. Accordingly, we have expanded our discussion of multimodal integration in Sections 2.1.1–2.1.5 with specific clinical scenarios and technical frameworks. Key additions include:

1.Gastric Cancer LNM Risk Stratification:

Fused endoscopic mucosal microstructure analysis (Kyoto Scoring) + CT radiomics (tumor heterogeneity) + serum ctDNA profiles → algorithmically combined to refine LNM prediction beyond unimodal limits (e.g., overcoming low PPV of endoscopy-only models).

2.Colorectal Cancer Phenotype-Biology Bridging:

MRI texture features (peritumoral edema) + WSI-derived immune patterns + proteomic signatures → integrated to link radiomic "habitats" with immune-microenvironment subtypes, elucidating biological underpinnings of imaging phenotypes.

3.Pathology-Imaging-Liquid Biopsy Synergy:

AI-predicted molecular subtypes (H&E slides) + serum biomarkers (BH-index) → dynamically validated as "liquid biopsy correlates" to reduce reliance on costly assays.

TIL spatial features (WSI) + CT/MRI radiomics → fused to overlay TIL heatmaps onto PET/CT scans, resolving "black box" limitations.

These additions (detailed in Sections 2.1.1, 2.1.2, and 2.1.3) explicitly demonstrate how modalities are meaningfully combined, with endoscopy/pathology serving as anchors for workflow-compatible fusion (Figure 1). We believe this significantly enhances translational coherence and underscores our core innovation: clinically actionable, multimodal AI frameworks.

---

7. “...there is some redundancy, especially in the discussions of AI applications in gastric and colorectal cancers. Streamlining repetitive content and grouping similar themes might improve readability.”

Response:Thank you for your valuable feedback highlighting redundancy in discussions of AI applications for gastric/colorectal cancers (particularly regarding "black box" and interpretability issues). We fully agree this repetition impacted readability.

---

1. Consolidated Core Discussion:

We identified that the most detailed and comprehensive discussion of the "black box" problem and the need for interpretable AI solutions (including specific examples like attention heatmaps for gastric cancer and correlation with molecular pathways like

EGFR) was located in Section 2.1.2 (Imaging Analysis). We have retained this core paragraph in its entirety as the primary exposition of this critical challenge and its potential solutions.

2. Systematic Removal of Redundancy:

We have deleted all repetitive descriptions of the "black box" issue and the general need for interpretability from the following subsections:

Section 2.1.1 (Endoscopic Applications): Removed the specific sentences mentioning the "black box" weakening trust and the future need for heatmaps/ranking.

Section 2.1.3 (Pathological Analysis): Removed the sentences discussing "black box" reducing trust and the need for tools like Grad-CAM.

Section 2.1.4 (Biomarker Translation): Removed the sentences focusing on interpretability hindering translation and the mention of heatmaps for MSI.

Section 2.1.5 (Multi-omics Integration): Removed the sentence concerning "ambiguous mechanistic interpretation" and "black-box operations" hindering target discovery.

3.Implemented Concise Cross-Referencing:

In place of the deleted redundant text in each of the subsections listed above, we have inserted brief reference statements that direct the reader to the consolidated discussion in Section 2.1.2. For example:

In Section 2.1.1, we replaced the deleted text with: "As noted in Section 2.1.2, enhancing model interpretability remains a priority to address the 'black box' limitation and build clinical trust."

In Section 2.1.3: "Similar to imaging analysis (Section 2.1.2), improving interpretability is critical for pathological AI adoption."

In Section 2.1.4: "Overcoming the 'black box' challenge (as discussed in Section 2.1.2) is essential for biomarker translation."

In Section 2.1.5: "Mechanistic ambiguity in AI predictions (Section 2.1.2) necessitates causal inference approaches."

4.Resulting Improvement:

This revision strategy has eliminated approximately 120 words of redundant text while preserving all the key concepts related to the "black box" challenge and interpretability solutions. The discussion is now significantly more streamlined and logically structured. Readers encounter the full depth of the issue only once (in 2.1.2), and subsequent sections efficiently reference this core point where relevant, avoiding unnecessary repetition and improving overall readability.

---

8. “In terms of methodology, the review lacks a clear explanation of how literature was selected or assessed. Inclusion and exclusion criteria, search databases, and time frames are not described, which limits reproducibility and transparency. A PRISMA-style flow diagram or methodological appendix would be beneficial.”

Response:We appreciate the reviewer’s feedback on the methodology. Noting that this is a narrative review rather than a systematic review, it does not adhere to a formal systematic literature search protocol. The literature was primarily retrieved through databases like PubMed and Web of Science, complemented by articles downloaded from journal websites. The references reflect the cumulative expertise of the author team based on their routine literature review in the field. All cited references are directly linked to the original works, ensuring authenticity and reliability, and focus on cutting-edge research to highlight the latest advancements in the discipline. We acknowledge the suggestion and would be happy to provide additional details on the literature search strategy if needed.