Evidence-based Practice Center Systematic Review Protocol

Project Title: ***Medical Therapies for Locally Advanced Gastric Adenocarcinoma***

I. Objectives for the Systematic Review

Purpose of the Systematic Review

The primary goal of this systematic review (SR) is to summarize the pertinent evidence so that patients and clinicians can select the best treatment option and policymakers can make sound recommendations on optimal treatment for non-metastatic, locally advanced gastric adenocarcinoma. This review will be utilized to inform a guideline by the American Society of Clinical Oncology (ASCO) on this topic. Specifically, the SR will 1) determine comparative effectiveness of medical therapies, and 2) determine how treatment effectiveness varies by tumor-specific variables and subpopulations. Given an evolving standard of care in the treatment paradigm for non-metastatic, locally advanced gastric adenocarcinoma, the focus of this review will be on comparative effectiveness of medical therapies. Interventions of particular interest include chemotherapy, external beam radiation, targeted therapy, and immunotherapy. This review aims to also study the comparative effectiveness of various combinations and sequences of therapies. Outcomes of interest include survival (overall and progression-free), nutritional assessment, quality of life, and comparative harms (direct and indirect). This SR is intended to inform primarily medical oncologists, alongside members of multidisciplinary teams and tumor boards, including radiation oncologists, surgical oncologists, gastroenterologists, radiologists, pathologists, pharmacists, oncology nurses, and social workers.

II. Key Questions (KQs)

KQ1: What is the comparative effectiveness and comparative harms of medical therapies for management of non-metastatic, locally advanced gastric adenocarcinoma?

KQ2: Do treatment effectiveness and harms vary by cancer stage, histology (*e.g.* intestinal, diffuse, signet ring cell), biomarkers (*e.g.* microsatellite instability-high [MSI-H] or mismatch repair-deficient [MMR-deficient], claudin, human epidermal growth factor receptor 2 [HER-2], programmed death-ligand 1 [PDL1], Epstein–Barr virus [EBV]), or genetic predisposition (*e.g.* cadherin-1 [CDH1])?

KQ3: Do treatment effectiveness and harms vary by age, functional status (*e.g.* Karnofsky score, Eastern Cooperative Oncology Group [ECOG] Performance Status score), medical comorbidities or conditions that increase risk of toxicity with specific therapy (*e.g.* existing neuropathy, prior radiation therapy, history of autoimmune disease)?

Table 1. Population, Interventions, Comparators, Outcomes, Timing, and Setting

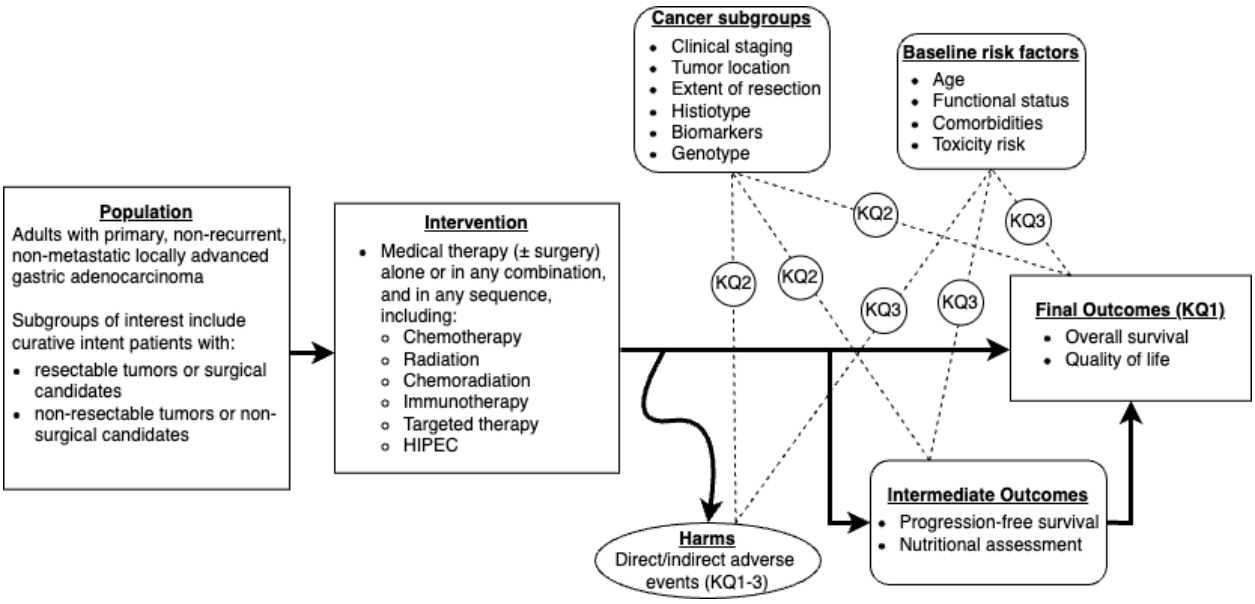
PICOTS	Inclusion	Exclusion
Population	**All KQs:** Adults (18 years or older) with primary, non-recurrent, non-metastatic locally advanced gastric adenocarcinoma stage T2N0 or higher. KQ1: Subgroups of interest may include patients who previously received endoscopic therapy or surgery, patients who are non-surgical candidates, and patients with initially unresectable disease. KQ2: Subgroups of interest may include patients with gastroesophageal junction (GEJ) cancer.	Recurrent cancer, metastatic cancer, early stage (T1aN0 and T1bN0), stage 4 cancer, GEJ cancer patients treated in a predominantly esophageal cancer cohort with an esophageal treatment paradigm, gastrointestinal stromal tumors (GIST), neuroendocrine tumors, gastric lymphoma, MALToma, other rare gastric cancers.
Interventions	**All KQs** Cancer-directed medical therapies administered either alone or in any combination, and may be neoadjuvant, adjuvant, or perioperative (neoadjuvant and adjuvant) and in any sequence, including: • Chemotherapy including but not limited to: Fluoropyrimidine-based therapy: FOLFOX, XELOX, FLOT, SOX, ECF • Radiation including but not limited to external beam radiation, intra-operative electron radiation • Chemoradiation • HIPEC • Immunotherapy (e.g., ipilimumab, nivolimumab) • Targeted therapy (e.g., anti-HER2 monoclonal antibodies)	• Surgical management exclusively • Intervention is not well specified (e.g., study reports intervention as "adjuvant chemotherapy" without describing the regimen) • Palliative interventions
Comparators	**All KQs** • Any comparator • No comparator (for biomarker-targeted interventions)	N/A
Outcomes	**All KQs** • Overall survival • Progression-free survival • Nutritional assessment • Quality of life, using validated scales • Direct moderate-severe treatment adverse events (grade 3, 4, 5) • Direct mild treatment adverse events (grade 1, 2) • Indirect adverse events from treatment (e.g.., long-term opioid use for pain management)	N/A
Timing	**All KQs:** Any follow-up duration for grade 3-5 or indirect adverse events and quality of life; minimum of 1 year for grade 1-2 adverse events; minimum of 3 months for remaining outcomes	N/A
Setting	**All KQs:** • Countries rated as very high on the 2024 Human Development Index (if study is multinational, at least one study center is in a country rated very high)	N/A

Study Design and Other Criteria	**All KQs:** • Randomized controlled trials • Non-randomized studies of interventions (experimental or observational) with a concurrent comparator and well-controlled for confounding (at minimum account for age, stage, functional status, and comorbidities) • Single-arm studies (for biomarker-targeted interventions) • Published in English-language • Published in 2006 or later	Case reports, case series, commentaries, cross-sectional studies, reviews, qualitative studies

Abbreviations: ECF = epirubicin, cisplatin, fluorouracil; FLOT = fluorouracil, leucovorin, oxaliplatin and docetaxel; FOLFOX = leucovorin, fluorouracil, and oxaliplatin; HER2 = human epidermal growth factor receptor 2; HIPEC = hyperthermic intraperitoneal chemotherapy; KQ = key question; SOX = tegafur, gimeracil, oteracil, and oxaliplatin; XELOX = capecitabine and oxaliplatin

Logic Model

Figure 1: Analytic framework for treatment of locally advanced gastric adenocarcinoma



Abbreviations: HIPEC = hyperthermic intraperitoneal chemotherapy; KQ = key question

III. Methods

Criteria for Inclusion/Exclusion of Studies in the Review: We will include randomized trials (RCTs) and comparative well-controlled non-randomized studies of interventions (NRSIs) evaluating chemotherapy, radiation, immunotherapy, targeted therapy, and combinations of these interventions among adults with primary, non-recurrent, non-metastatic, locally advanced gastric adenocarcinoma. We will use the following definition for locally advanced gastric adenocarcinoma: any gastric adenocarcinoma of stage T2N0 or higher, or unstaged (if pre-operative therapy), without evidence of metastatic disease (M0). To be considered well-controlled, comparative NRSIs will have to account for confounding by age, stage, functional status, and comorbidities at a minimum. To help ensure the applicability of the review's findings to standard practice in the United States, we will restrict eligibility to studies conducted in countries rated very high on the 2024 Human Development Index and published in 2006 or later.[1] The full inclusion and exclusion criteria by population, interventions, comparators, outcomes, timing, and settings (PICOTS) and study design are provided in Table 1.

We will also include single-arm studies for interventions targeting specific biomarkers, such as MSI-H. Feedback from the Technical Expert Panel indicated the populations for biomarker-targeted interventions can be small and lack comparative studies, and that single-arm studies have been sufficient to change practice in these circumstances.[2,3]

Full-length articles: The article must be published as a full-length, peer-reviewed study. Abstracts and meeting presentations do not contain sufficient details about experimental methods to permit an evaluation of study design and conduct; they may also contain only a subset of measured outcomes.[4,5] Additionally, it is not uncommon for abstracts to have inconsistencies when compared with the final study publication, or to describe studies that are never published as full articles.[6-9] In order to account for the assessment of publication bias, we will identify relevant abstract presentations during the screening process, but not formally extract from them or assess them for risk of bias.

Literature Search Strategies to Identify Relevant Studies to Answer the Key Questions: A research librarian will develop a comprehensive search strategy. The research librarian will search MEDLINE and EMBASE (via EMBASE.com) and PubMed (publisher-supplied records only) for RCTs and NRSIs published from 2006 to the present. The search start date for 2006 was selected given the 2006 MAGIC trial, which established that surgical management alone was no longer the standard of care, with subsequent trials assessing specific chemotherapy regimens.[10-13]

The search strategy will be independently peer-reviewed by a second librarian using the Peer Review of Electronic Search Strategies (PRESS) Checklist. The proposed search strategy for EMBASE and Medline (via EMBASE.com) is included in Appendix A. Additionally, we will review reference lists of other systematic reviews for inclusion in the current review.

The research librarian will also conduct a grey literature search of relevant stakeholder organizations (e.g., American Cancer Society, the American College of Gastroenterology, the American Gastroenterological Association, the American Society for Gastrointestinal Endoscopy, the American Society of Clinical Oncology, National Comprehensive Cancer Network,), clinical trial registries (e.g., ClinicalTrials.gov), government agencies (e.g., National Institutes of Health [NIH], Agency for Healthcare Research and Quality [AHRQ], Food and Drug Administration [FDA]), Patient-Centered Outcomes Research Institute (PCORI), and other resources identified by the other team members, Key Informants (KIs), and Technical Expert Panel (TEP). We will update our literature search during the peer review of the draft report. A Supplemental Evidence And Data for Systematic review (SEADS) portal will be available for this review, and a Federal Register Notice will be posted to submit additional studies. Studies identified from the grey literature, SEADS packets, or public and peer review will be evaluated against the eligibility criteria for potential inclusion in the review.

We will screen identified records using prespecified criteria to guide study selection. Screening will be conducted with the DistillerSR software (DistillerSR Inc., Ottawa, Canada). We will independently dual screen each title/abstract and full-text article. Discrepancies will be resolved by discussion and consensus among the review team. To

improve the calibration of screening decisions, we will conduct pilot exercises with a minimum of 20 records for title/abstracts and 10 full text articles. If the number of records retrieved from the search is sufficiently large (approximately 10,000 or more) we will consider incorporating artificial intelligence (AI) to provide input for some title/abstract screening decisions to expedite the screening process. More detail is provided in the "Use of Artificial Intelligence and/or Machine Learning" section below.

We will contact authors of original studies if we determine that additional information is needed. If we are unable to reach the study authors or do not receive a response within 6 weeks, we will assess the study without additional input.

Data Abstraction and Data Management: We will develop a structured data extraction form in DistillerSR to capture relevant data from eligible studies. Each study will be extracted by one reviewer, and a minimum of 10% of the extracted studies will be randomly sampled and checked by a second reviewer. Discrepancies will be resolved by discussion and consensus among the review team. We will pilot the data extraction form on a minimum of 4 studies to improve the consistency of extractions and correct any issues arising with the form. As secondary reports (e.g., publications on secondary outcomes or longer-term results) are identified, they will be linked to the primary publication, with relevant data being extracted with the primary publication, to prevent potential double-counting of studies. When several reports of the same or overlapping groups of patients are available, we will prioritize extracting data from the report with the largest number of patients. We will include data from other reports if they provide data on relevant outcomes, subgroups, or time points not provided by the largest report.

For each included study, we will extract the following elements: study characteristics (including study design, country of conduct, and funding), population characteristics (including age, sex, race and ethnicity, socioeconomic status, insurance status, functional status, and medical comorbidities), cancer variables (including staging, histology, and biomarkers, genetic predisposition), any prior intervention (including type and extent of endoscopic or surgical resection), intervention and comparator details (including type, dosage, and frequency), details for relevant outcomes (including outcome domain and specific measurement), and relevant study results (including subgroup and interaction analyses for effect modification variables of interest).

Assessment of Methodological Risk of Bias of Individual Studies: Two reviewers will independently assess the risk of bias for included studies, with discrepancies being resolved by discussion and consensus among the review team. We will assess the risk of bias for RCTs using the Cochrane Risk of Bias 2 (ROB2) and NRSIs (including non-randomized single-arm trials for biomarkers) using the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool.[14,15]

ROB2 comprises five domains:

- Bias arising from the randomization process
- Bias due to deviations from intended interventions
- Bias due to missing outcome data
- Bias in measurement of the outcome
- Bias in selection of the reported result

Each domain includes multiple signaling questions which inform the judgment of risk of bias as "low," "some concerns," or "high" for that domain as well as the overall risk of bias.

ROBINS-I comprises seven domains which will each be assessed:

- Bias due to confounding
- Bias in selection of participants into the study
- Bias in classification of interventions
- Bias due to deviations from intended interventions
- Bias due to missing outcome data
- Bias in measurement of the outcome
- Bias in selection of the reported result

Each domain includes multiple signaling questions which inform the judgment of risk of bias as "low," "moderate," "serious," or "critical" for that domain as well as the overall risk of bias. A study with an overall risk of bias rating of "low" is considered comparable to a well-performed randomized trial.

For using both tools, we will consider the effect of interest to be the effect of assignment to the intervention. As different outcomes and results from the same study may reflect different risks of bias (e.g., objective versus subjective outcomes, different levels of missingness in data across time points), we will conduct assessments at the outcome level. We will only assess risk of bias for outcomes which will be included in the strength of evidence (SOE) assessment (see "Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes" below).

Data Synthesis: We will summarize the evidence qualitatively and quantitatively (e.g., meta-analysis, network meta-analysis) when possible. Decisions about whether a quantitative synthesis is appropriate will be based on the number of studies, and our judgment regarding population, research design, and outcome measures heterogeneity across the included studies. To assess whether pairwise meta-analysis will be appropriate, we will consider population factors such as comorbidities, intervention factors such as intervention dose and frequency, and whether outcomes are the same or address the same concept (most relevant to quality of life and harms) among studies addressing the same types of comparisons. In considering network meta-analysis, we will evaluate how these factors are distributed across direct comparisons, as well as systematic differences in intervention implementation based on the comparator (e.g., whether

fluorouracil, leucovorin, oxaliplatin and docetaxel [FLOT] regimen is implemented similarly in trials comparing it to radiation and those comparing to another chemotherapy regimen).

For data synthesis, we will categorize the interventions following the general categorizations reported in Table 1: chemotherapy, radiation, hyperthermic intraperitoneal chemotherapy (HIPEC), immunotherapy, and targeted therapies. For chemotherapies, we will further subcategorize therapies as non-operative, adjuvant, neoadjuvant, or perioperative. Combinations of these interventions, including the sequence of administration, will also be analyzed separately. Chemoradiation will be considered independently from chemotherapy + radiation, as chemotherapy used in chemoradiation is administered at a lower dose for sensitizing the tumor for radiation than doses administered for traditional chemotherapy. In addition to comparing broader intervention categories, we will also compare specific regimens to each other. We will consider surgery-alone and other comparators outside of the interventions of interest primarily in the context of facilitating indirect comparisons between interventions of interest, but intervention effects will not be framed relative to these comparators.

To reduce clinical heterogeneity, we will analyze studies of patients with resectable tumors or who are surgical candidates separately from those with non-resectable tumors or who are non-surgical candidates. For other subgroups which may be effect modifiers such as biomarkers and functional status, we will analyze them together for KQ1 but separately for KQs 2 and 3, as the purpose of these questions is to determine whether these factors are actually effect modifiers. We will use a best evidence approach to address the key questions.[16] RCTs at low risk of bias or with some concerns and low risk of bias NRSIs will be prioritized for use in the analysis. If these evidence bases are insufficient to provide conclusions on relevant comparisons or outcomes, we would consider incorporating high risk of bias RCTs and moderate, serious, or critical risk of bias NRSIs into the analysis. RCTs with high risk of bias will not be excluded from the review, but these studies will not be prioritized for analysis if better evidence is available. We will analyze RCTs and NRSIs separately unless there is reason to believe the studies are sufficiently consistent in terms of effects and clinically homogenous.

For direct treatment adverse events (AEs) and indirect treatment-related AEs, we will prioritize using summary measures for the purpose of analysis (e.g., incidence of any moderate to severe AE) rather than individual AEs, as these are likely to be rare to infrequent and result in underpowered analyses. Grade 1-2 AEs will be characterized as mild, and grades 3-5 characterized as moderate.

We will summarize time-to-event outcomes (overall survival and progression-free survival) with hazard ratios (HRs). Dichotomous outcomes (incidence of AEs) will be summarized as odds ratios (ORs). For continuous outcomes (quality of life), we anticipate using standardized mean differences (SMDs) as scales reported across studies are likely to differ.

If the data are amenable to meta-analysis, we plan to use random-effects models, anticipating that even with efforts made to reduce heterogeneity in the analyses, there will likely be remaining variability in the study populations and administration of interventions. We will use the restricted maximum likelihood approach to estimate the heterogeneity variance parameter. We will quantify heterogeneity with τ^2, as well as consider using prediction intervals if there are at least 10 studies included in a pairwise comparison. If network meta-analysis is conducted, we will quantify incoherence, the statistical agreement between direct and indirect evidence.

To assess how treatment effects vary by cancer and demographic subgroups for KQs 2 and 3, we will prioritize analyzing within-study comparisons if sufficient information is reported to do so. This data includes the intervention by subgroup interaction term derived from a regression model or subgroup-specific effects which could be used to estimate the interaction.[17,18] If multiple studies evaluating similar comparisons and subgroups report this data, we will consider meta-analyzing the interaction terms. Lacking such data, we will consider meta-regression to estimate associations between these factors and intervention effectiveness. For example, meta-regressing the difference in overall survival between two chemotherapy regimens on the percentage of participants at a particular cancer stage per study. If this approach is taken, results will be interpreted very cautiously, as it is prone to aggregation bias. Additionally, we plan to assess the credibility of potentially relevant effect modification using the Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) tool.[19]

We do not anticipate social determinants of health (e.g., race and ethnicity, gender identity, socioeconomic status) acting as effect modifiers for this review's interventions of interest but acknowledge the importance of addressing racial and ethnic health equity. We plan to describe the distribution of social determinants among included studies as well as report any subgroup or other analyses conducted by the studies which take these factors into account. We additionally plan to employ the PROGRESS-Plus framework (which includes Place; Race/ethnicity/culture/language; Occupation; Out of Work, Gender and sex; Religion, Education, Socioeconomic status, Social capital; and other determinants) to provide clarity and transparency in order to adequately consider health equity in this systematic review.[20] Similarly, we anticipate findings to be applicable to all groups, including underrepresented racial and ethnic groups, given that there are no anticipated biologic or physiologic mechanisms to suggest otherwise.[21]

Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes: We will grade the SOE for the following outcomes:

- Overall survival
- Quality of life
- Direct moderate-severe treatment AEs
- Indirect AEs from treatment

We will grade the SOE according to Evidence-based Practice Center (EPC) Methods Guide recommendations.[22] Two reviewers will independently rate SOE for the relevant bodies of evidence, with discrepancies being resolved by discussion and consensus among the review team.

The primary domains assessed include risk of bias, directness, consistency, precision, and reporting bias. Additional domains may be used when appropriate, including dose-response association, strength of association, and the possibility that controlling for plausible confounders would increase the effect size. The output is a rating of the SOE: high, moderate, low, or insufficient. This rating is made separately for each outcome of each comparison of each KQ.

We will assign a rating of Insufficient when the evidence does not permit a conclusion for the outcome of interest for that KQ (for example, when a difference is not statistically significant, and the 95% confidence is too wide to permit a conclusion that there is no important difference (based on the minimally important difference [MID]). If the evidence is sufficient to permit a conclusion, the rating is deemed high, moderate, or low. Below, we discuss the primary domains and how we will assess them:

Study limitations. This concerns internal validity: the extent to which post-treatment outcomes can be attributed to the treatments themselves rather than other factors. The assessment of study limitations for an evidence base is based on the aggregation of risk of bias ratings for individual studies (described above). If the evidence permits a conclusion, then, all else being equal, a set of studies at low risk of bias yields a higher SOE rating than a set of studies at moderate or high risk of bias. We will not apply different initial SOE ratings based on study design, as ROBINS-I was designed so an NRSI at low risk of bias is comparable to an RCT at low risk of bias.

Directness. Directness relates to (a) whether evidence links interventions directly to a health outcome of specific importance for the review and (b) for comparative studies, whether the comparisons are based on head-to-head studies. If network meta-analytic estimates are used, we will consider the weight of direct versus indirect evidence for the comparison, the plausibility of the transitivity assumption, and evidence of statistical incoherence when assessing this domain.

Consistency. Consistency is the degree to which included studies find either the same direction or similar magnitude of effect.

Precision. Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome, based on the width of confidence intervals relative to a clinically important effect estimate, sufficiency of sample size, and number of events.

Reporting bias. Reporting bias will be addressed by examining the funding source of included studies, the direction and magnitude of effects identified in included studies, possible selective outcome reporting, and noting the presence of abstracts or

ClinicalTrials.gov entries describing studies that did not subsequently appear as full-length published articles. If the evidence base includes at least 10 studies that present data for critical outcomes, review of funnel plots may be used to help ascertain publication bias.

Assessing Applicability: Several factors may limit the applicability of findings, including the extent to which the results from included studies may or may not apply to the full spectrum of populations, interventions, and comparators for this clinical area. Based on EPC guidance, the SOE rating will be uninfluenced by these factors. Instead, we will discuss applicability in a separate section, using PICOTS as a guiding framework to ensure that we consider several components of applicability. Factors important to consider include the type or extent of surgery participants in included studies undergo and the Siewert typing of gastroesophageal junction tumors.

Use of Artificial Intelligence and/or Machine Learning: We will use AI in DistillerSR to automatically re-rank records during title/abstract screening based on their predicted probability of being included in the full-text review. This process sorts the records so that the most likely inclusions are presented first, facilitating faster identification of relevant studies. References identified as potentially relevant during the preliminary literature scan will be included as part of the training set for the AI to help it prioritize records accurately. As human screeners dual-screen each 2% subset of the total records, the AI will iteratively re-order the records, incorporating information from newly identified inclusions or exclusions. This process allows the AI to inform the screening without making any final decisions.

If the search yields a large number of records (approximately 10,000 or more), we may consider allowing the AI to make screening decisions for a subset of records. When references are screened in order of their predicted inclusion probability, the AI uses the rate of inclusion to predict how many of the total relevant records have been identified. Once the AI predicts that 95% of the included records have been screened, we would switch to an accelerated screening approach. In this approach, the AI will act as one screener, excluding records it predicts to be irrelevant, while the final decision will be made by a human screener.

We do not plan on using AI or machine learning in full text review, data extraction, or risk of bias assessment.

IV. References

1. United Nations Development Programme. Human Development Report 2023-24. *UNDP (United Nations Development Programme)*. 2024.
2. André T, Tougeron D, Piessen G, et al. Neoadjuvant Nivolumab Plus Ipilimumab and Adjuvant Nivolumab in Localized Deficient Mismatch Repair/Microsatellite Instability–

High Gastric or Esophagogastric Junction Adenocarcinoma: The GERCOR NEONIPIGA Phase II Study. *Journal of Clinical Oncology*. 2023-01-10 2023;41(2):255-265. doi:10.1200/jco.22.00686.

3. Lin F, Chen Y, Huang B, et al. Application of immune checkpoint inhibitors for resectable gastric/gastroesophageal cancer. *Frontiers in Pharmacology*. 2024-05-09 2024;15doi:10.3389/fphar.2024.1391562.
4. Chalmers I, Adams M, Dickersin K, et al. A cohort study of summary reports of controlled trials. *JAMA*. Mar 9 1990;263(10):1401-5. PMID:2304219.
5. Neinstein LS. A review of Society for Adolescent Medicine abstracts and Journal of Adolescent Health Care articles. *J Adolesc Health Care*. Mar 1987;8(2):198-203. doi:10.1016/0197-0070(87)90265-8. PMID:3818406.
6. Dundar Y, Dodd S, Williamson P, Dickson R, Walley T. Case study of the comparison of data from conference abstracts and full-text articles in health technology assessment of rapidly evolving technologies: does it make a difference? *Int J Technol Assess Health Care*. Summer 2006;22(3):288-94. doi:10.1017/s0266462306051166. PMID:16984055.
7. De Bellefeuille C, Morrison CA, Tannock IF. The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. *Ann Oncol*. Mar 1992;3(3):187-91. doi:10.1093/oxfordjournals.annonc.a058147. PMID:1586615.
8. Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev*. Nov 20 2018;11(11):MR000005. doi:10.1002/14651858.MR000005.pub4. PMID:30480762. PMCID:PMC7073270.
9. Yentis SM, Campbell FA, Lerman J. Publication of abstracts presented at anaesthesia meetings. *Can J Anaesth*. Jul 1993;40(7):632-4. doi:10.1007/BF03009700. PMID:8403137.
10. Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med*. Jul 6 2006;355(1):11-20. doi:10.1056/NEJMoa055531. PMID:16822992.
11. Bang YJ, Kim YW, Yang HK, et al. Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial. *Lancet*. Jan 28 2012;379(9813):315-21. doi:10.1016/S0140-6736(11)61873-4. PMID:22226517.
12. Smalley SR, Benedetti JK, Haller DG, et al. Updated analysis of SWOG-directed intergroup study 0116: a phase III trial of adjuvant radiochemotherapy versus observation after curative gastric cancer resection. *J Clin Oncol*. Jul 1 2012;30(19):2327-33. doi:10.1200/JCO.2011.36.7136. PMID:22585691. PMCID:PMC4517071.
13. Al-Batran SE, Homann N, Pauligk C, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal

junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet*. May 11 2019;393(10184):1948-1957. doi:10.1016/S0140-6736(18)32557-1. PMID:30982686.
14. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019-08-28 2019:l4898. doi:10.1136/bmj.l4898.
15. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. Oct 12 2016;355:i4919. doi:10.1136/bmj.i4919. PMID:27733354. PMCID:PMC5062054.
16. Treadwell JR, Singh S, Talati R, McPheeters ML, Reston JT. AHRQ Methods for Effective Health Care. *A Framework for "Best Evidence" Approaches in Systematic Reviews*. Agency for Healthcare Research and Quality (US); 2011.
17. Godolphin PJ, White IR, Tierney JF, Fisher DJ. Estimating interactions and subgroup-specific treatment effects in meta-analysis without aggregation bias: A within-trial framework. *Research Synthesis Methods*. 2023-01-01 2023;14(1):68-78. doi:10.1002/jrsm.1590.
18. Brantner CL, Chang TH, Nguyen TQ, Hong H, Stefano LD, Stuart EA. Methods for Integrating Trials and Non-experimental Data to Examine Treatment Effect Heterogeneity. *Stat Sci*. Nov 2023;38(4):640-654. doi:10.1214/23-sts890. PMID:38638306. PMCID:PMC11025720.
19. Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *Canadian Medical Association Journal*. 2020-08-10 2020;192(32):E901-E906. doi:10.1503/cmaj.200077.
20. Welch V, Dewidar O, Tanjong Ghogomu E, et al. How effects on health equity are assessed in systematic reviews of interventions. *Cochrane Database Syst Rev*. Jan 18 2022;1(1):MR000028. doi:10.1002/14651858.MR000028.pub3. PMID:35040487. PMCID:PMC8764740.
21. Lin JS, Webber EM, Bean SI, Evans CV. Development of a Health Equity Framework for the US Preventive Services Task Force. *JAMA Network Open*. 2024-03-11 2024;7(3):e241875. doi:10.1001/jamanetworkopen.2024.1875.
22. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville (MD): Agency for Healthcare Research and Quality (US); November 18, 2013.

V. Definition of Terms

None

VI. Summary of Protocol Amendments

If the EPC needs to amend the protocol, provide a numbered list of versions with the date of posting, which will be hyperlinked to previous versions; and a table with the date of each amendment, description of the change, and the rationale. Changes will be incorporated into the protocol.

VII. Previous Versions of the Protocol

Not applicable

VIII. Review of Key Questions

The Agency for Healthcare Research and Quality (AHRQ) posted the Key Questions on the AHRQ Effective Health Care Website for public comment. The Evidence-based Practice Center (EPC) refined and finalized them after reviewing of the public comments and seeking input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the Key Questions are specific and relevant.

IX. Key Informants

Key Informants are the end-users of research; they can include patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into the decisional dilemmas and help keep the focus on Key Questions that will inform health care decisions. The EPC solicits input from Key Informants when developing questions for the systematic review or when identifying high-priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report. They do not review the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The AHRQ Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. The Technical Expert Panel is selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that fosters a thoughtful, relevant systematic review. Therefore,

study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts.

Technical Experts provide information to the EPC to identify literature search strategies and suggest approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind; neither do they contribute to the writing of the report. They do not review the report, except as given the opportunity to do so through the peer or public review mechanism.

Members of the TEP must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The AHRQ TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparing the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers.

The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published 3 months after publication of the evidence report.

Potential peer reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Invited peer reviewers with any financial conflict of interest greater than $5,000 will be disqualified from peer review. Peer reviewers who disclose potential business or professional conflicts of interest can submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Direct financial conflicts of interest that cumulatively total more than $1,000 will usually disqualify an EPC core team investigator.

XIII. Role of the Funder

This project was funded under Contract No. 75Q80120D00002 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The AHRQ Task Order Officer reviewed the EPC response to contract deliverables for adherence to contract

requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by either the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

XIV. Registration

This protocol will be registered in the international prospective register of systematic reviews (PROSPERO).

Appendix A. Search strategies

Embase.com Strategy: (Combines Medline and EMBASE)
Syntax
* = truncation
/exp = explode to include all terms in the tree
/mj = limit to terms indexed as major concepts
/de = search term without exploding
:ti = search in the title field
:kw = search in the author keywords field
:ab = search in the abstract field
NEAR/# - search the terms within # of each other in any order
NEXT/# - search terms within # of each other in the specified order

1 'stomach carcinoma'/exp OR 'stomach cancer'/exp OR (((antral OR antrum OR cardia OR 'esophago-gastric' OR esophagogastric OR fundic OR fundus OR gastric OR 'gastro-esophageal' OR gastroesophageal OR 'gastro-oesophageal' OR gastrooesophageal OR 'oesophago-gastric' OR oesophagogastric OR pyloric OR pylorous OR pylorus OR stomach) NEAR/2 (adenocarcinoma* OR cancer* OR carcinoma* OR malignan* OR neoplas* OR tumor* OR tumour*)):ab,ti,kw)

2 'cancer palliative therapy'/mj OR 'cancer recurrence'/mj OR metastasis/mj OR 'palliative chemotherapy'/mj OR ((advanced:ti OR metasta*:ti OR recur*:ti OR reoccur*:ti) NOT ('locally advanced':ti OR nonadvanced:ti OR 'non-advanced':ti OR nonmetasta*:ti OR 'non-metasta*':ti OR nonrecur*:ti OR 'non-recur*':ti)) OR (palliat* AND (cura* OR cure*)):ti

3 #1 NOT #2

4 capecitabine/de OR 'capecitabine plus oxaliplatin'/de OR 'cancer chemotherapy'/exp OR cisplatin/de OR docetaxel/de OR epirubicin/de OR 'folinic acid'/de OR fluorouracil/de OR 'gimeracil plus oteracil potassium plus tegafur'/de OR oxaliplatin/de OR (capecitabine OR cisplatin OR docetaxel OR epirubicin OR fluorouracil OR 'folinic acid' OR leucovorin OR oxaliplatin OR 'S 1'):ab,ti,kw OR ('5fu lv' OR '5 fu lv' OR 'cape ox' OR capeox OR capox OR chemo* OR ecf OR flot OR folfox* OR hipec OR 'hyperthermic intraperitoneal' OR oxcap OR sox OR xelox):ab,ti,kw

5 'cancer radiotherapy'/exp OR 'external beam radiotherapy'/exp OR ('3d crt' OR imrt OR radiochemo* OR radioimmuno*):ti,ab,kw OR (radiation OR radiotherapy):ti OR (((conformal OR 'external beam' OR 'image guided' OR 'intraoperative electron' OR particle OR proton) NEXT/2 (radiat* OR radiot* OR therap*)):ab,ti,kw)

6 apoptosis/de OR 'cancer immunotherapy'/exp OR 'cancer vaccine'/de OR chemoimmunotherapy/de OR 'checkpoint inhibitor therapy'/de OR 'immune checkpoint inhibitor'/de OR 'microsatellite instability'/de OR 'mismatch repair'/de OR 'tumor mutational burden'/de OR (dmmr OR ctla* OR ebv OR mmr OR msi* OR pd1 OR 'pd-1' OR 'pd-l*' OR pdl* OR tmb*) OR ('checkpoint inhibitor*' OR 'chemo immuno*' OR chemoimmuno* OR 'immune checkpoint*' OR 'immune therap*' OR immunochemo* OR 'immuno-chemo*' OR immunoradio* OR 'immuno-radio*' OR immunotherap* OR 'immuno-therap*' OR 'microsatellite instability' OR 'mismatch repair' OR 'programmed cell death' OR 'programmed death ligand' OR 'radio-immuno*' OR radioimmuno*) OR (atezolizumab OR ipilimumab OR nivolumab OR opdivo OR pembrolizumab OR keytruda OR sintilimab OR tecentriq OR yervoy)

7 'drug targeting'/exp OR 'molecularly targeted therapy'/de OR 'personalized cancer therapy'/exp OR ((claudin* OR 'cldn 18.2' OR cldn18.2 OR egfr OR fgfr OR her2 OR 'her-2' OR (met NOT metasta*) OR vegfr)) OR ((personal* OR target* OR precision) NEAR/2 (treatment* OR medicine OR therap*)) OR (apatinib OR cyramza OR herceptin OR lapatinib OR perjeta OR pertuzumab OR ramucirumab OR

	regorafenib OR rivoceranib OR stivarga OR trastuzumab OR tucatinib OR tukysa OR tykerb) OR (molecular* NEAR/2 alter*)
8	'biological marker'/exp OR 'tumor marker'/de OR (biomarker* OR marker*):ab,ti,kw
9	('book'/de OR 'case report'/de OR 'conference paper'/de OR 'editorial'/de OR 'letter'/de OR (book OR chapter OR conference OR editorial OR letter):it OR [conference abstract]/lim OR [conference paper]/lim OR [conference review]/lim OR [editorial]/lim OR [letter]/lim OR (abstract OR annual OR conference OR congress OR meeting OR proceedings OR sessions OR symposium):nc OR ((book NOT series) OR 'conference proceeding'):pt OR ('case report' OR comment* OR editorial OR letter OR news):ti)
10	(([animals]/lim NOT [humans]/lim) OR ((animal OR animals OR canine* OR dog OR dogs OR feline OR hamster* OR lamb OR lambs OR mice OR monkey OR monkeys OR mouse OR murine OR pig OR piglet* OR pigs OR porcine OR primate* OR rabbit* OR rat OR rats OR rodent* OR sheep* OR swine OR veterinar* OR (vitro NOT vivo)) NOT (human* OR patient*)):ti)
11	((adolescen* OR babies OR baby OR boy* OR child* OR girl* OR infancy OR infant* OR juvenile* OR neonat* OR newborn* OR nurser* OR paediatric* OR pediatric* OR preschool* OR 'school age*' OR schoolchildren* OR teen* OR toddler* OR youth*):ti NOT (adult*:ti,ab OR father*:ti OR matern*:ti,ab OR men:ti,ab OR mother*:ti OR parent*:ti OR patern*:ti,ab OR women:ti,ab))
12	('random sample'/de OR 'randomization'/de OR 'randomized controlled trial'/exp OR 'phase 3':ti,ab OR 'phase iii':ti,ab OR random*:ti,ab OR RCT:ti,ab)
13	'controlled clinical trial'/de OR 'cohort analysis'/de OR 'experimental study'/de OR 'observational study'/de OR 'controlled clinical trial*':ti,ab OR 'prospective study'/de OR 'retrospective study'/de OR (('non randomized' OR nonrandomized) NEXT/1 (compar* OR control* OR trial* OR stud*)) OR ((cohort* OR experimental OR observational OR prospectiv* OR retrospectiv*) NEXT/1 (stud* OR trial*))
14	'comparative effectiveness'/de OR 'comparative study'/de OR compar*:ti
15	'phase 2 clinical trial'/de OR 'case series' OR 'single arm*' OR 'single arm' OR (('un controlled' OR 'non controlled' OR uncontrolled OR noncontrolled OR 'phase 2' OR 'phase two' OR 'phase ii') NEXT/1 (stud* OR trial*))
16	#3 AND (#4 OR #5 OR #6 OR #7) AND [2006-2024]/py AND [english]/lim
17	#16 NOT (#9 OR #10 OR #11)
18	#17 AND #12
19	#17 AND #13
20	#17 AND #14
21	(#3 AND (#6 OR #7 OR #8) NOT (#9 OR #10 OR #11) AND #15) AND [2006-2024]/py AND [english]/lim
22	#18 OR #19 OR #20 OR #21