

Radiographic and endoscopic screening to reduce gastric cancer mortality: a systematic review and meta-analysis

Masaya Hibino,^a Chisato Hamashima,^b Mitsunaga Iwata,^a and Teruhiko Terasawa^{a,*}

^aDepartment of Emergency Medicine and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, Japan

^bHealth Policy Section, Department of Nursing, Faculty of Medical Technology, Teikyo University, Tokyo, Japan



Summary

Background Previous systematic reviews naïvely combined biased effects of screening radiography or endoscopy observed in studies with various designs. We aimed to synthesize currently available comparative data on gastric cancer mortality in healthy, asymptomatic adults by explicitly classifying the screening effects through study designs and types of intervention effects.

Methods We searched multiple databases through October 31, 2022 for this systematic review and meta-analysis. Studies of any design that compared gastric cancer mortality among radiographic or endoscopic screening and no screening in a community-dwelling adult population were included. The method included a duplicate assessment of eligibility, double extraction of summary data, and validity assessment using the Risk Of Bias In Non-randomized Studies of Interventions tool. Bayesian three-level hierarchical random-effects meta-analysis synthesized data corrected for self-selection bias on the relative risk (RR) for per-protocol (PP) and intention-to-screen (ITS) effects. The study registration number at PROSPERO is CRD42021277126.

Findings We included seven studies in which a screening program was newly introduced (median attendance rate, 31%; at moderate-to-critical risk of bias), and seven cohort and eight case-control studies with ongoing screening programs (median attendance rate, 21%; all at critical risk of bias); thus, data of 1,667,117 subjects were included. For the PP effect, the average risk reduction was significant for endoscopy (RR 0.52; 95% credible interval: 0.39–0.79) but nonsignificant for radiography (0.80; 0.60–1.06). The ITS effect was not significant for both radiography (0.98; 0.86–1.09) and endoscopy (0.94; 0.71–1.28). The magnitude of the effects depended on the assumptions for the self-selection bias correction. Restricting the scope to East Asian studies only did not change the results.

Interpretation In limited-quality observational evidence from high-prevalence regions, screening reduced gastric cancer mortality; however, the effects diminished at a program level.

Funding National Cancer Center Japan; and Japan Agency for Medical Research and Development.

Copyright © 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Gastric cancer; Cancer screening; Endoscopy; Radiography; East Asia

Introduction

Gastric cancer is the fifth most frequently diagnosed cancer and the fourth most common cause of cancer-specific mortality, with a worldwide estimated prevalence of over a million cases and approximately 769,000 associated deaths in 2020.¹ The highest incidence rate of 45.7 cases per 100,000 persons was observed in East Asia.¹ The risk factors associated with noncardia gastric cancer include *Helicobacter pylori* infection, high alcohol consumption, tobacco smoking, and the consumption of

salt-preserved foods,² whereas the risk factors associated with cardia gastric cancer include gastroesophageal reflux disease and a high body mass index.²

South Korea and Japan have implemented population-based gastric cancer screening programs.³ The current national guidelines of these countries^{4,5} recommend screening for gastric cancer via either radiography or endoscopy. However, their supporting evidence was limited and largely based on biased, naïvely estimated screening effects on gastric cancer

The Lancet Regional Health - Western Pacific 2023;35: 100741

Published Online 11 March 2023

<https://doi.org/10.1016/j.lanwpc.2023.100741>

*Corresponding author. Section of General Internal Medicine, Department of Emergency Medicine and General Internal Medicine, Fujita Health University School of Medicine, Toyoake, 470-1192, Japan.

E-mail address: terasawa@fujita-hu.ac.jp (T. Terasawa).

Research in context

Evidence before this study

Individual studies performed in high-prevalence regions suggest that screening with both radiography and endoscopy reduces gastric cancer mortality. However, the majority of these studies are not based on designs that rigorously and appropriately assess the effects of screening on gastric cancer mortality in a healthy, community-dwelling population. Such inappropriate designs include studies of patients, studies comparing gastric cancer mortality between a screened cohort and its background general population regardless of the screening status from a vital statistics, and survival analyses between patients with screen-detected gastric cancer and patients with symptom-detected gastric cancer. Even studies exclusively assessing healthy, general populations naïvely estimated screening effects on gastric cancer mortality for the screening attendees vs. nonattendees among the screening-invited population without accounting for self-selection bias. In our PubMed search with no language restrictions from database inception to Oct 31, 2022, for systematic reviews and meta-analyses on gastric cancer screening and gastric cancer mortality, we used the following search strategy: “(gastric cancer screening) AND (meta-analysis OR (systematic review)) AND (endoscopy OR (upper gastrointestinal series)).” Three identified systematic reviews performed meta-analysis and consistently reported that screening by radiography and/or endoscopy decreased risks of gastric cancer mortality (range of the average effect estimates: 0.56–0.63) compared with no screening. However, these results were based on naïvely performed syntheses of inaccurate, unadjusted data reported in the studies with abovementioned designs. To the best of our knowledge, no systematic review has ever addressed these critical points.

Added value of this study

In this meta-analysis, we provided generalized evidence synthesis on gastric cancer mortality in healthy, asymptomatic adults by accounting for different study designs, screening attendance rates, and self-selection bias and calculated effect type-specific, model-corrected estimates on gastric cancer mortality reduction due to screening. Our corrected per-protocol effect (i.e., the effect observed in screened individuals) indicated a risk reduction in gastric cancer mortality, which seemed greater for endoscopy than radiography. By contrast, the mortality reduction in the intention-to-screen effect (i.e., the effect observed at the program level, including both screening invitation and actual attendance) was substantially diluted due to the low attendance rates. Direct comparative data between radiographic and endoscopic screening were limited.

Implications of all the available evidence

Although our model-based corrected results aid and add to the accuracy of results from previous studies, our results further warrant the need for randomized trials to provide reliable evidence on the absolute benefits and harms of gastric cancer screening programs. In countries with widespread implementation of population-based screening programs, we would like to suggest more reliable observational evidence employing research-oriented data collection systems as well as sophisticated analytical approaches for better utilization of real-world data as feasible options.

mortality observed between screening attendees and nonattendees under the already-implemented population-based screening programs. Despite several methodological weaknesses (including self-selection bias),³ subsequent meta-analyses have also naïvely synthesized these inaccurate effect estimates.^{6–8}

Since the publication of the abovementioned guidelines^{4,5} and subsequently-reported systematic reviews,^{6–8} several relevant studies have been published. This systematic review aimed to reanalyze currently available comparative data on gastric cancer mortality in healthy, average risk, asymptomatic adults by explicitly classifying the screening effects by study designs, modalities (radiography vs. endoscopy vs. no screening), and types of intervention effects (effects derived from screening invitation vs. attendance).

Methods

We conducted this study per the Preferred Reporting Items for Systematic reviews and Meta-Analyses

(PRISMA) 2020 statement.⁹ Ethical approval is not required for systematic reviews.

Search strategy and selection criteria

This review repurposed the literature search conducted for the Japanese Guidelines for Gastric Cancer Screening 2015.⁵ We updated the search using PubMed, EMBASE, CENTRAL, and [ClinicalTrials.gov](https://www.clinicaltrials.gov) databases to identify pertinent reports published between January 1, 2012 and October 31, 2022, without language restrictions. In addition, we perused the reference lists of eligible studies and previous systematic reviews and meta-analyses. Further details are listed in [Supplementary Materials](#).

Two reviewers (MH and TT) independently screened abstracts and examined the selected full-text publications. We included nonrandomized comparative studies that compared mortality from gastric cancer among radiographic screening, endoscopic screening, and no screening in a community-dwelling adult (age ≥ 18

years) population at average risk (Table 1). In the case of multiple publications from a given study cohort, we used the publication with the largest sample size or most representative results in the main analysis and the results from the other reports in sensitivity analyses. No randomized clinical trials (RCTs) addressing this research question were eligible. Discrepancies regarding study inclusion were resolved via discussions between the assessors and a third reviewer (CH). The inclusion and exclusion criteria are further defined in the [Supplementary Materials](#).

Data extraction

One reviewer (TT) extracted descriptive data from each eligible paper and the other (MH) confirmed the extracted data ([Supplementary Methods](#)). Both reviewers (MH, TT) independently extracted quantitative outcome data. Adjusted estimates for which the full set of study-specified covariates was accounted were preferred over adjusted estimates with fewer, selected covariates or unadjusted estimates. For case-control studies, estimates based on the standard conditional logistic regression were preferred over unconditional logistic regression. When only sex-specific estimates were reported for a study, we simply pooled the two results under the common-effect assumption to derive a study-level average estimate. We contacted the authors for additional data if a study did not report on the pertinent data ([Supplementary Methods](#)).

We first classified study designs into two groups: screening introduction studies for screen-naïve populations and studies of population-based screening programs (with either a cohort or case-control design) where an ongoing screening program was already established at the study sites. Then, we categorized screening effects into three types: the intention-to-screen (ITS) effect, per-protocol (PP) effect, and *post-hoc* effect as previously defined.^{10,11} The ITS effect assessed the effect of screening invitation regardless of attendance while the PP effect assessed the effect of screening attendance. In both contexts, the effects were estimated in contrast to a

control, screen-naïve population, for which a screening program had never been available (i.e., an uninvited population). In contrast, the *post-hoc* effect was the effect for the screening attendees vs. nonattendees under an ongoing screening program. The *post-hoc* effect is susceptible to self-selection bias due to inconsistencies between group risks of gastric cancer mortality caused by self-selected screening attendance.^{10,11} Operational definitions are included in the [Supplementary Methods](#).

In theory, ITS and PP effects are not directly estimable in studies of population-based screening programs. To address this, we performed a statistical correction to convert the *post-hoc* effect into the corresponding ITS and PP effects.¹⁰ We specified the conversion factor, the D_r ; i.e., the relative risk (RR) of the screening nonattendees compared with the control, uninvited population, as 1.05 in the main analysis and explored a range between 0.8 and 1.1 in the sensitivity analysis (further details in [Supplementary Methods](#)).

Risk of bias assessment

Two reviewers independently rated the risk of bias in screening introduction studies and cohort-type studies of population-based screening programs using the Risk of Bias in Non-randomized Studies of Interventions tool.¹² For case-control studies, we applied the tool's prototype version specifically designed for case-control studies.¹³ Discrepant ratings were resolved via discussions. The specific items considered while rating for each risk of bias domain are presented in [Supplemental Table S1](#).

Synthesis methods

Our primary outcomes were the effects of PP and ITS on gastric cancer mortality. We also assessed the *post-hoc* effect as the reference. We used RRs for the incidence rates as the effect measure. Under the rare event assumption, we deemed odds ratios (ORs) estimated in case-control studies to approximate RRs.

We calculated summary estimates and their 95% credible intervals (CrIs) and prediction intervals (PIs)

PICOTS item	Specific details
Population	Healthy, asymptomatic community-dwelling adults (age ≥ 18 years)
Interventions and Comparators	Radiographic screening, endoscopic screening, and no screening
Outcomes	Mortality from gastric cancer
O1	PP effect: the effect observed among people who attended screening ^a
O2	ITS effect: the effect observed among people who were invited to screening regardless of actual screening attendance ^a
Timing	Not specified
Setting	Population-based or opportunistic gastric cancer screening programs

ITS = intention-to-screen; PICOTS = patient, intervention, comparator, outcome, timing, and setting; PP = per-protocol. ^aFor both PP and ITS effects, the outcomes were compared with people who were not invited to (or could not access) screening. In cohort and case-control studies conducted in the context of an ongoing screening program, the only estimable *post-hoc* effect (i.e., naively estimated effect) observed among people who attended screening, compared with people who did not, was converted into respective PP and ITS effects corrected for self-selection.

Table 1: Inclusion criteria and clinical outcomes of interest based on the PICOTS framework.

using a Bayesian study-level pairwise hierarchical random-effects model meta-analysis when at least two studies were deemed to be appropriately combined.¹⁴ First, we performed meta-analysis separately by study design for each type of screening effect. Subsequently, we performed a generalized evidence synthesis using a three-level hierarchical random-effects model to allow for variation across the results derived from different study designs.^{15,16} For the between-study variance (τ^2), we used an evidence-based informative prior distribution¹⁷ for the main analysis and another weakly informative half-normal prior for the sensitivity analysis.¹⁸ Details of the model specifications, fitting, convergence, and choice of the prior distributions for parameters are reported in [Supplementary Methods](#).

We graphically assessed the between-study statistical heterogeneity and quantified it using the tau and I^2 statistics, along with the 95% PIs of the treatment effects. We did not perform the planned tests for funnel plot asymmetry because there were <10 eligible studies per study design.¹⁹ Limited available data per specific effect type restricted the planned subgroup analyses exclusively to studies conducted in East Asia. For the post-hoc sensitivity analyses, we excluded studies published before 2000 or used a standard hierarchical

random-effects model. In the absence of randomized and high-quality observational evidence, we did not perform the preplanned assessment of certainty in the body of evidence. Meta-analyses were performed using OpenBUGS V.3.2.3 from Stata V.17.²⁰

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the manuscript, or the decision to submit the manuscript for publication.

Results

Study selection

Our literature review identified 22 eligible studies involving 1,667,117 community-dwelling adults in high-prevalence regions^{21–42} (12 from the previous reviews^{21,28–32,35–40} and nine newly identified studies^{22–27,33,34,41,42}) ([Fig. 1](#); [Table 2](#)). Unpublished information was obtained from two studies.^{21,33} A list of excluded publications and reasons for exclusion are available in the [Appendix](#).

Study and participant characteristics

Seven studies (one quasi-experimental study from Costa Rica²¹ and six cohort studies, three from China,^{22,23,27} two

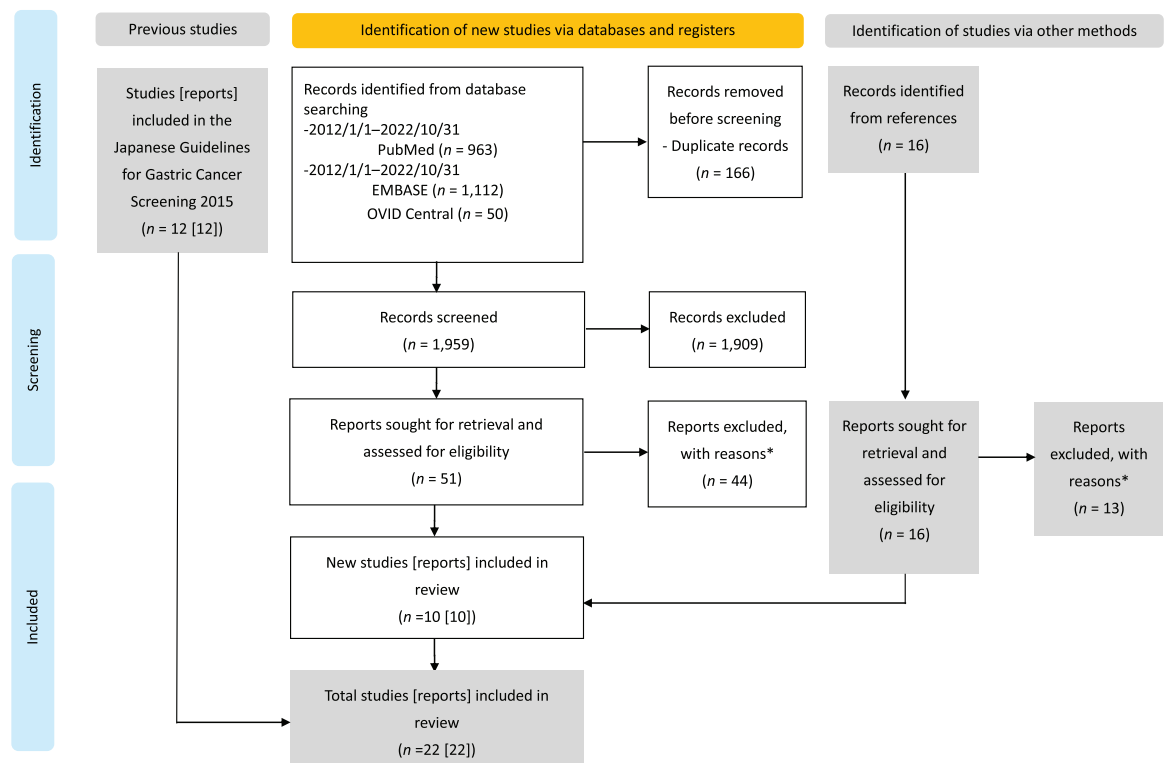


Fig. 1: Study flow diagram*. *See [Supplementary Documents](#) (List of Excluded Publications) for excluded publications and the reasons for exclusion.

Author year	Study location	Target age, year	Screening modality	Screening invitation	Availability of screening program	Exposure, cycles [year]	Attendance rate ^a , %	Follow-up period	Analyzed subjects, n
<i>Studies in screen-naïve population with control group</i>									
<i>Quasi-experimental studies</i>									
Rosero-Bixby 2007 ²¹	Costa Rica (Cartago)	50–75	UGIS	Twice only, biennial	1996–	≥1 [1996–2004]	76	1996–2004	18,777
<i>Cohort studies</i>									
Chen 2009 ²²	China (Ci-xian)	40–69	EGD	Single	2001–2002	1 [2001–2002]	53	2002–2008	15,325
Chen 2021 ²³	China (Multi regions)	40–69	EGD	Single	2005–2006	1 [2005–2012]	34	2005–2015	637,500
<i>Studies in screen-naïve population</i>									
<i>Cohort studies</i>									
Nakamura 1977 ²⁴	Japan (Fukuoka)	40–	UGIS	Annual	1964–	≥1 [1964–1975]	29	1964–1975	1553
Hisamichi 1984 ²⁵	Japan (Miyagi)	40–69	UGIS	Annual	1960–	≥1 [1960–1977]	14	1960–1977	7008
Kim 2018 ²⁶	South Korea (four regions)	40–	UGIS; EGD	Biennial	2002–	≥1 [2002–2014]	31	1993–2014	15,682 ^d
Li 2022 ²⁷	China (Linqi)	40–69	EGD	Single	2012–	1 [2012–2018]	4	2012–2019	375,800
<i>Studies in regions with population-based screening programs</i>									
<i>Cohort studies</i>									
Inaba 1999 ²⁸	Japan (Gifu)	41–	UGIS	Annual	1960s–	1 [1991–1992]	14	1992–1995	24,134
Mizoue 2003 ²⁹	Japan (Multi regions)	40–79	UGIS	Annual	1960s–	1 [1988–1990]	13	1988–1997	87,312
Lee 2006 ³⁰	Japan (Multi regions)	40–59	UGIS	Annual	1960s–	1 [1990]	13	1990–2003	42,150
Miyamoto 2007 ³¹	Japan (Miyagi)	40–64	UGIS	Annual	1960–	1 [2003]	20	1990–2001	41,394
Hamashima 2015 ³²	Japan (Tottori)	40–79	UGIS; EGD	Annual	1960s– ^b 2000– ^c	1 [2007–2008]	25	2007–2013; ^b 2008–2013 ^c	14,274
Hagiwara 2021 ³³	Japan (Gunma)	40–79	UGIS; EGD	Annual	1960s– ^b 2004– ^c	1 [2006]	21	2006–2012	21,802
Narii 2022 ³⁴	Japan (Multi regions)	50–	UGIS	Annual	1960s–	2 [1995–1998; 2000–2003]	13	2000–2015	80,272
<i>Case-control studies</i>									
Oshima 1986 ³⁵	Japan (Osaka)	40–	UGIS	Annual	1962–	≥1 [1962–1981]	21	1969–1981	351
Pisani 1994 ³⁶	Venezuela (Tachira)	35–	UGIS	Annual to biennial	1980–	1981–1989	12	1985–1989	2651
Abe 1995 ³⁷	Japan (Chiba)	40–	UGIS	Annual	1968–	1968–1989	14	1981–1989	3233
Fukao 1995 ³⁸	Japan (Miyagi)	40–	UGIS	Annual	1960–	1980–1991	22	1980–1991	775
Hamashima 2013 ³⁹	Japan (Tottori + Niigata)	40–79	UGIS; EGD	Annual	1960s– ^b 2000–and 2003– ^{c,e}	1960s; ^b 2000–2006 and 2003–2010 ^{c,d}	25	2003–2006 and 2006–2010 ^d	2702
Matsumoto 2014 ⁴⁰	Japan (Nagasaki)	40–	EGD	Annual	1960s–1996 (UGIS); ^{b,f} 1996–(EGD) ^c	1996–2008	22	2000–2008	143
Chen 2016 ⁴¹	China (Linzhou)	40–69	EGD	1-off	2005–	2005–2013	51	2005–2015	2189
Jun 2017 ⁴²	South Korea (Nationwide)	40–	UGIS; EGD	Biennial	2002–	2002–2009	21	2004–2012	272,090

EGD = esophagogastroduodenoscopy; NA = not applicable; ND = no data; UGIS = upper gastrointestinal series. ^aAveraged attendance rates during the study period. Data were extrapolated from other resources when not presented in the study report (see [Supplemental Table S4](#)). ^bFor radiographic screening. ^cFor endoscopic screening. ^dSome overlapping cases are possible. ^eData are for Tottori and Niigata, respectively. ^fRadiographic screening was no longer available since 1996.

Table 2: Included studies of gastric cancer screening.

from Japan,^{24,25} and one from South Korea²⁶) assessed the effect of introducing screening programs in screen-naïve populations (Table 2; [Supplemental Table S2](#)). Of these, three had a contemporaneous control population,^{21–23} whereas in the other studies, the comparison was between attendees and nonattendees of the newly introduced programs.^{24–27}

The other 15 studies (seven population-based cohort studies in Japan^{28–34} and eight case-control studies [five

in Japan^{35,37–40} and one each in Venezuela,³⁶ China,⁴¹ and South Korea⁴²) were conducted in the context of ongoing screening programs (Table 2). These studies retrospectively assessed data from local screening registries or parental cohort studies; the latter was typically designed for assessing risk factors for non-communicable diseases, and the screening attendance relied on data from self-reported questionnaires or interviews ([Supplemental Table S2](#)).

Studies generally targeted community-dwelling adults aged over 40–50 years and excluded those with a prior history of gastric cancer (Table 2). For cohort-type studies, the mean or median follow-up duration range was 3.2–18.5 years. Case-control studies similarly selected gastric cancer death cases from the local residents registered in the local death and/or cancer registries and then selected age-, sex-, and resident area-matched controls from the same resident populations (Supplemental Table S2).

Studies inconsistently addressed signs and/or symptoms in screened subjects (Supplemental Table S3). Only one prospective cohort study²³ explicitly excluded symptomatic subjects upon enrollment. Subjects diagnosed with gastric cancer immediately after the recruitment (within 6–18 months) were *post-hoc* excluded in one screening introduction study²⁷ and one cohort study²⁸ in the main analysis and in three other cohort studies in the sensitivity analysis only^{21,29,30} because of the possibility of being symptomatic cases. Similarly, three case-control studies^{35,36,38} operationally disregarded the screening attendance within 6–12 months before gastric cancer diagnosis in the case group. Sparse and inconsistent reports on risk factors for gastric cancer precluded meaningful across-study comparisons (Supplemental Table S3).

Screening tests and attendance

Studies typically reported radiographic screening as 6- to 8-film double-contrast direct or indirect radiography, whereas the reported methods of endoscopic screening generally lacked sufficient details, except for one recent study (Supplemental Table S4).²³ Details of the performers and/or interpreters of the screening results, positivity rates and criteria, and adherence to the recall and confirmatory investigations were scarcely reported.

The number of screening cycles provided in the screening introduction studies varied, including a single-cycle endoscopic screening in three Chinese studies,^{22,23,27} a two-cycle radiographic screening in the quasi-experimental study conducted in Puerto Rico,²¹ and annual^{24,25} or biennial²⁶ screening programs (Table 2). The attendance rates were generally low (4%–34%), except in two studies (53%–76%).^{21,22}

In cohort and case-control studies conducted under ongoing screening programs, screening was locally available at annual or biennial intervals (Table 2). The attendance rates reported were low (12%–25%), except in one study conducted in China (51%).⁴¹

Other interventions

Only two screening introduction studies provided the follow-up methods for subjects with specific screening results (Supplemental Table S4).^{23,27} Sufficient details of the treatment for diagnosed gastric cancer were reported in two screening introduction studies only.^{21,23} No

studies reported on the attendance of opportunistic screening programs.

Risk of bias

The ratings of the risk of bias assessment varied for screening introduction studies: One was rated as having a moderate risk of bias,²³ three were rated as having a severe risk of bias,^{21,22,27} and three others were rated as having a critical risk of bias^{24–26} (Supplemental Fig. S1; Supplemental Tables S6–S8). In contrast, all cohort^{28–34} and case-control studies^{35–42} conducted under ongoing screening programs were rated as having a critical risk of bias. The risk of bias was consistently serious for all studies in two specific domains, confounding and classifications of screening attendance (Supplemental Figs. S2 and S3; Supplemental Tables S6–S11).

Effect of radiographic screening vs. no screening

A total of 14 studies on radiographic screening (four screening introduction studies,^{21,24–26} four cohort studies,^{28–31} and six case-control studies^{35–39,42} under ongoing screening programs) reported findings on the *post-hoc* effect (Supplemental Figs. S4–S6). While the summary estimate was significant with only small across-study heterogeneity for screening introduction studies (RR: 0.42; 95% CrI: 0.29–0.59; 95% PI: 0.25–0.68) and cohort studies (RR: 0.61; 95% CrI: 0.47–0.79; 95% PI: 0.40–0.94), the summary estimate for case-control studies was not significant, with moderate-to-substantial across-study heterogeneity (RR: 0.69; 95% CrI: 0.45–1.02; 95% PI: 0.24–1.92). After excluding one overlapping nonrepresentative study,²⁶ the generalized synthesis yielded a significant overall estimate associated with a reduced risk by an average of 43% (RR: 0.57; 95% CrI: 0.43–0.75; 95% PI: 0.35–0.90) (Supplemental Fig. S7).

The PP effect was reported by one screening introduction study only,²¹ which suggested a significant risk reduction of 37% (RR: 0.63; 95% CrI: 0.44–0.90) (Fig. 2). Additional, nonoverlapping two screening introduction,^{24,25} four cohort,^{28–31} and six case-control^{35–39,42} studies provided data amenable to adjustment for self-selection bias, the point estimates of which ranged from 0.52 to 2.14. The overall result based on the generalized synthesis suggested an average risk reduction of 20% with only small-to-moderate across-design heterogeneity (RR: 0.80; 95% CrI: 0.60–1.06; 95% PI: 0.50–1.25). The summary estimate was not significant.

The ITS effect was, again, reported by one screening introduction study only,²¹ which was no longer significant (RR: 0.84; 95% CrI: 0.62–1.13) (Fig. 2). Also, nonoverlapping screening ($n = 2$),^{24,25} cohort ($n = 4$),^{28–31} and case-control ($n = 6$)^{35–39,42} studies provided data amenable to adjustment for self-selection bias, with a point estimate range of 0.87–1.09. The generalized synthesis calculated an overall estimate, suggesting an

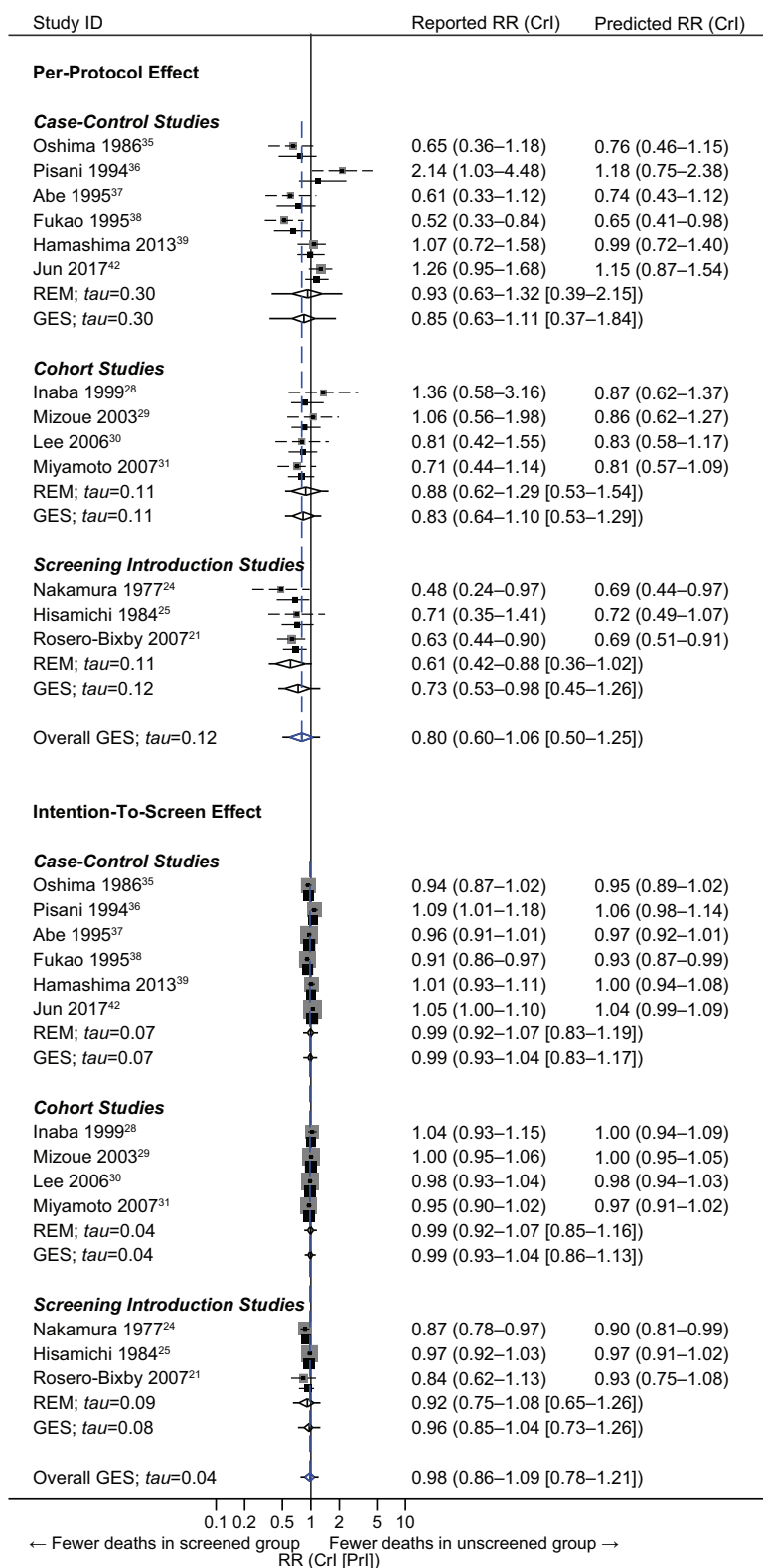


Fig. 2: Per-protocol (upper) and intention-to-screen (lower) effects for gastric cancer mortality by radiographic screening. The diamond represents the summary risk ratio (RR) centered on a combined estimate and extending to a 95% credible interval (CrI), with estimated 95%

almost null effect with small across-design heterogeneity (RR: 0.98; 95% CrI: 0.86–1.09; 95% PI: 0.78–1.21).

Effect of endoscopic screening vs. no screening

A total of seven studies on endoscopic screening (three screening introduction studies^{23,26,27} and four case-control studies^{39–42} under ongoing screening programs) reported data on the *post-hoc* effect (Supplemental Figs. S8 and S9). All studies consistently reported significant estimates associated with a reduced risk of gastric cancer mortality. The summary estimates were significant for both study designs and suggested an average risk reduction of 49% with only small-to-moderate across-study heterogeneity for screening introduction studies (RR: 0.51; 95% CrI: 0.29–0.72; 95% PI: 0.21–0.97), and a reduction of 42% with small-to-moderate between-study heterogeneity for case-control studies (RR: 0.58; 95% CrI: 0.45–0.78; 95% PI: 0.36–0.99). After excluding two overlapping, nonrepresentative studies,^{26,41} the overall result based on the generalized synthesis also suggested a significant average risk reduction of 46% with small across-design heterogeneity (RR: 0.54; 95% CrI: 0.39–0.72; 95% PI: 0.34–0.83) (Supplemental Fig. S10).

The PP effect was reported by only one screening introduction study²³; the estimate suggested a significant risk reduction of 54% (RR: 0.46; 95% CI: 0.41–0.52) (Fig. 3). Additional three case-control studies^{39,40,42} provided data amenable to adjustment for self-selection bias. The overall result based on the generalized synthesis calculated a significant summary estimate associated with a risk reduction of 48% with small-to-moderate across-design heterogeneity (RR: 0.52; 95% CrI: 0.39–0.79; 95% PI: 0.32–0.99).

The ITS effect was reported by two screening introduction studies: the estimate from one study²³ suggested a significant risk reduction (RR: 0.72; 95% CI: 0.68–0.77), whereas the other study reported a significant and contradictory estimate suggesting increased risk (RR: 1.72; 95% CrI: 1.12–2.65) (Fig. 3). Additional three case-control studies^{39,40,42} provided data amenable to adjustment for self-selection bias. The generalized synthesis yielded a nonsignificant summary estimate suggesting a risk reduction of 6% (RR: 0.94; 95% CrI: 0.71–1.28; 95% PI: 0.64–1.44).

Comparative effect of radiographic screening vs. endoscopic screening

Only three cohort studies under ongoing screening programs compared the *post-hoc* effects of these two

modalities (Supplemental Fig. S11).^{26,32,33} We did not perform a meta-analysis because of the wide-ranging effect sizes suggesting possible superiority and inferiority for both modalities and the critical risk of bias particularly for the classification of interventions (disregard for the participants' switching between the two modalities during the study period).

Sensitivity analysis

In the PP effect for radiographic screening, decreasing the Dr (i.e., increasing the average risk of gastric cancer mortality in screening attendees in comparison to screening-nonattendees) substantially lowered the summary effect estimates (e.g., from 0.81 [Dr = 1.05] in the main analysis to 0.26 [Dr = 0.8] in the sensitivity analysis), which yielded significantly reduced risks (Supplemental Fig. S12). Decreasing the Dr similarly lowered the ITS effect for radiographic screening (e.g., from 0.98 [Dr = 1.05] in the main analysis to 0.75 [Dr = 0.8] in the sensitivity analysis), yielding significantly reduced risks. Although decreasing the Dr lowered the summary point estimates for the PP effect as well as the ITS effect for endoscopic screening, the results remained nonsignificant in the ITS effect (Supplemental Fig. S13). The results did not differ significantly when studies conducted outside Asia or studies reported before 2000 were excluded or the alternative random-effects model was specified. Furthermore, the results were in general agreement when alternative priors for the between-study and/or across-design heterogeneity were specified (Supplemental Figs. S12–S15). Finally, replacing the data on the radiographic screening from a Japanese cohort study³⁰ with their subsequent report³⁴ did not substantially alter the conclusions (Supplemental Figs. S16–S18).

Discussion

In this systematic review and meta-analysis, we analyzed 22 nonrandomized comparative studies of radiologic and/or endoscopic screening involving 1,667,117 community-dwelling adults. This research included only studies of community-dwelling adults and used a refined framework of different types of screening effects accounting for both different study designs and inherent bias due to self-selection, which previous meta-analyses failed to address (Supplemental Table S12). First, regarding the theoretically expected effect itself on gastric cancer mortality observed in the screened individuals (i.e., the PP effect), radiographic screening was

prediction intervals (PrIs) depicted as extending solid horizontal lines. Gray squares and dashed horizontal lines indicate self-selection bias-corrected prior estimates of risk ratios and their 95% CIs. Black squares and solid horizontal lines indicate (study-specific) predicted risk ratios and 95% CrIs based on the posterior distribution of individual studies. The size of the square is proportional to the inverse of the variance of the log risk ratio of each study. GES = generalized evidence synthesis; REM = random-effects model.

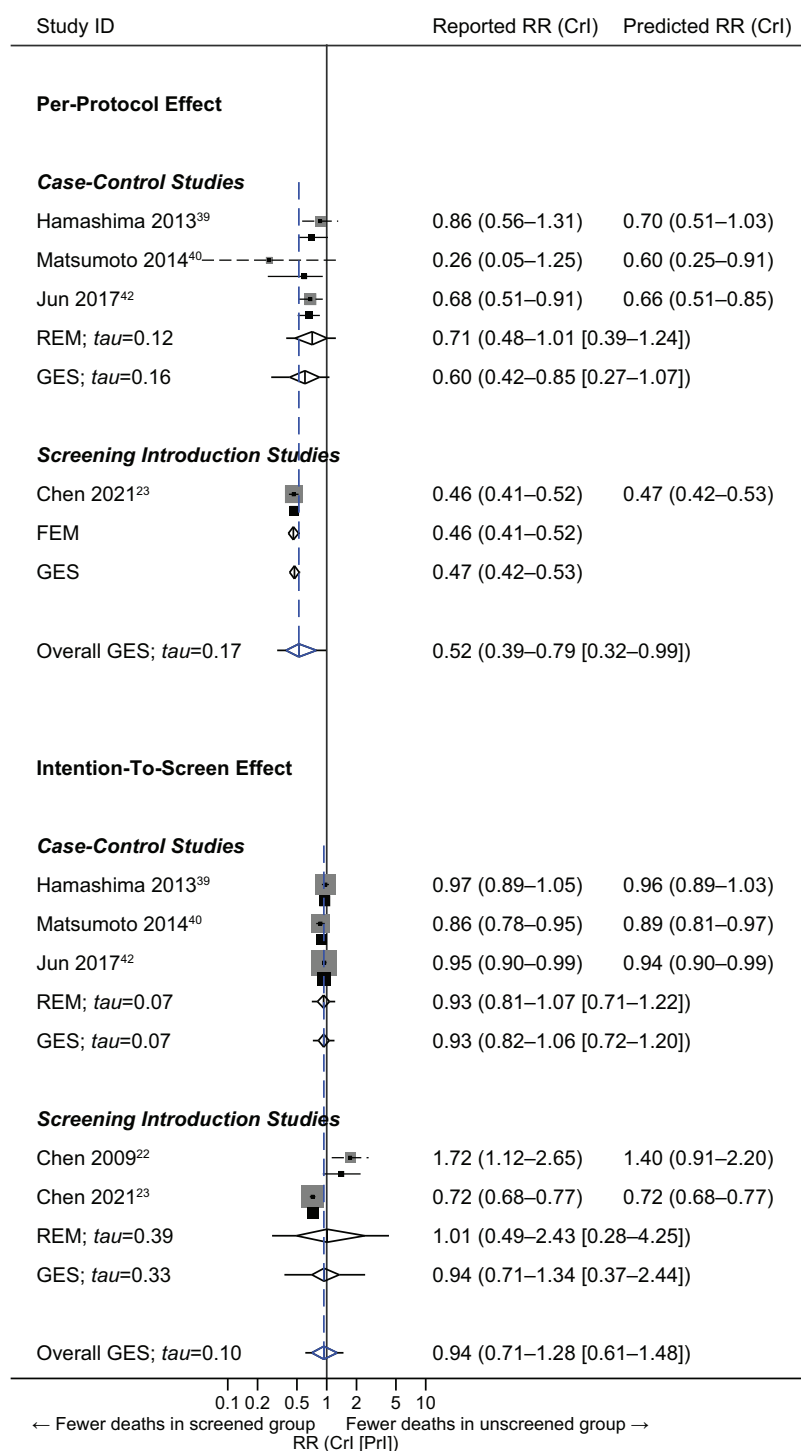


Fig. 3: Per-protocol (upper) and intention-to-screen (lower) effects for gastric cancer mortality by endoscopic screening. The diamond represents the summary risk ratio (RR) centered on a combined estimate and extending to a 95% credible interval (CrI), with estimated 95% prediction intervals (PrIs) depicted as extending solid horizontal lines. Gray squares and dashed horizontal lines indicate self-selection bias-corrected prior estimates of risk ratios and their 95% CIs. Black squares and solid horizontal lines indicate (study-specific) predicted risk ratios and 95% CrIs based on the posterior distribution for individual studies. The size of the square is proportional to the inverse of the variance of the log risk ratio of each study. FEM = fixed-effect model; GES = generalized evidence synthesis; REM = random-effects model.

associated with an average reduced risk of 20%, which was only marginally significant, whereas endoscopic screening was associated with a significant average reduced risk of 48%. Second, when the screening effect was assessed at the whole program level addressing the joint effect of both screening invitation and attendance (i.e., the ITS effect), both radiographic and endoscopic screening programs were no longer significantly associated with a decreased risk of gastric cancer mortality—the low attendance rates diluted the effects observed at an individual level. Third, these results did not significantly change in sensitivity analyses. Importantly, both the PP and ITS effects for both modalities were based on limited-quality observational evidence and appeared sensitive to the correction factor used in the adjustment for self-selection bias. Fourth, no trustworthy direct comparative data existed between radiographic and endoscopic screening.

Our review has several limitations. First, our results relied on data from nonrandomized studies only. Second, the utilized model-adjustment for self-selection bias¹⁰ has the potential to address unmeasured confounders,⁴³ which has been successfully applied in other cancer screening disciplines.^{44,45} Nevertheless, due to the lack of pertinent data, we had to rely on several assumptions and multiple sensitivity analyses based on hypothetical data—potentially missing important information including attendance to opportunistic screening and surveillance testing performed after positive screening, seroprevalence and/or the eradication history of *H. pylori*, and improvement in the treatment of gastric cancer *per se*. Third, we did not include risk-stratified screening programs based on *H. pylori* infectious status as the comparator strategy. Fourth, our review failed to address population- and program-specific factors that can modify the screening effects such as the target age group and screening intervals. Regardless, RCTs, are the best sources to provide reliable evidence on these effect modifiers to assist individualized clinical decision-making.⁴⁶ Finally, the comparative effectiveness of gastric cancer screening should be based not only on mortality benefits but also on harms attributable to screening,⁴⁷ which include recall rates for additional diagnostic tests (including biopsy with histopathology) that were proven to be unnecessary, and overdiagnosis and its consequences, if any.⁴⁸ The lack of these data precluded a formal assessment of harms.

Although the magnitude of the corrected screening effects is not as reliable as that of rigorously-conducted RCTs, given the mortality benefit observed for screening attendees coupled with the large disease burden in East Asia, our synthesized nonsignificant ITS effects due to low attendance rates should not deter people from participating in recommended radiographic and endoscopic screenings.

Currently, two gastric cancer screening RCTs, a Chinese cluster RCT comparing single-cycle

endoscopic screening to a control⁴⁹ and a small-sized Japanese RCT comparing a risk-stratified endoscopic screening program to the standard radiographic screening program,⁵⁰ are ongoing. Regrettably, both are unlikely to answer the key questions targeted in this review—what is the magnitude of gastric cancer mortality benefit from population-based radiographic or endoscopic screening programs (such as those currently provided annually or biennially in South Korea or Japan) compared with no screening at all. Large, long-term RCTs certainly provide the most reliable comparative evidence of both benefits and harms of cancer screening.⁴⁷ However, cohort and case-control studies are the only realistic options in regions where ongoing screening programs already exist. Therefore, given their limited-quality, observational evidence also needs to be refined. Resources to achieve this goal should be multifaceted—efficient and rigorous methodologies to validate contemporary screening effectiveness using routinely collected data, including *de novo*, a more accurate and upfront data collection system fully linked with other sources of information such as medical records at the individual level. This is particularly relevant to reliably classifying the receipt and objective for implementing a test both within and outside population-based programs, confounders to be adjusted, and the cause of death. The use of sophisticated analytical techniques^{51,52} to address biases inherent in real-life data derived from the already-implemented screening programs is another key point.

Contributors

CH and TT conceptualized the study. All authors were involved in the design of the study. TT performed the literature search. MH, CH, and TT determined the study eligibility. MH and TT performed the data extraction and assessed the quality of included studies. MH, CH, and TT analyzed the data. All authors interpreted the results. TT drafted the first version of the manuscript; all co-authors contributed to the writing of the manuscript and approved the final version. MH and TT accessed and verified the data. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. All authors had access to the data in the study and had final responsibility for the decision to submit for publication.

Data sharing statement

All data relevant to the study are either included in the article or uploaded as supplementary information.

Declaration of interests

We declare no competing interests.

Acknowledgments

This research was supported by the National Cancer Center Research and Development Fund from the National Cancer Center, Tokyo, Japan (Grant Number 29-A-16); and Japan Agency for Medical Research and Development (AMED) under Grant Number JP22ck0106729. The authors thank Drs. Luis Rosero-Bixby and Hiroaki Hagiwara for providing unpublished data from their original studies. The English language editing was provided by MARUZEN-YUSHODO Co., Ltd. (<https://kw.maruzen.co.jp/kousei-honyaku/>). This assistance was funded by AMED (Grant Number JP22ck0106729).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lanwpc.2023.100741>.

References

- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.
- Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet*. 2020;396(10251):635–648.
- Huang RJ, Koh H, Hwang JH. A summary of the 2020 gastric cancer summit at stanford university. *Gastroenterology*. 2020;159(4):1221–1226.
- Park HA, Nam SY, Lee SK, et al. The Korean guideline for gastric cancer screening. *J Kuwait Med Assoc*. 2015;58(5):373–384.
- Hamashima C, Group SR. Guidelines GDGfGCS. Update version of the Japanese guidelines for gastric cancer screening. *Jpn J Clin Oncol*. 2018;48(7):673–683.
- Khanderia E, Markar SR, Acharya A, Kim Y, Kim YW, Hanna GB. The influence of gastric cancer screening on the stage at diagnosis and survival: a meta-analysis of comparative studies in the far East. *J Clin Gastroenterol*. 2016;50(3):190–197.
- Zhang X, Li M, Chen S, et al. Endoscopic screening in Asian countries is associated with reduced gastric cancer mortality: a meta-analysis and systematic review. *Gastroenterology*. 2018;155(2):347–354.e9.
- Faria L, Silva JC, Rodríguez-Carrasco M, Pimentel-Nunes P, Dinis-Ribeiro M, Libânio D. Gastric cancer screening: a systematic review and meta-analysis. *Scand J Gastroenterol*. 2022;57:1–11.
- Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
- Duffy SW, Cuzick J, Tabar L, et al. Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. *J Roy Stat Soc C*. 2002;51(2):235–243.
- Spix C, Berthold F, Hero B, Michaelis J, Schilling FH. Correction factors for self-selection when evaluating screening programmes. *J Med Screen*. 2016;23(1):44–49.
- Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
- Sterne JAC, Higgins JPT, Reeves BC, on behalf of the development group for ACROBAT-NRSI. A Cochrane Risk of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI). 2014, Version 1.0.0, 24 September 2014. https://www.bristol.ac.uk/media-library/sites/social-community-medicine/images/centres/cresyda/ACROBAT-NRSI%20Version%201_0_0.pdf. Accessed April 16, 2022.
- Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making*. 2013;33(5):607–617.
- Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. *Stat Med*. 2000;19(24):3359–3376.
- Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med*. 2003;22(23):3687–3709.
- Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Stat Med*. 2015;34(6):984–998.
- Röver C, Bender R, Dias S, et al. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Res Synth Methods*. 2021;12(4):448–474.
- Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
- Thompson J. *Bayesian analysis with Stata*. Stata Press; 2014.
- Rosero-Bixby L, Sierra R. X-ray screening seems to reduce gastric cancer mortality by half in a community-controlled trial in Costa Rica. *Br J Cancer*. 2007;97(7):837–843.
- Chen Z, Song G, Hou J, et al. The relationship between a cohort endoscopic screening of 15,000 subjects and the death of patients with esophageal and gastric cancers in high-risk areas. *J Cancer Res Clin Oncol*. 2009;6(4):271–276.
- Chen R, Liu Y, Song G, et al. Effectiveness of one-time endoscopic screening programme in prevention of upper gastrointestinal cancer in China: a multicentre population-based cohort study. *Gut*. 2021;70(2):251–260.
- Nakamura Y, Takeshita M, Hirota Y, et al. An evaluation of gastric cancer screening program in Hisayama, Japan. *Gastroenterol Jpn*. 1977;12(6):427–434.
- Hisamichi S, Sugawara N. Mass screening for gastric cancer by X-ray examination. *Jpn J Clin Oncol*. 1984;14(2):211–223.
- Kim H, Hwang Y, Sung H, et al. Effectiveness of gastric cancer screening on gastric cancer incidence and mortality in a community-based prospective cohort. *Cancer Res Treat*. 2018;50(2):582–589.
- Li WQ, Qin XX, Li ZX, et al. Beneficial effects of endoscopic screening on gastric cancer and optimal screening interval: a population-based study. *Endoscopy*. 2021;54(9):848–858.
- Inaba S, Hirayama H, Nagata C, et al. Evaluation of a screening program on reduction of gastric cancer mortality in Japan: preliminary results from a cohort study. *Prev Med*. 1999;29(2):102–106.
- Mizoue T, Yoshimura T, Tokui N, et al. Prospective study of screening for stomach cancer in Japan. *Int J Cancer*. 2003;106(1):103–107.
- Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S. Gastric cancer screening and subsequent risk of gastric cancer: a large-scale population-based cohort study, with a 13-year follow-up in Japan. *Int J Cancer*. 2006;118(9):2315–2321.
- Miyamoto A, Kuriyama S, Nishino Y, et al. Lower risk of death from gastric cancer among participants of gastric cancer screening in Japan: a population-based cohort study. *Prev Med*. 2007;44(1):12–19.
- Hamashima C, Shabana M, Okada K, Okamoto M, Osaki Y. Mortality reduction from gastric cancer by endoscopic and radiographic screening. *Cancer Sci*. 2015;106(12):1744–1749.
- Hagiwara H, Moki F, Yamashita Y, Saji K, Iesaki K, Suda H. Gastric cancer mortality related to direct radiographic and endoscopic screening: a retrospective study. *World J Gastroenterol*. 2021;27(33):5595–5609.
- Narii N, Sobue T, Zha L, et al. Effectiveness of endoscopic screening for gastric cancer: the Japan public health center-based prospective study. *Cancer Sci*. 2022;113(11):3922–3931.
- Oshima A, Hirata N, Ubukata T, Umeda K, Fujimoto I. Evaluation of a mass screening program for stomach cancer with a case-control study design. *Int J Cancer*. 1986;38(6):829–833.
- Pisani P, Oliver WE, Parkin DM, Alvarez N, Vivas J. Case-control study of gastric cancer screening in Venezuela. *Br J Cancer*. 1994;69(6):1102–1105.
- Abe Y, Mitsushima T, Nagatani K, Ikuma H, Minamihara Y. Epidemiological evaluation of the protective effect for dying of stomach cancer by screening programme for stomach cancer with applying a method of case-control study—a study of a efficient screening programme for stomach cancer. *Nihon Shokakibyō Gak-kai Zasshi*. 1995;92(5):836–845.
- Fukao A, Tsubono Y, Tsuji I, Hisamichi S, Sugahara N, Takano A. The evaluation of screening for gastric cancer in Miyagi Prefecture, Japan: a population-based case-control study. *Int J Cancer*. 1995;60(1):45–48.
- Hamashima C, Ogoshi K, Okamoto M, Shabana M, Kishimoto T, Fukao A. A community-based, case-control study evaluating mortality reduction from gastric cancer by endoscopic screening in Japan. *PLoS One*. 2013;8(11):e79088.
- Matsumoto S, Yoshida Y. Efficacy of endoscopic screening in an isolated island: a case-control study. *Indian J Gastroenterol*. 2014;33(1):46–49.
- Chen Q, Yu L, Hao CQ, et al. Effectiveness of endoscopic gastric cancer screening in a rural area of Linzhou, China: results from a case-control study. *Cancer Med*. 2016;5(9):2615–2622.
- Jun JK, Choi KS, Lee HY, et al. Effectiveness of the Korean national cancer screening program in reducing gastric cancer mortality. *Gastroenterology*. 2017;152(6):1319–1328.e7.
- Autier P, Boniol M. Pitfalls in using case-control studies for the evaluation of the effectiveness of breast screening programmes. *Eur J Cancer Prev*. 2013;22(5):391–397.
- Dibden A, Offman J, Duffy SW, Gabe R. Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. *Cancers (Basel)*. 2020;12(4):976.
- Jansen EEL, Zielonke N, Gini A, et al. Effect of organised cervical cancer screening on cervical cancer mortality in Europe: a systematic review. *Eur J Cancer*. 2020;127:207–223.

- 46 Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. *Ann Intern Med.* 2020;172(1):35–45.
- 47 Harris RP, Wilt TJ, Qaseem A. A value framework for cancer screening: advice for high-value care from the American College of Physicians. *Ann Intern Med.* 2015;162(10):712–717.
- 48 Hamashima C. Overdiagnosis of gastric cancer by endoscopic screening. *World J Gastrointest Endosc.* 2017;9(2):55–60.
- 49 Zeng H, Sun K, Cao M, et al. Initial results from a multi-center population-based cluster randomized trial of esophageal and gastric cancer screening in China. *BMC Gastroenterol.* 2020;20(1):398.
- 50 Gotoda T, Ishikawa H, Ohnishi H, et al. Randomized controlled trial comparing gastric cancer screening by gastrointestinal X-ray with serology for *Helicobacter pylori* and pepsinogens followed by gastrointestinal endoscopy. *Gastric Cancer.* 2015;18(3):605–611.
- 51 Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016;183(8):758–764.
- 52 Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ.* 2017;359:j4587.