The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality

Published in:
Gastroenterology

Publication status and date:
Published: 01/03/2024

DOI (link to publisher):
10.1053/j.gastro.2023.11.286

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality: a Synthetic Control Study

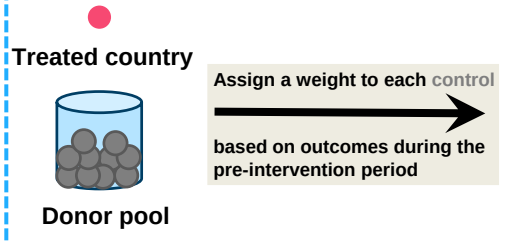
Dianqin Sun, MMed, Duco T. Mülder, MSc, Yige Li, MSc, Daan Nieboer, PhD, Jin Young Park, PhD, Mina Suh, MD PhD, Chisato Hamashima, MD, Weiran Han, MSc, James F. O'Mahony, PhD, Iris Lansdorp-Vogelaar, PhD

PII: S0016-5085(23)05562-2
DOI: https://doi.org/10.1053/j.gastro.2023.11.286
Reference: YGAST 66024

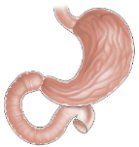
To appear in: *Gastroenterology*
Accepted Date: 19 November 2023

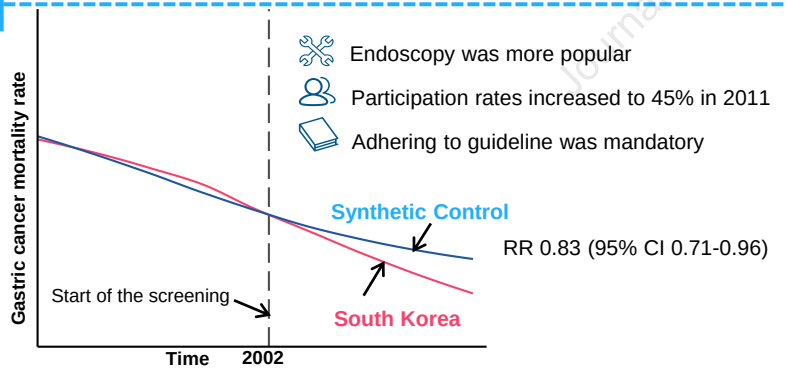
Please cite this article as: Sun D, Mülder DT, Li Y, Nieboer D, Park JY, Suh M, Hamashima C, Han W, O'Mahony JF, Lansdorp-Vogelaar I, The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality: a Synthetic Control Study, *Gastroenterology* (2023), doi: https://doi.org/10.1053/j.gastro.2023.11.286.

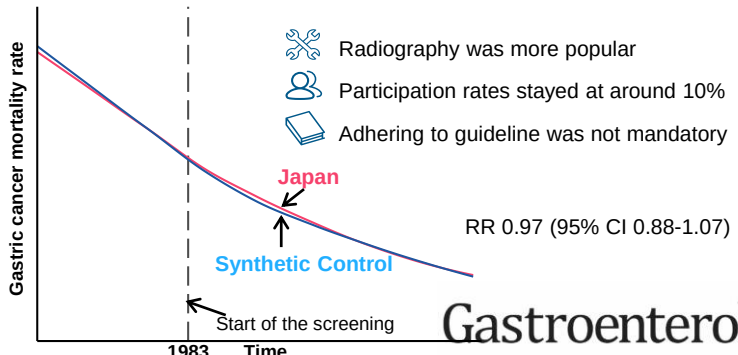
Methods



Results







Gastroenterology

Target Journal: Gastroenterology

The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality: a Synthetic Control Study

Dianqin Sun MMed,[1] Duco T. Mülder MSc,[1] Yige Li MSc,[2] Daan Nieboer PhD,[1] Jin Young Park PhD,[3] Mina Suh MD PhD,[4] Chisato Hamashima MD,[5] Weiran Han MSc,[1] James F. O'Mahony PhD,[1,6] Iris Lansdorp-Vogelaar PhD [1]

[1]Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, the Netherlands

[2]Department of Biostatistics and CAUSALab, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States

[3]Early Detection, Prevention, and Infections Branch, International Agency for Research on Cancer, Lyon, France

[4]National Cancer Control Institute, National Cancer Center, Goyang, Korea

[5]Division of Health Policy, Department of Nursing, Faculty of Medical Technology, Teikyo University, Tokyo, Japan

[6]School of Economics, University College Dublin, Dublin, Ireland

Short Title: Effect of National Gastric Cancer Screening (43, limit 45 characters with space)

Word count: 6522 words (limit 7000 words inclusive of main text, abstract, references, table/figure legends); 2 tables and 2 figures (limit 7 tables/figures)

Abbreviations: COPD, chronic obstructive pulmonary disease; GC, gastric cancer; H. pylori, Helicobacter pylori; KNCSP, Korean National Cancer Screening Program; RMSPE, root mean square percentage errors; SCM, synthetic control method; SDID, synthetic difference-in-differences; UGI, upper gastrointestinal

Conflicts of interest: The authors disclose no conflicts. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer /World Health Organization.

Funding: Dianqin Sun has been supported by the China Scholarship Council. Duco Mülder, Weiran Han, and Iris Lansdorp-Vogelaar have been supported by National Institutes of Health/National Cancer Institute grants U01CA265729. James O'Mahony is funded by Ireland's Health Research Board under grant EIA-2017-054. The funders had no role in the design, implementation, or interpretation of results of this study.

Data Transparency Statement: All data used in this study are from publicly available databases.

CRediT Authorship Contributions: **Dianqin Sun, MMed** (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Writing-original draft: Lead). **Duco T. Mülder, MSc** (Conceptualization: Supporting; Formal analysis: Supporting; Investigation: Lead; Writing - original draft: Lead). **Yige Li, MSc** (Conceptualization: Supporting; Formal analysis: Supporting; Methodology: Lead; Writing - original draft: Supporting) **Daan Nieboer, PhD** (Methodology: Supporting; Formal analysis: Supporting; Writing - original draft: Supporting). **Jin Young Park, PhD** (Investigation: Supporting; Writing - original draft: Supporting). **Mina Suh, MD PhD** (Writing - original draft: Supporting). **Chisato Hamashima, MD** (Writing - original draft: Supporting). **Weiran Han, MSc** (Writing - original draft: Supporting). **James F. O'Mahony, PhD** (Formal analysis: Supporting; Investigation: Supporting; Supervision: Lead; Writing - original draft: Supporting). **Iris Lansdorp-Vogelaar, PhD** (Conceptualization: Lead; Formal analysis: Supporting; Investigation: Lead; Methodology: Supporting; Supervision: Lead; Writing - original draft: Supporting).

Corresponding author:

Dianqin Sun, MMed

Department of Public Health, Erasmus MC, University Medical Center, Rotterdam

d.sun@erasmusmc.nl

Dr. Molewaterplein 40, 3015 GD Rotterdam

Abstract (258 words, limit 260)

Background & Aims: Nationwide organized gastric cancer (GC) screening programs have been running for decades in South Korea and Japan. This study aimed to conduct a quasi-experimental analysis to assess the population impact of these programs on GC mortality.

Methods: We used the flexible synthetic control method (SCM) to estimate the effect of the screening programs on age-standardized GC mortality and other upper gastrointestinal (UGI) diseases (esophageal cancer and peptic ulcer) among people aged 40 years or above. World Health Organization mortality data and country-level covariates from the World Bank and the Global Burden of Diseases study were used for the analyses. We compared post-intervention trends in outcome with the counterfactual trend of the synthetic control, and estimated average post-intervention rate ratios (RRs) with associated 95% confidence intervals (CIs). A series of sensitivity analyses were conducted.

Results: The pre-intervention fits were acceptable for the analyses of South Korea and Japan's GC mortality but poor for Japan's other UGI disease mortality. The average post-intervention RRs are 0.83 (95% CI 0.71-0.96) for GC mortality and 0.72 (0.57-0.90) for other UGI disease mortality in South Korea. The RR reached 0.59 by the 15th year after the initiation of nationwide screening. For Japan, the average RRs were 0.97 (0.88-1.07) for GC mortality and 0.93 (0.68-1.28) for other UGI disease mortality. Sensitivity analysis reveals the result for Japan may potentially be biased.

Conclusion: South Korea's nationwide GC screening has apparent benefits while the Japanese program's effectiveness is uncertain. The experiences of South Korea and Japan could serve as a reference for other countries.

Keywords: Stomach Cancer; Screening; Endoscopy; Upper Gastrointestinal Series

Introduction

Gastric cancer (GC) ranks as the fifth most common cancer and the fourth leading cause of cancer-related mortality worldwide. [1] Despite some improvements in treatment, the prognosis of GC is still poor as it is commonly diagnosed at an advanced stage, [2-4] underlining the need for primary prevention and early detection. [5]

Endoscopy for the early detection of GC and precancerous lesions in people at high risk is recommended by several clinical guidelines. [6-8] Based on meta-analyses of observational studies, endoscopic screening was related to an approximately 40% reduction in GC mortality risk. [9, 10] Two cohort studies in China's high-incidence areas further endorsed the effectiveness of one-time and repeated endoscopic screening in decreasing GC mortality risk. [11, 12] Evidence from randomized controlled trials (RCTs), however, is still lacking. The upper gastrointestinal (UGI) series is the other main GC screening modality, though it exhibits lower sensitivity and worse performance in organized screening in comparison to endoscopy. [13-15] A recent meta-analysis found that screening with UGI series did not lead to a statistically significant reduction in the risk of GC death. [10]

Despite the lack of evidence from RCTs, South Korea and Japan, two countries with high GC-incidence, have been at the forefront of GC secondary prevention and have implemented nationwide organized GC screening programs for decades using endoscopy or UGI series. [16, 17] Individual-level data from these programs have indicated the effectiveness of mass screening in decreasing the individual risk of GC death. [14, 18, 19] These studies, however, are susceptible to volunteer bias. Furthermore, few studies have investigated the population-level impact of nationwide organized GC screening on GC mortality.

In the absence of clear evidence on the effectiveness of GC screening, this study aimed to use a quasi-experimental design to complement existing evidence on a population-level basis. Using the synthetic control method (SCM), we assessed the changes in population GC mortality due to the start of the nationwide screening programs in South Korea and Japan.

Materials and Methods

Research setting

South Korea and Japan have both implemented national GC screening programs within their healthcare systems, though there are large differences between these two contexts. The Korean National Cancer Screening Program (KNCSP) for GC started in 1999, targeting low-income Medical Aid recipients [14] and then expanded to the National Health Insurance beneficiaries in 2002, almost covering all citizens. Since its initiation, the KNCSP has provided biennial screening for individuals aged 40 years and older with UGI series or endoscopy. [16] In Japan, locally organized GC radiographic screening began in the 1960s and expanded nationwide in 1983 following the Health Service Law for the Aged. [17] At the inception of the Japanese nationwide screening, the national government policy recommended GC screening for people aged 40 years and older each year, using UGI series[17]. However, the revised 2014 Japanese guidelines restricted the target population to those aged 50 years and older and recommended either UGI series or endoscopic screening every two years[6]. Please see **Supplementary Table S1** for more details about these two countries' screening tools, program organization, and participation rates.

Study design

We used the SCM[20-22] to estimate the effects of nationwide screening programs on gastric cancer mortality in South Korea and Japan. The concept of SCM is to construct a synthetic control for the treated country by deriving a weighted average of multiple control countries without intervention. In this case, South Korea and Japan are the countries with the intervention of nationwide GC screening. The weight of controls is determined in a data-driven way to minimize the differences in pre-intervention outcomes (i.e. GC mortality prior to the introduction of nationwide screening) and other covariates associated with GC mortality between the treated country and the synthetic control. If a good pre-intervention fit is achieved, the difference in the post-intervention trend can plausibly be attributed to the nationwide screening program.

The SCM was first proposed by Abadie & Gardeazabal[22] in 2003 and has been widely applied in public health intervention research. [23, 24] The SCM is designed to interpolate the outcomes of control countries to avoid unrealistic counterfactuals. This means that the outcomes of the treated units must remain within the distribution of the control data, as known as the "convex hull" condition. South Korea and Japan, however, have very high gastric cancer mortality rates globally[25], which fail to meet the convex hull

requirement for interpolation. [26] Therefore, we applied the flexible SCM approach developed by Bonander. [27] The flexible SCM relaxes the constraint that weights sum to one to permit extrapolation when necessary but applies an accompanying penalization. In this way, this alternative relaxes the convex hull condition and avoids potentially unrealistic estimates. Please see **Supplementary Methods** for technical details about the flexible SCM. This study was exempt from ethical review, as it only utilized aggregated and publicly accessible data.

Outcome data

The primary outcome in the SCM was age-standardized GC (International Classification of Diseases code: C16) mortality among people aged 40 years or above. Our secondary outcome was age-standardized mortality rates of other UGI diseases, including esophageal cancer (C15) and peptic ulcer (K25-K27), which might be detected during GC screening. We did not include gastritis and duodenitis (K29) since mortality data were unavailable before 1994. We combined mortality rates of esophageal cancer (C15) and peptic ulcer (K25-K27) because deaths from these causes are rare and therefore susceptible to random fluctuations. [20]

We extracted age-specific death rates from the World Health Organization's (WHO) mortality database[28] and then calculated age-standardized rates using the age structure from the WHO standard population. South Korea and Japan's GC mortality records were available from 1985 to 2019 and 1950 to 2019 respectively, though the most recent years of data were missing for many other countries. Covariate data were also not available for many counties in earlier years. Thus, to balance the pre-intervention period length and the number of countries to include as potential control countries (donor pools), we determined our study period as 1985 to 2017 for South Korea and 1965 to 2017 for Japan. Accordingly, approximately 20 pre-intervention years of data were available for both countries.

Covariates data

Our selection of potential covariates associated with GC mortality was based on prior literature and relevant guidelines [29-31], as well as data availability: covariates needed at least one observation in each third of the pre-intervention period to allow proper functioning of the flexible SCM (see Statistical Analysis section for more details). This led to eight country-level indicators of demographic (rurality and education), socioeconomic development (per capita gross national income), health service capacity (unsafe water sanitation and handwashing), and GC risk factor (high sodium diet, tobacco use, alcohol consumption, overweight) prevalence as covariates in the analysis for South Korea. We included only two demographic covariates (rurality and education) in the analysis for Japan due to the limited data available for the other covariates in the early pre-intervention period (1965-1975). **Table S2** shows the sources, definitions, and years of available data for these covariates.

Synthetic control group selection

Countries were included in the donor pool if they met the following criteria: 1) having outcome data with a missing percentage of less than 10% during the study period, 2) having covariate data during the years listed in **Table S2**, 3) being classified as upper-middle-income or high-income countries as defined by the World Bank. To improve the homogeneity between control and treated countries, we used the income classification to restrict the donor pool in the intervention year 2002 for South Korea. Since the income classification was not available for 1983, we applied the country category in the nearest available year (1987) to select control countries for Japan. Japan and South Korea were excluded from donor pools. **Figure S1** shows the selection procedure for the donor pools. The final set of control countries constructing the synthetic control is selected from the donor pool in a data-driven process, and the weight is determined by the algorithm as described in **statistical analysis** section.

Statistical analysis

Data imputation

If mortality rates of specific age groups in certain countries and certain years were missing from the WHO database, but age-standardized rates for the whole population and age-specific rates for people aged under 40 years were available, we calculated age-standardized rates for 40 years and older based on these data. For countries with less than 10% missing mortality rates, the missing records were imputed using spline interpolation.

SCM

Like prior studies [32-34] that used SCM to estimate multiple outcomes, we conducted separate analyses for the primary and secondary outcomes. Similarly, analyses were stratified for South Korea and Japan owing to the substantial differences between the two programs and the country contexts. First, we built the synthetic control following the three steps. [27] 1), covariates associated with the outcome were included to predict the post-intervention trend and avoid overfitting the pre-intervention outcome. The relative importance of covariates and the lagged outcome was determined to prioritize balance on those with better predictive power. The mean values of covariates over available years were used for the analysis. 2), the original control weight is calculated by the original SCM algorithm (**Equation S1** in the **Supplementary Methods**). 3), the relaxed control weight was determined by adding a modified loss function that allows for extrapolation and penalizes the sum of squared differences from the original weights (**Equation S2** in the **Supplementary Methods**). The final synthetic control was then constructed as the weighted average of control counties based on the relaxed weight.

Second, we plotted the observed and counterfactual trends of the outcome based on the average from the synthetic control group to estimate the effects of nationwide GC screening. We employed segmented regression models to examine trends before and after nationwide screening for both synthetic control and treated units. The segmented regression models were fitted by linear regressions, including a continuous variable for time since the observation, a dummy indicator for before/after screening, and a continuous variable for time after screening. To address the first-order autoregressive serial correlation of errors, we used the Prais-Winsten estimator. [35] In addition, we assessed pre-intervention fit by plotting the difference in the observed outcomes and their synthetic controls for all pre-intervention years.

Finally, we calculated the average post-intervention rate ratios (RRs) and 95% confidence intervals (CIs) using the cross-fitting method proposed in a pre-print article by Chernozhukov et al. [36] Following the recommendations of Chernozhukov et al., we chose the three-fold cross-fitting to balance efficiency and accuracy of the estimates. In short, the pre-intervention phase was divided into training and holdout samples. In each iteration, two-thirds of the pre-intervention period was used to train the model as the training sample. The holdout sample of the remaining third was used to assess the bias of effect estimate during each iteration. Then the final average debiased estimates and the associated CI were calculated

as described by Bonander.[27] We also calculated the rate difference (RD) and RRs in each year during the post-intervention period. As an alternative to establish the intervention effect, we performed a placebo examination to evaluate the probability that the observed effect for South Korea and Japan could occur by chance. Specifically, we repeated the SCM for each control country as if it were the intervention country and plotted the difference between the observed and counterfactual trends for each control country, South Korea, and Japan. Then we calculated and ranked the ratios of post-intervention to pre-intervention root mean square percentage errors (RMSPE) for each control country, South Korea, and Japan. A higher rank of the post-/pre- intervention RMSPE ratio indicates a lower probability that the observed effect occurs at random. In addition, to objectively assess the models' performance regarding pre-intervention fit, we calculated the RMSPE between predicted and observed outcomes during the pre-intervention period, as well as the average RMSPE during holdout periods in the cross-fitting procedure.

Sensitivity analysis

We conducted several sensitivity analyses to test the robustness of our results. First, we repeated the analyses using the synthetic difference-in-differences (SDID) method.[37] While still being developed, the novel SDID method offers a promising alternative for conducting robust checks in our study. The SDID avoids the convex hull requirement by combining SCM and the difference-in-differences methods. Unlike the SCM, the SDID adjusts time-varying covariates during the study period directly rather than seeking to match the pre-intervention average values of covariates. The SDID method estimated the average post-intervention RD as the effect measure. Second, we performed two negative control outcome analyses, by repeating the SCM in two settings in which we would not expect a gastric screening program to make an impact. We conducted SCM with the same input variables among people aged 35 to 39, who did not undergo screening, and for chronic obstructive pulmonary disease (COPD, J40-J44), which is not expected to be impacted by a GC screening program. If in these settings no difference in post-intervention trend is seen compared to the synthetic control, this adds confidence that the observed differences in the main analyses trend are associated with the gastric cancer screening program. Third, we conducted the leave-one-out examination to test the influence of each single control country's weight.

Data analyses were performed using R version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria). See **Supplementary Methods** for more details about the sensitivity analyses.

Results

South Korea

The donor pool for South Korea consisted of 37 control countries (**Figure S1**). Chile, Costa Rica, Singapore, and Italy made up the synthetic control for GC mortality in South Korea (**Table 1**), with Chile being the most important contributor. For the analysis of other UGI disease mortality, Czechia, Greece, Mauritius, Poland, and Singapore were used to construct the synthetic control. The synthetic control resulted in reasonably good pre-intervention fits for GC and other UGI disease mortality in South Korea (**Figure 1**, **Figure 2, Figure S2**). In particular, the difference was mostly centered around zero between 1996 and 2002. For most covariates, the values of the synthetic control resembled those of South Korea more closely than the corresponding donor pool averages (**Table S3**).

The differences between the synthetic control and South Korea during the pre-intervention period were relatively small compared to the differences after the implementation of nationwide screening (**Figure 1**, **Figure 2**). The average post-intervention RRs are 0.83 (95% CI 0.71-0.96) for GC mortality and 0.72 (95% CI 0.57-0.90) for other UGI disease mortality. The beneficial effect on GC mortality and UGI disease mortality increased over the post-intervention period (2002-2017), with the RR reaching 0.59 and 0.47 respectively, by the 15th year following the introduction of nationwide screening (**Table S4**). Mortality trends for GC and other UGI diseases were the same between the synthetic control and South Korea prior to screening. These trends diverged after the start of screening (**Figure S3**). The post-intervention differences between the synthetic control and South Korea became distinct from the placebo tests after 2008. Consistent with the clear post-intervention differences, the ratios of post-intervention to pre-intervention RMSPEs ranked top among the placebo tests (**Figure S4**).

Japan

The GC mortality in Japan was reproduced as the weighted combination of Venezuela, Singapore, Uruguay, Malta, and Italy. For other UGI disease mortality, Iceland, Portugal, Singapore, Switzerland, and the United States made up the synthetic control. The pre-intervention fits were acceptable for GC mortality but poor for other UGI disease mortality (**Figure 1**, **Figure 2, Figure S2**).

In contrast to South Korea, no statistically significant differences in post-intervention levels were observed between Japan and the synthetic control, with the average RRs of 0.97 (95% CI 0.88-1.07) for GC mortality and 0.93 (95% CI 0.68-1.28) for other UGI disease mortality between 1983 and 2017 (**Figure 1**, **Figure 2**). Also, no apparent increase in the difference in GC and UGI mortality rates was observed between Japan and its synthetic control across the post-intervention period (**Table S5**). Mortality trends in GC between the synthetic control and Japan remained consistent before and after screening (**Figure S3**). Correspondingly, the post-/pre-intervention ratios of RMSPEs in the placebo analysis indicated no post-intervention outcome differences between the synthetic control and Japan (**Figure S5**).

Sensitivity analyses

Table 2 summarizes the effect estimates from the sensitivity analyses. For the SDID analysis of South Korea, the pre-intervention parallel trend was not satisfied for GC mortality but acceptable for other UGI disease mortality. Findings of SDID indicated that nationwide GC screening was associated with an average decrease of 22.70 deaths (95% CI -28.82 to -16.79) in GC and 3.86 deaths (95% CI -6.19 to -0.89) in other UGI diseases per 100,000 people across the post-intervention period (**Figure S6, Figure S7**). The effect for GC mortality estimated by SDID was larger than that by SCM, while the results for other UGI disease mortality were similar between these two methods (**Table S4**). For Japan, the pre-intervention parallel trend was acceptable for GC mortality but poor for other UGI disease mortality. Unlike the results of SCM, the SDID study indicated that the national program in Japan induced an average decrease of 20.41 (95% CI -29.71 to -11.05) deaths in GC per 100,000 people across the post-intervention period (**Figure S4**). Negative control outcome analyses showed no post-intervention differences in both GC mortality in 35-39 year-olds and COPD mortality in South Korea (**Figure S8-**

S10). Conversely, differences in post-intervention trends for these negative controls were seen for Japan. Excluding each weighted control country did not change the results of the base case analyses substantially for either South Korea or Japan (**Figure S11-S14**). After removing Chile from the analysis for South Korea, the estimated RR changed and was no longer statistically significant (**Figure S11**). However, this result cannot be considered reliable due to the inadequate pre-intervention fit.

Discussion

This study assessed the population impact of the national GC screening programs in South Korea and Japan. We found that the national program in South Korea was associated with a reduction in GC mortality of 41% and other UGI disease mortality of 53% by the 15th year following the start of the program. No effect was found in Japan in the base-case SCM analysis, although the supplementary SDID analysis indicated a reduction in GC mortality in Japan after the introduction of screening. The effects for South Korea were robust across different settings, while the results of Japan had relatively large uncertainties due to potential bias indicated by negative control outcome analysis.

By concentrating on the two early adopters, we not only provide novel evidence on the effectiveness of organized GC screening but also evidence on the varying degrees of benefits between South Korea and Japan. The negative control outcome analysis identified an unexpected decreasing effect on COPD mortality after the introduction of GC screening in Japan. This suggested that certain interventions targeting COPD risk factors might coincide with the start of the nationwide screening. Given the shared risk factor of smoking between COPD and GC, the analysis for Japan could be susceptible to bias, potentially leading to an overestimation of screening's advantageous impact. While the bias resulting from not accounting for opportunistic screening rates also tended to favor a beneficial effect in Japan, the SCM analysis still did not find such an effect. One possible explanation for the different effects estimated in South Korea and Japan is the various GC screening circumstances in these two countries (**Table S1**). Although Japan started nationwide screening about 20 years earlier than South Korea, the Japanese GC screening guidelines did not recommend endoscopic screening until 2014. Previous studies have demonstrated that endoscopy has higher sensitivity than the UGI series in organized screening. [10, 13-15] But only 19% of municipalities in Japan adopted endoscopic screening in 2015, while 72.55% of the participants in South Korea chose endoscopic screening instead of UGI series in 2011. [16, 38] In addition to the difference in the coverage of endoscopic screening, guideline adherence is lower in Japan than in South Korea. [39] The participation rate of national organized GC screening is much lower in Japan than in South Korea. [40] Even when combined with opportunistic screening, Japan's participation rates in

GC screening still fall short of those in South Korea. [16, 41] Therefore, the findings in our study may have been expected. However, it is important to note that certain covariates were unavailable for the analysis in Japan, which may have introduced potential biases, the directions of which are unclear. Further studies are needed to compare the screening impact in South Korea and Japan.

SDID analysis tended to estimate larger effect sizes than the SCM in our study, indicating a degree of uncertainty around the evidence. The differences between results from SDID and SCM might be due to the fact that the SDID adjusts the time-varying covariates directly. The time-varying nature of covariates should be treated cautiously in our analysis due to the long study period. Besides, the relative effect on other UGI disease mortality was more pronounced than on GC mortality in our study. This suggests that early detection might be more effective for patients with esophageal cancer or peptic ulcers. A cohort study investigating one-time endoscopic screening found that the effect on esophageal cancer mortality (RR 0.49, 95% CI 0.43-0.56) was more pronounced than on cardia gastric cancer (RR 0.69 95% CI 0.56-0.85). [11]

Our results support the findings of a beneficial effect of organized GC screening on GC mortality in previous observational studies. [9, 11, 12] One Chinese cohort study with 637500 participants concluded that one-time endoscopic screening could decrease 42% mortality risk in cardia GC and 62% in non-cardia GC. [11] Another large cohort observed the intensifying beneficial effects of repeated endoscopic screening on GC mortality after the one-time screening. [12] Observational studies nested in the national programs in South Korea and Japan also reported a reduction in the risk of GC mortality among participants receiving screening compared with the control population. [14, 18, 19] The changes in GC stage distribution and survival in South Korea and Japan suggested the potential population benefit of the expansion of GC screening. The percentage of gastric cancers diagnosed at an early stage in South Korea has grown from 39% in 2001 to 73% in 2016. [42] During the same period, the five-year GC survival increased from 48.6% to 68.9% in South Korea and grew from 50.5% to 60.3% in Japan. [3] These encouraging changes, however, cannot be directly ascribed to the nationwide screening due to the absence of comparable controls. Our study constructs synthetic controls for these two countries and address the evidence limitations on the population impact of organized CG screening.

In addition to the effect on GC mortality, we also found that nationwide screening in South Korea might decrease the mortality of other UGI diseases, including esophageal cancer and peptic ulcers. A cluster randomized trial in China indicated that endoscopic screening could decrease esophageal cancer mortality with a hazard ratio of 0.45 (95% CI 0.54-0.95). [43] Besides, GC screening via UGI series or

endoscopy can detect silent ulcers, accounting for a major proportion of ulcers[44], before developing complications. Complications of peptic ulcers, such as bleeding, are the leading causes of peptic ulcer mortality. The enhancement of living conditions, the advent of proton pump inhibitors, and the application of effective *Helicobacter pylori* (*H. pylori*) eradication therapies have collectively reduced peptic ulcer mortality over time, as seen in mortality trends in South Korea, Japan, and other nations. The influences of these factors on South Korea and its synthetic control, as well as Japan and its synthetic control, were expected to be similar, as reflected by the acceptable pre-intervention fits in both cases. Therefore, the observed differences during the post-intervention period might largely be attributed to GC screening. The impact of endoscopic or radiographic screening on peptic ulcer mortality remains unexplored but is crucial for the economic evaluation of GC screening, and further research is necessary. However, conducting such a study could be challenging due to the low incidence of peptic ulcer deaths.

Our study has some limitations. First, we did not include the prevalence of *H. pylori* infection as a covariate due to the limited data availability. Globally, most countries have seen a decline in *H. pylori* infections, [45] including South Korea[46] and Japan[47]. However, few countries have conducted multi-round surveys on infection rates, making direct cross-country comparisons challenging. Japan has taken a proactive stance in combatting *H. pylori*. In 2013, it became the first country to provide national health insurance coverage for eradication therapy in gastritis patients associated with *H. pylori*. This led to a rise in eradication cases. [48] Considering the time lag for *H. pylori* eradication's impact on GC risk, the effect on GC mortality rates likely exceeds our study period and therefore not influenced our results. [47] [49] [49]Second, we did not have indicators of opportunistic screening rates. Opportunistic GC screening rates in South Korea and Japan have increased following the start of their national screening programs[16, 50], which might exaggerate the estimated effect in our study. However, to some extent, the beginning of the nationwide GC screening programs might have affected the opportunistic screening rates by establishing the screening guidelines, increasing public awareness, and strengthening the infrastructure, and this could be seen as an indirect effect of national programs. Third, some covariates were not available in the analysis for Japan. SCM and SDID also generated conflicting results for Japan. Therefore, more investigation is required to evaluate the impact of national GC screening programs in Japan. Fourth, Chile has a large weight in the GC mortality analysis for South Korea, rendering the results vulnerable to the influence of the single country. After excluding Chile from the analysis, the pre-intervention fit deteriorated, making the outcome insufficient to provide reliable information. Fifth, limited data availability hinders including countries more comparable to Korea and Japan. Nonetheless, we have tried to enhance homogeneity by only including upper-middle-income or high-income countries in the donor pool.

The screening programs in South Korea and Japan provide practical evidence for other countries with high GC burdens. The observed effect, however, might not be easily generalized to other settings, especially considering that disparities between South Korea and Japan also exist. The quality of GC screening influences the effectiveness of screening programs, and Korean and Japanese gastroenterologists are known to perform better in detecting early-stage GC and precancerous lesions when compared to specialists in western countries. [51, 52] Nevertheless, population-level benefits identified in this study can still be expected in other countries with high GC burdens if high-quality screening is achieved through training and quality control. Implementing such nationwide screening programs will probably not be cost-effective for western European and North American countries as their GC incidence is low. Besides, it is important to note that endoscopic screening is not without potential harm. Complications such as bleeding or perforation can occur, albeit rarely. Nevertheless, also for these countries, this study now lends some support to the provision of screening to their high-risk populations, for example, those with precursor lesions[7, 8, 53].

In conclusion, our findings provide an essential addition to the existing evidence on GC screening effectiveness from a population-level perspective. South Korea's nationwide GC screening has apparent benefits on the reduction of GC mortality while the Japanese program's effectiveness was uncertain. The effects for South Korea were robust across different specifications and analyses, while the results of Japan were susceptible to bias and warrant more investigation. The disparities in screening programs between South Korea and Japan suggest that the effectiveness of GC screening might be influenced by factors like screening modality, participation rates, and organizational strategies. This highlights the significance of a well-planned organizational structure and evidence-based decision making when starting the organized screening. With a quasi-experimental design, this study will facilitate triangulating current observational evidence and provide valuable insights while the GC screening RCTs[54, 55] are still underway. The data and experience from South Korea and Japan will inform GC screening policy in other countries.

References

1. 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;71:209-249.
2. Arnold M, Rutherford MJ, Bardot A, et al. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–2014 (ICBP SURVMARK-2): a population-based study. Lancet Oncol 2019;20:1493-1505.
3. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391:1023-1075.
4. Arnold M, Morgan E, Bardot A, et al. International variation in oesophageal and gastric cancer survival 2012-2014: differences by histological subtype and stage at diagnosis (an ICBP SURVMARK-2 population-based study). Gut 2022;71:1532-1543.
5. Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet 2020;396:635-648.
6. Hamashima C, Systematic Review G, Guideline Development Group for Gastric Cancer Screening G. Update version of the Japanese Guidelines for Gastric Cancer Screening. Jpn J Clin Oncol 2018;48:673-683.
7. Banks M, Graham D, Jansen M, et al. British Society of Gastroenterology guidelines on the diagnosis and management of patients at risk of gastric adenocarcinoma. Gut 2019;68:1545-1575.
8. Pimentel-Nunes P, Libanio D, Marcos-Pinto R, et al. Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European Helicobacter and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019. Endoscopy 2019;51:365-388.
9. **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;155:347-354 e9.
10. Hibino M, Hamashima C, Iwata M, et al. Radiographic and endoscopic screening to reduce gastric cancer mortality: a systematic review and meta-analysis. Lancet Reg Health West Pac.
11. **Chen R, Liu Y, Song GH**, 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;70:251-260.
12. **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 2022;54:848-858.
13. Hamashima C, Shabana M, Okada K, et al. Mortality reduction from gastric cancer by endoscopic and radiographic screening. Cancer Sci 2015;106:1744-1749.

14. Jun JK, Choi KS, Lee HY, et al. Effectiveness of the Korean National Cancer Screening Program in Reducing Gastric Cancer Mortality. Gastroenterology 2017;152:1319-1328 e7.
15. Hamashima C, Okamoto M, Shabana M, et al. Sensitivity of endoscopic screening for gastric cancer by the incidence method. Int J Cancer 2013;133:653-659.
16. Lee S, Jun JK, Suh M, et al. Gastric cancer screening uptake trends in Korea: results for the National Cancer Screening Program from 2002 to 2011: a prospective cross-sectional study. Medicine (Baltimore) 2015;94:e533.
17. Hamashima C. Cancer screening guidelines and policy making: 15 years of experience in cancer screening guideline development in Japan. Jpn J Clin Oncol 2018;48:278-286.
18. Hamashima C, Ogoshi K, Narisawa R, et al. Impact of endoscopic screening on mortality reduction from gastric cancer. World J Gastroenterol 2015;21:2460-6.
19. Matsumoto S, Yamasaki K, Tsuji K, et al. Results of mass endoscopic examination for gastric cancer in Kamigoto Hospital, Nagasaki Prefecture. World J Gastroenterol 2007;13:4316-20.
20. Bonander C, Humphreys D, Degli Esposti M. Synthetic Control Methods for the Evaluation of Single-Unit Interventions in Epidemiology: A Tutorial. Am J Epidemiol 2021;190:2700-2711.
21. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. J Am Stat Assoc 2010;105:493-505.
22. Abadie A, Gardeazabal J. The economic costs of conflict: A case study of the Basque Country. Am Econ Rev 2003;93:113-132.
23. Craig P, Katikireddi SV, Leyland A, et al. Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research. Annu Rev Publ Health 2017;38:39-56.
24. Abadie A, Cattaneo MD. Econometric Methods for Program Evaluation. Annu Rev Econ 2018;10:465-503.
25. Lin Y, Zheng Y, Wang HL, et al. Global Patterns and Trends in Gastric Cancer Incidence Rates (1988-2012) and Predictions to 2030. Gastroenterology 2021;161:116-127 e8.
26. Abadie A. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. J Econ Lit 2021;59:391-425.
27. Bonander C. A (Flexible) Synthetic Control Method for Count Data and Other Nonnegative Outcomes. Epidemiology 2021;32:653-660.

[dataset] 28. World Health Organization. WHO Mortality Database, WHO Data Platform, 2023. https://platform.who.int/mortality.

29. Uthman OA, Jadidi E, Moradi T. Socioeconomic position and incidence of gastric cancer: a systematic review and meta-analysis. J Epidemiol Community Health 2013;67:854-60.
30. World Cancer Research Fund International. Diet, nutrition, physical activity and cancer: a global perspective (the Third Expert Report): World Cancer Research Fund International, 2018.

31. Rota M, Alicandro G, Pelucchi C, et al. Education and gastric cancer risk-An individual participant data meta-analysis in the StoP project consortium. Int J Cancer 2020;146:671-681.
32. Rieger M, Wagner N, Bedi AS. Universal health coverage at the macro level: Synthetic control evidence from Thailand. Soc Sci Med 2017;172:46-55.
33. Cohn E, Chimowitz M, Long T, et al. The effect of a proof-of-vaccination requirement, incentive payments, and employer-based mandates on COVID-19 vaccination rates in New York City: a synthetic-control analysis. Lancet Public Health 2022;7:e754-e762.
34. Wigley S, Dieleman JL, Templin T, et al. Autocratisation and universal health coverage: synthetic control study. BMJ 2020;371:m4040.
35. Prais GJ, Winsten CB. Trend Estimates and Serial Correlation. Cowles Commission Discussion Paper. Chicago, 1954.
36. Chernozhukov V, Wuthrich K, Zhu Y. Practical and robust t-test based inference for synthetic control and related methods: arXiv preprint, 2019.
37. Arkhangelsky D, Athey S, Hirshberg DA, et al. Synthetic Difference-in-Differences. Am Econ Rev 2021;111:4088-4118.
38. Mabe K, Inoue K, Kamada T, et al. Endoscopic screening for gastric cancer in Japan: Current status and future perspectives. Dig Endosc 2022;34:412-419.
39. Goto R, Hamashima C, Mun S, et al. Why screening rates vary between Korea and Japan--differences between two national healthcare systems. Asian Pac J Cancer Prev 2015;16:395-400.
40. Sano H, Goto R, Hamashima C. What is the most effective strategy for improving the cancer screening rate in Japan? Asian Pac J Cancer Prev 2014;15:2607-12.
41. National Cancer Center Cancer Information Service. Cancer screening rates in Japan. Volume 2023, 2023.
42. Huang RJ, Koh H, Hwang JH, et al. A Summary of the 2020 Gastric Cancer Summit at Stanford University. Gastroenterology 2020;159:1221-1226.
43. Wei WQ, Chen ZF, He YT, et al. Long-Term Follow-Up of a Community Assignment, One-Time Endoscopic Screening Study of Esophageal Cancer in China. J Clin Oncol 2015;33:1951-7.
44. Lu CL, Chang SS, Wang SS, et al. Silent peptic ulcer disease: frequency, factors leading to "silence," and implications regarding the pathogenesis of visceral symptoms. Gastrointest Endosc 2004;60:34-8.
45. Li Y, Choi H, Leung K, et al. Global prevalence of Helicobacter pylori infection between 1980 and 2022: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2023;8:553-564.

46. Lim SH, Kim N, Kwon JW, et al. Trends in the seroprevalence of Helicobacter pylori infection and its putative eradication rate over 18 years in Korea: A cross-sectional nationwide multicenter study. PLoS One 2018;13:e0204762.
47. Hirayama Y, Kawai T, Otaki J, et al. Prevalence of Helicobacter pylori infection with healthy subjects in Japan. J Gastroenterol Hepatol 2014;29 Suppl 4:16-9.
48. Hiroi S, Sugano K, Tanaka S, et al. Impact of health insurance coverage for Helicobacter pylori gastritis on the trends in eradication therapy in Japan: retrospective observational study and simulation study based on realworld data. Bmj Open 2017;7.
49. **Hooi JKY, Lai WY, Ng WK**, et al. Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. Gastroenterology 2017;153:420-429.
50. Hong S, Lee YY, Lee J, et al. Trends in Cancer Screening Rates among Korean Men and Women: Results of the Korean National Cancer Screening Survey, 2004-2018. Cancer Res Treat 2021;53:330-338.
51. Coda S, Lee SY, Gotoda T. Endoscopic Mucosal Resection and Endoscopic Submucosal Dissection as Treatments for Early Gastrointestinal Cancers in Western Countries. Gut Liver 2007;1:12-21.
52. Veitch AM, Uedo N, Yao K, et al. Optimizing early upper gastrointestinal cancer detection at endoscopy. Nat Rev Gastro Hepat 2015;12:660-667.
53. Huang RJ, Epplein M, Hamashima C, et al. An Approach to the Primary and Secondary Prevention of Gastric Cancer in the United States. Clin Gastroenterol Hepatol 2022;20:2218-2228 e2.
54. Zeng HM, Sun KX, Cao MM, 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.
55. 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;18:605-611.

Author names in bold designate shared co-first authorship.

Figure legends

Figure 1. Trends in age-standardized gastric cancer mortality in South Korea (A) and Japan (B) versus their synthetic control countries as well as the intervention effects and placebo tests (C, D). The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening. In panels C and D, the red lines denote the estimated intervention effects in South Korea and Japan, respectively. Meanwhile, the grey lines correspond to the effects estimated in placebo tests, where each control country was treated as the unit of interest. RR indicates rate ratio. CI indicates confidence interval.

Figure 2. Trends in age-standardized mortality of other upper gastrointestinal diseases in South Korea (A) and Japan (B) versus their synthetic control countries as well as the intervention effects and placebo tests (C, D). The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening. In panels C and D, the red lines denote the estimated intervention effects in South Korea and Japan, respectively. Meanwhile, the grey lines correspond to the effects estimated in placebo tests, where each control country was treated as the unit of interest. RR indicates rate ratio. CI indicates confidence interval.

Table 1. The control countries contributing to synthetic controls and their corresponding weights[a]

Country	Country weights for South Korea		Country weights for Japan	
	Gastric cancer mortality	Other UGI disease mortality	Gastric cancer mortality	Other UGI disease mortality
Chile	0.90	—	—	—
Costa Rica	0.29	—	—	—
Czechia	—	0.29	—	—
Greece	—	0.04	—	—
Iceland	—	—	—	0.02
Italy	0.02	—	0.05	—
Malta	—	—	0.15	—
Mauritius	—	0.02	—	—
Poland	—	0.45	—	—
Portugal	—	—	—	0.22
Singapore	0.05	0.17	0.56	0.30
Switzerland	—	—	—	0.39
Uruguay	—	—	0.51	—
United States	—	—	—	0.16
Venezuela	—	—	0.63	—

UGI disease: upper gastrointestinal disease

[a]: The constraint that the weights sum up to one is relaxed in the flexible synthetic control method.

Table 2. Post-intervention period average rate ratios and their 95% confidence intervals in sensitivity analyses for South Korea and Japan

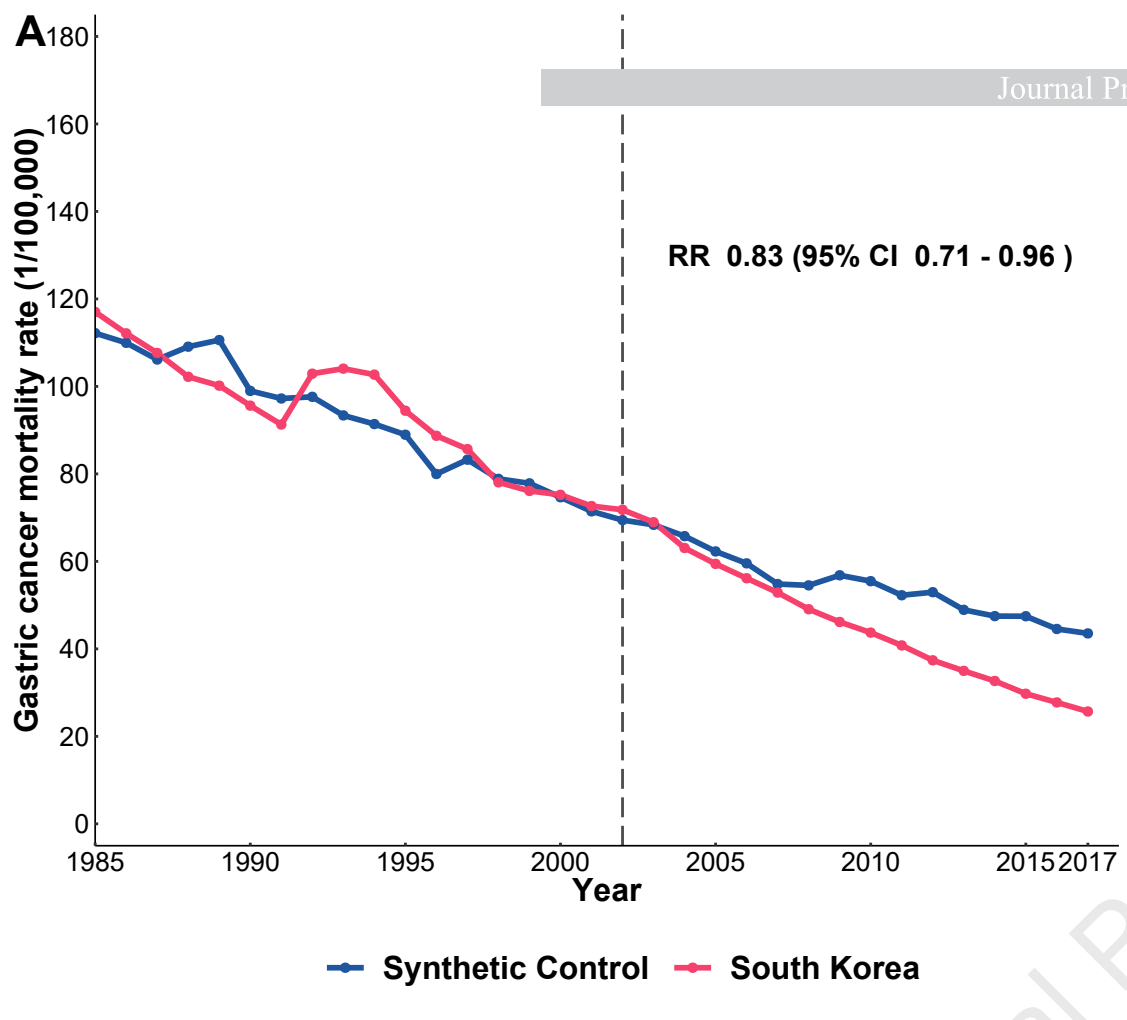
Types of sensitivity analyses	Gastric cancer mortality		Other UGI disease mortality	
	South Korea	Japan	South Korea	Japan
Synthetic difference-in-difference analysis (risk differences per 100,000 and 95% confidence intervals)	-22.70 (-28.82, -16.79) [a]	-20.41 (-29.71, -11.05)	-3.86 (-6.19, -0.89)	-1.92 (-11.07,7.32) [a]
Negative control age band (35-39 years) analysis	0.69 (0.47-1.01) [a,b]	0.70 (0.52-0.94) [a]	0.24 (0.06-0.79) [a,b,c]	0.68 (0.17-2.73) [a,b]
Negative control outcome analysis (Chronic obstructive pulmonary disease)	1.31 (0.50-3.41) [a,b,c]	0.46 (0.29-0.74) [a,b]	1.31 (0.50-3.41) [a,b,c]	0.46 (0.29-0.74) [a,b]
Leave-out-one examination	Exclude CHL: 1.02 (0.70-1.49) [a] Exclude CRI: 0.73 (0.65-0.81) Exclude ITA: 0.84 (0.70-0.99) Exclude SGP: 0.83 (0.73-0.94)	Exclude ITA: 0.97 (0.88-1.07) Exclude MLT: 0.98 (0.94-1.03) Exclude SGP: 0.98 (0.93-1.02) Exclude URY: 0.98 (0.90-1.06) Exclude VEN: 0.97 (0.85-1.12)	Exclude CZE: 0.78 (0.63-0.96) Exclude GRC: 0.72 (0.55-0.96) [a] Exclude MUS: 0.75 (0.56-1.00) [a] Exclude POL: 0.76 (0.66-0.87) Exclude SGP: 0.69 (0.65-0.73)	Exclude CHE: 0.88 (0.56-1.38) [a] Exclude ISL: 0.94 (0.69-1.28) [a] Exclude PRT: 0.88 (0.72-1.08) [a] Exclude SGP: 0.82 (0.62-1.08) [a] Exclude USA: 0.93 (0.68-1.28) [a]

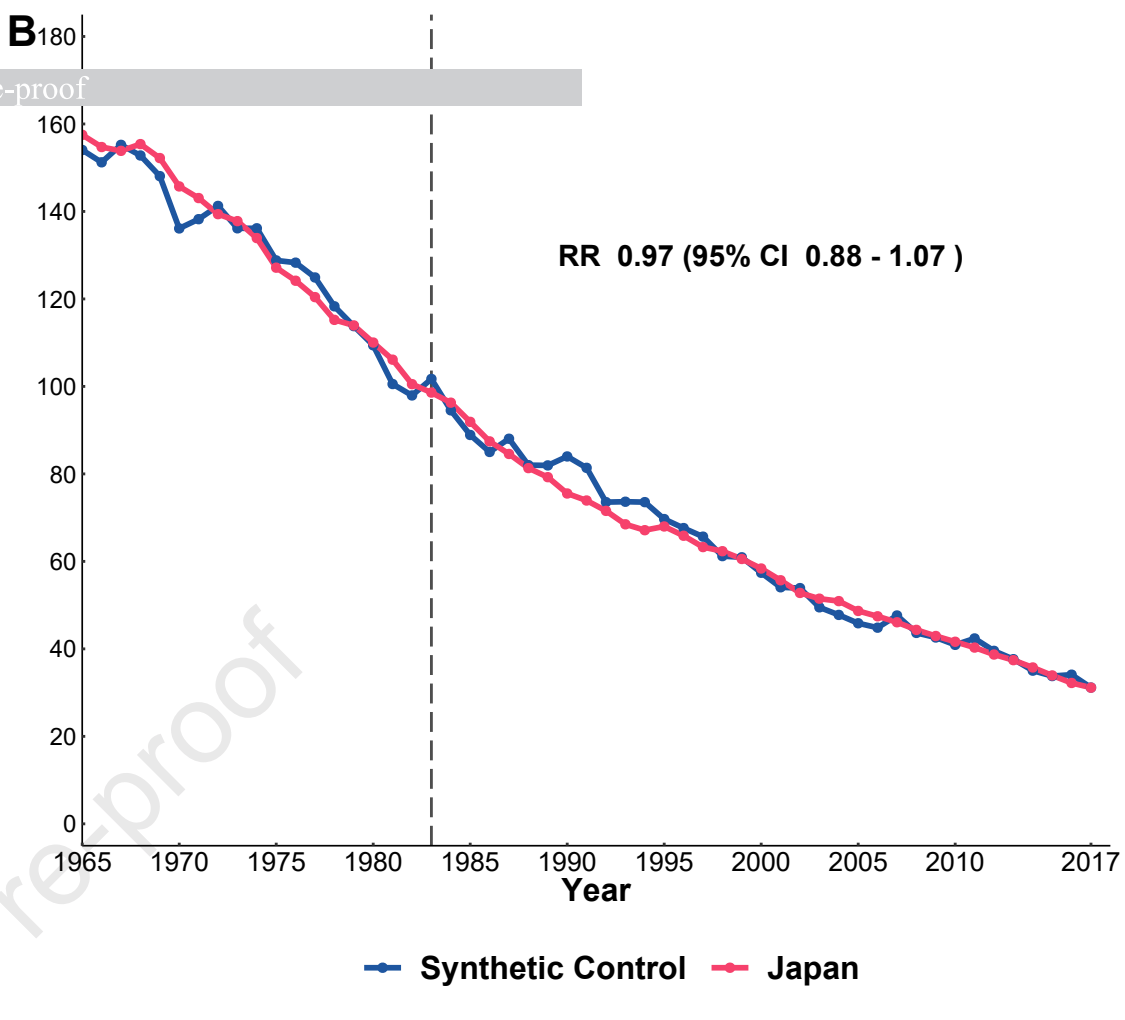
UGI disease: upper gastrointestinal disease; CHL: Chile; CRI: Costa Rica; CZE: Czechia; GRC: Greece; ISL: Iceland; ITA: Italy; MLT: Malta; MUS: Mauritius; POL: Poland; PRT: Portugal; SGP: Singapore; CHE: Switzerland; URY: Uruguay; USA: United States; VEN: Venezuela

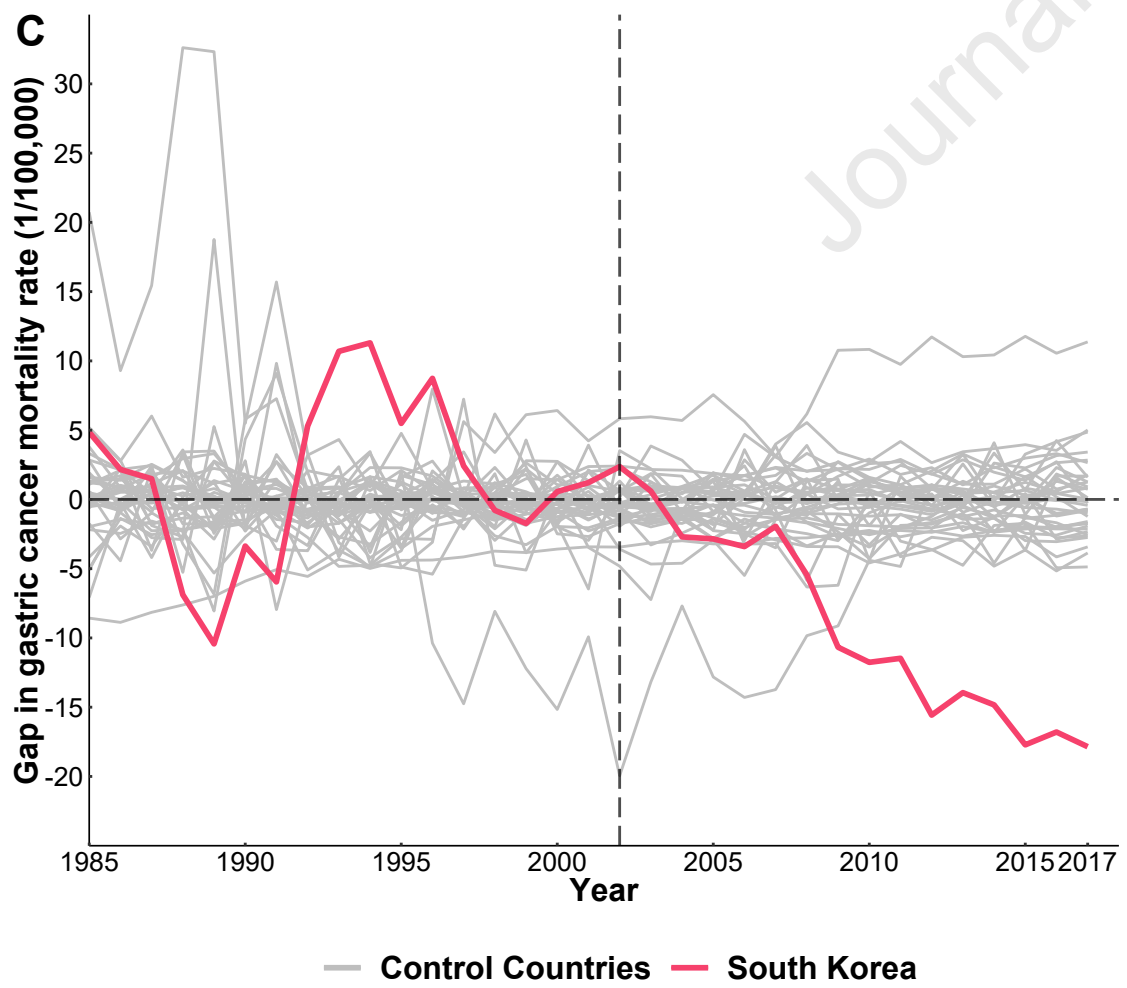
[a]: The effect estimate should be explained cautiously (the average root mean squared percentage errors from the cross-fitting>0.1).

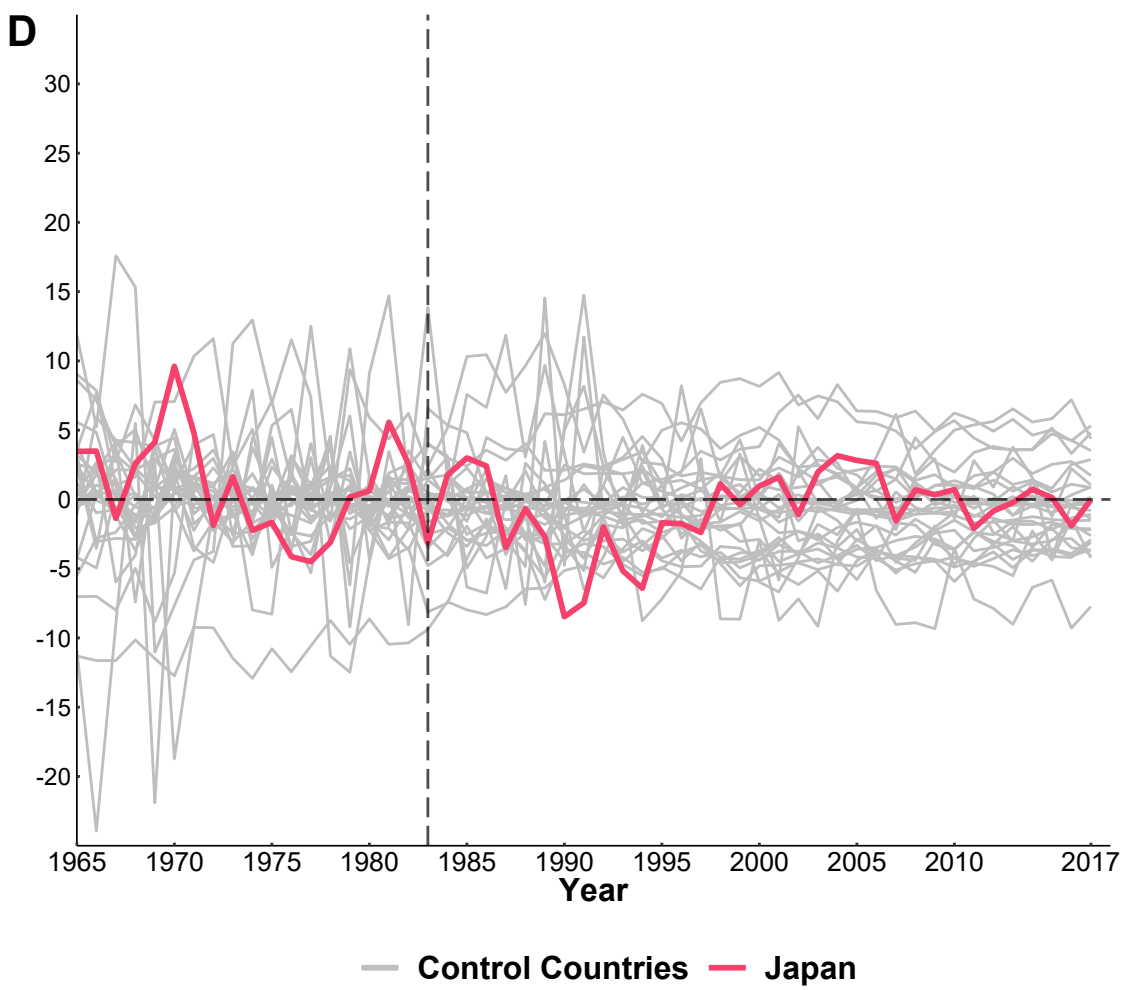
[b]: The effect estimate should be explained cautiously (the average root mean squared percentage errors from the cross-fitting>0.3).

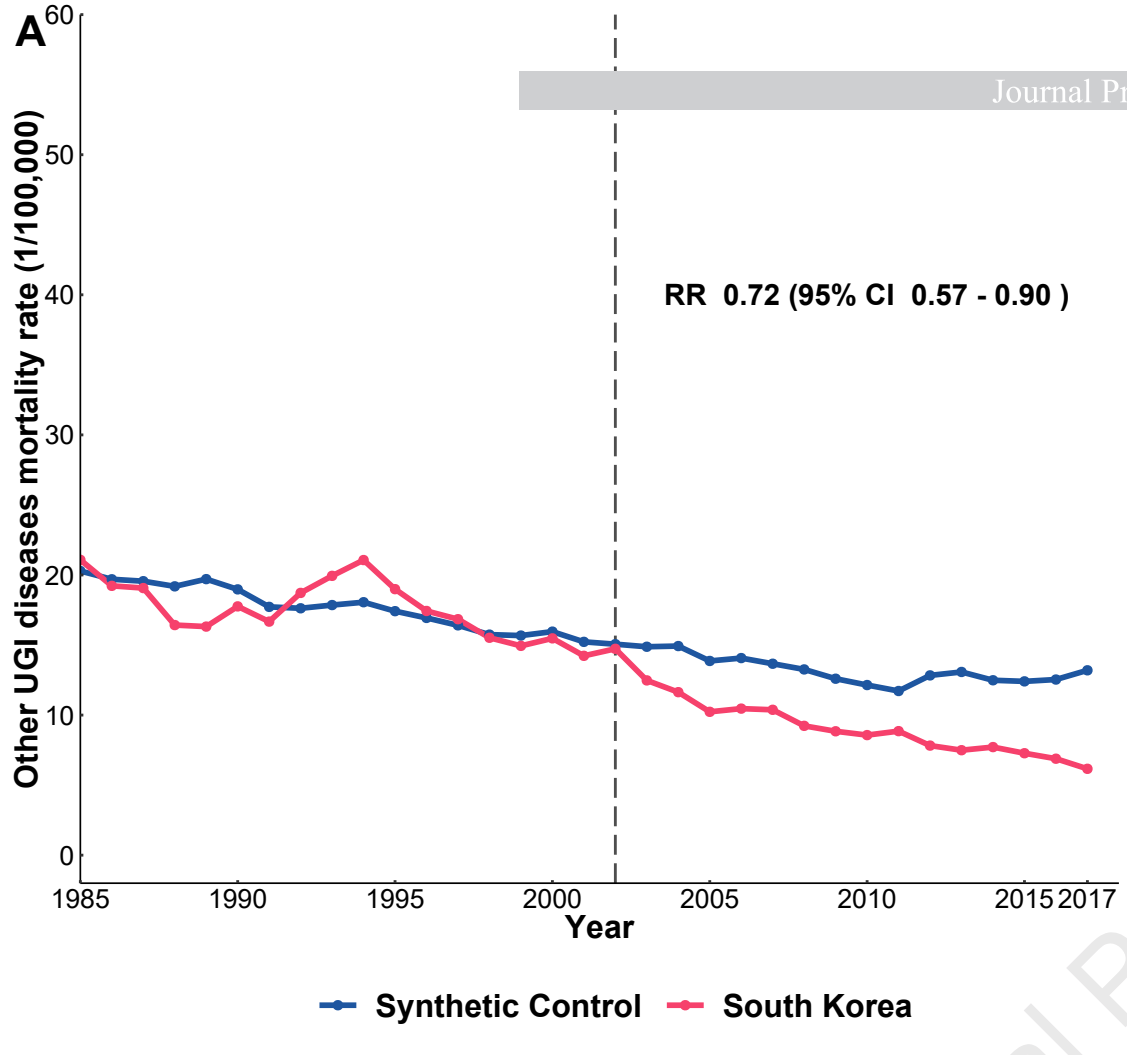
[c]: The effect estimate should be explained cautiously (the average root mean squared percentage errors from the cross-fitting>0.5).

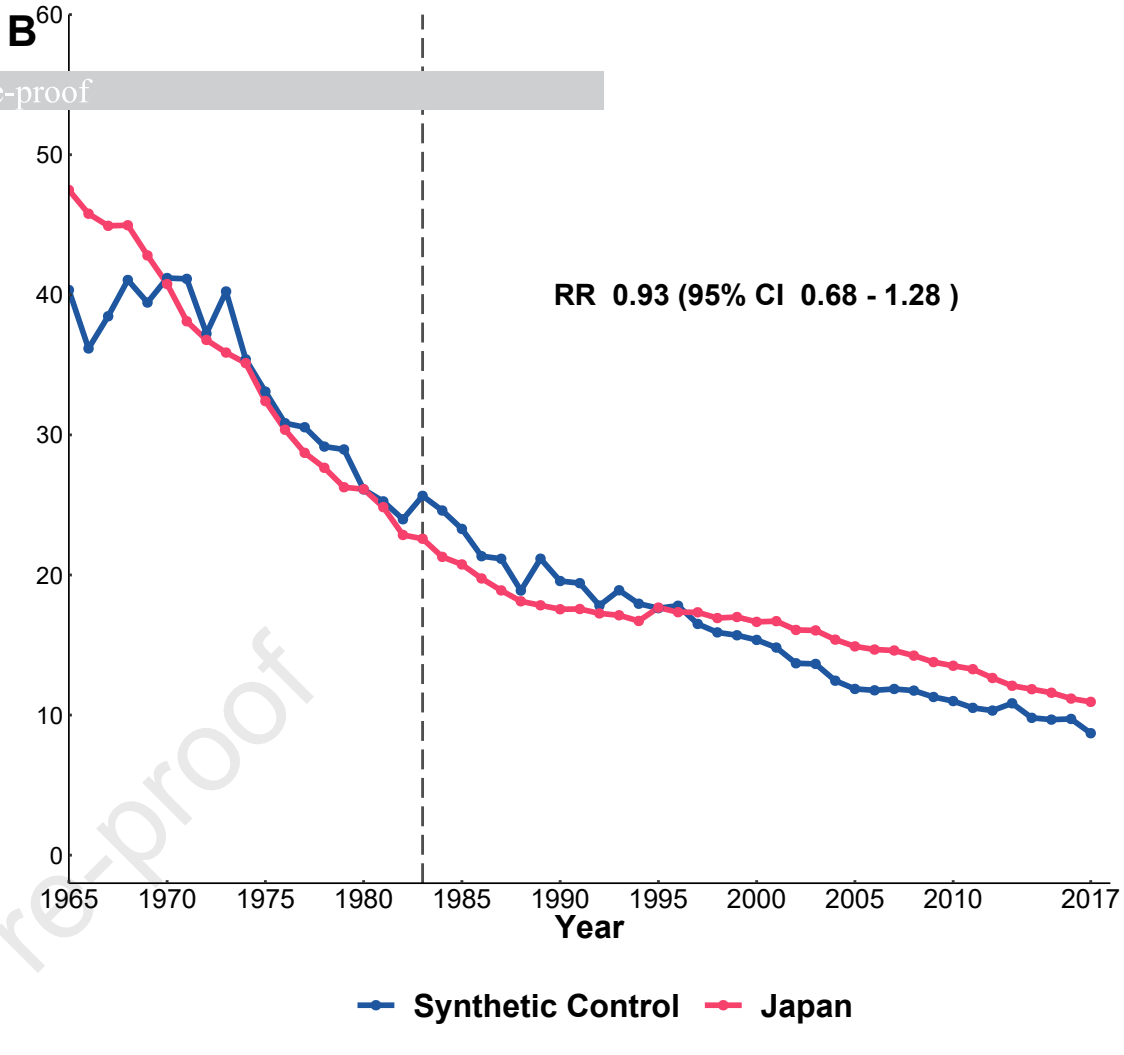


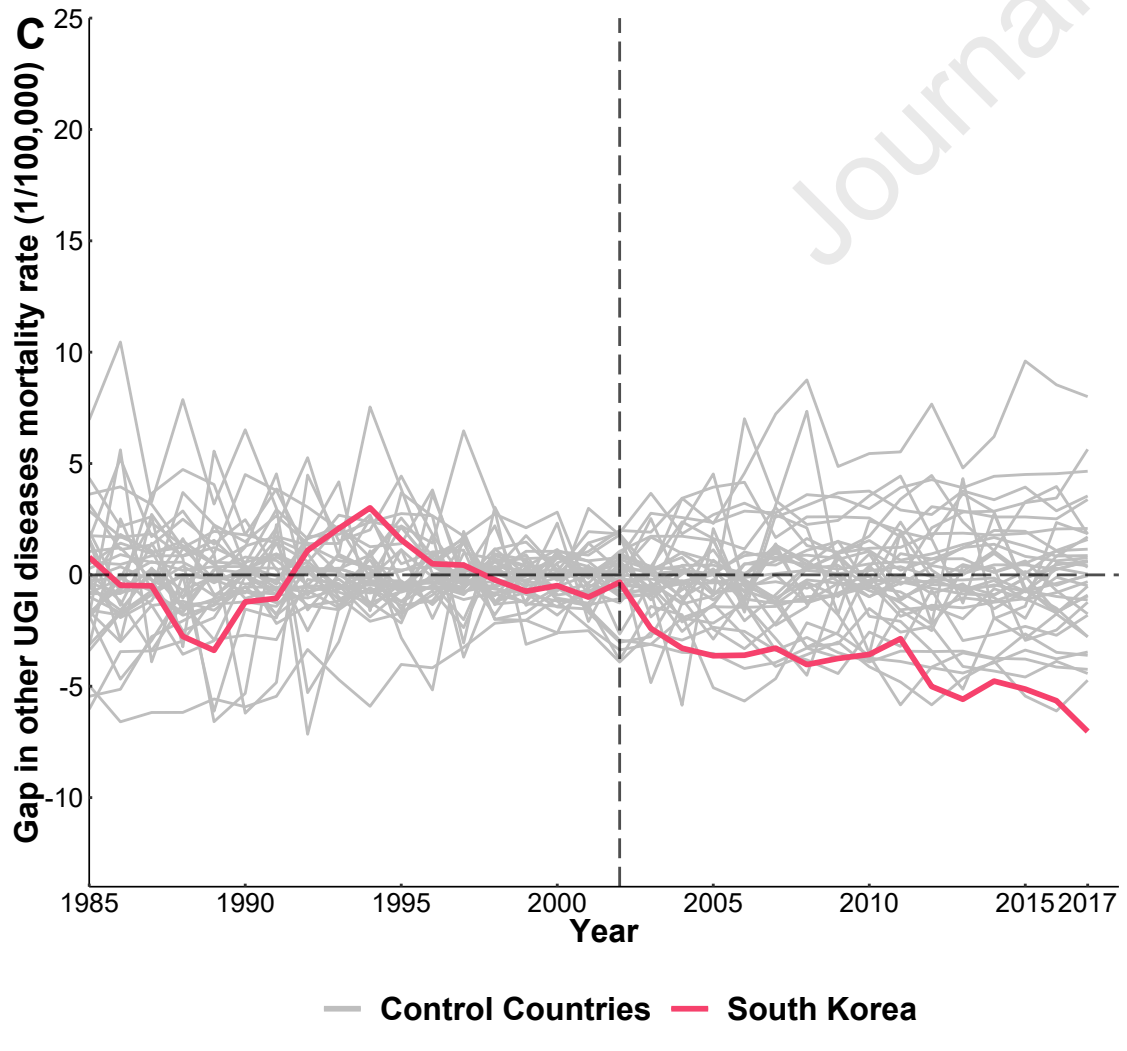


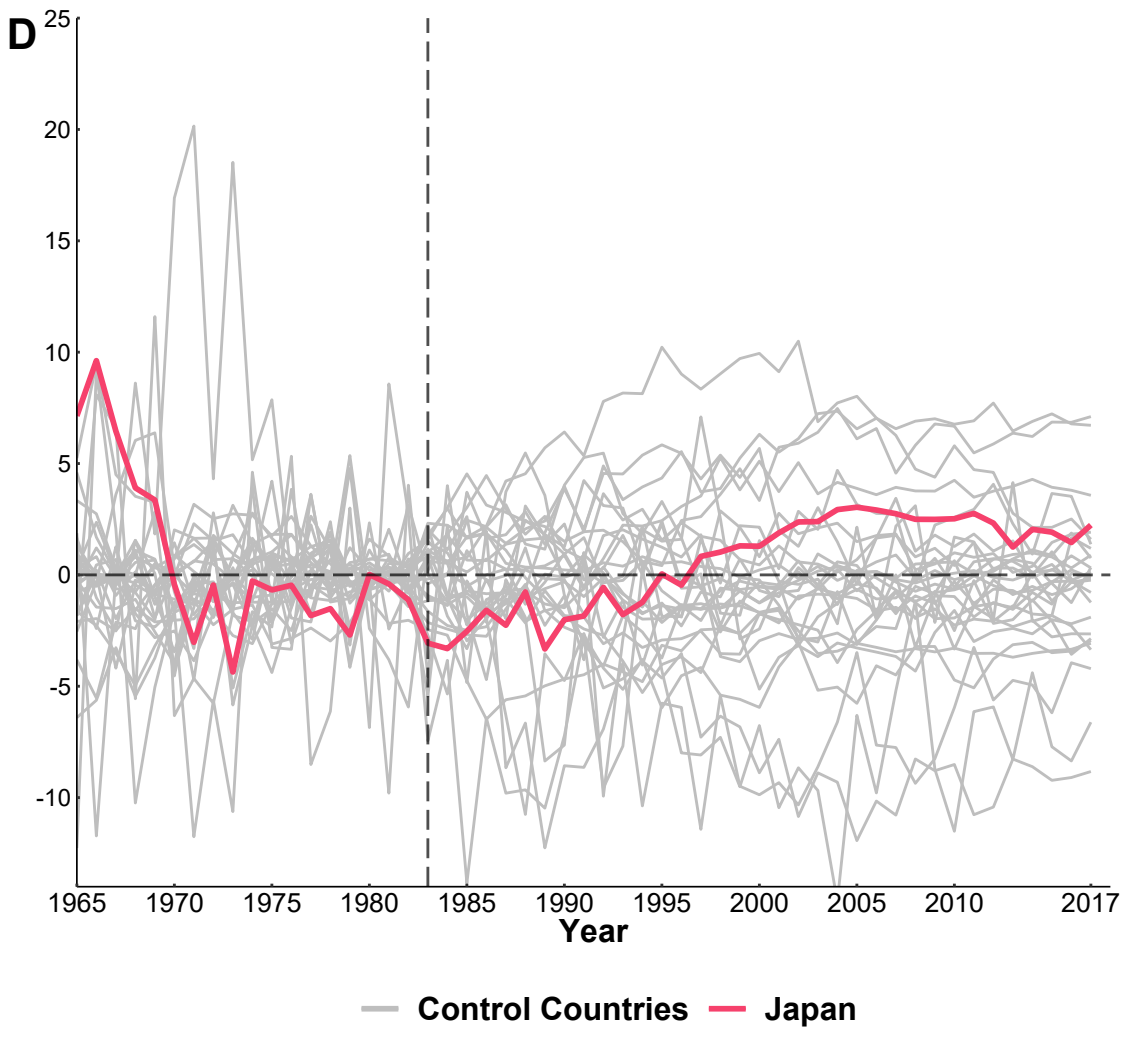












What You Need to Know (25-30 words under each of the first 3 headings. The last headings should be no more than 50-60 words)

Background and Context

Several individual-level and simulation studies have indicated the effectiveness of gastric cancer screening programs. Few studies, however, analyzed the population impact based on observed data.

New Findings

Nationwide screening in South Korea was associated with a reduction in gastric cancer mortality of 41% in the 15th year. The effect on gastric cancer mortality in Japan was uncertain

Limitations

The beneficial impact observed in South Korea may not be generalizable to other countries. The analysis of Japan is limited by the availability of covariate data, which may introduce potential bias in the results.

Clinical Research Relevance and Basic Research Relevance

Our quasi-experimental study complements existing evidence and offers a novel perspective on a population-level basis. It will facilitate triangulation of current available evidence on the efficacy of gastric cancer screening. The insights gained from the data and experiences in South Korea and Japan will serve as valuable references for informing gastric cancer screening programs in other countries.

Lay summary (25-30 words, about one sentence)

We found a significant reduction in gastric cancer mortality rates following the implementation of nationwide screening in South Korea. Further research is needed to validate the impact of screening in Japan.

Supplement to:

The Effect of Nationwide Organized Cancer Screening Programs on Gastric Cancer Mortality: a Synthetic Control Study

Supplementary Methods

Research setting

Table S1. Facts about nationwide gastric cancer screening programs in South Korea and Japan

Attributes	South Korea	Japan
Payment	In 1999, free for Medical Aid recipients In 2002, free for Medical Aid recipients and NHI beneficiaries in the 20% income bracket, whereas the remaining NHI beneficiaries were eligible with a copayment of 50%. In 2003, free of charge to NHI beneficiaries within the 30% income bracket. In 2005, free of charge to NHI beneficiaries in the lower 50% income bracket. In 2006, the copayment amount for the upper 50% was reduced to 20%. In 2010, the copayment was further reduced to 10%[1].	The budgetary responsibility was transferred from the central to local government in 1988. Local municipal governments can set the amount of out-of-pocket payment. About 8.3% of municipalities providing a free gastric cancer screening in 2009. [2]
Organizing systems	Centralized information system under the unified insurer. Screening service providers must adhere to the government recommendations to qualify for financial support.	Each municipality is free to choose whether or not to follow the recommendations of the ministry guidelines [3].
Screening tools	UGI series or endoscopy The proportion of endoscopic screening increased from 31.15% in 2002 to 72.55% in 2011. [1]	The 2005 guideline did not support endoscopic screening. The revised 2014 version approved gastric endoscopy or UGI series. [4] About 19% of municipalities reported having adopted endoscopic screening in 2015. [5]
Participation rates	From 7.40% in 2002 to 45.40% in 2011[1]	From 13% in 2002 to 9.2% in 2011, 8.4% in 2017[6-8]
Access to opportunistic screening	Most people choose population-based screening rather than opportunistic screening[3]	Over 3000 employee- and community-based insurers provide additional screening opportunities; Municipal cancer screening program is not the only option[3]

NHI: National Health Insurance; UGI series: upper gastrointestinal series

The original synthetic control function

The weight W_i^{SCM} in the original synthetic control method (SCM) is derived by optimizing the following function

$$\min_{w}(X_1 - X_0 W^{SCM})' V (X_1 - X_0 W^{SCM}) \quad \text{S1}$$

where X_1 denotes the vector containing covariates and pre-intervention outcomes in the treated unit. X_0 indicates the matrix of corresponding vectors for n_0 control units. V is a symmetric and positive semidefinite matrix where the diagonal elements indicate the relative importance of covariates and pre-intervention outcomes. The weight W_i^{SCM} is restricted by the constraints of being non-negative and having a sum of one. [9]

The flexible synthetic control method

The new flexible method drops the constraint that the weights sum up to one[10] to allow extrapolation and adds a penalty term $\lambda \sum_{i=1}^{n_0}(w_i - w_i^{SCM})^2$ to the original objective function to penalize the sum of squared difference between weights estimated and the original synthetic control weights. W indicates the vector of relaxed weights for n_0 control units from the flexible SCM. W^{scm} is the solution to the original SCM objective function. Thus, the optimization problem is to find the optimal $(n_0 \times 1)$ vector of unit weights $W = (w_1, \ldots, w_{n_0})$ that minimizes the following objective function

$$\min_{w}(X_1 - X_0 W)' V (X_1 - X_0 W) + \lambda \sum_{i=1}^{n_0} \left(w_i - w_i^{SCM}\right)^2 \quad \text{S2}$$

V is determined by the coefficients in the ridge regression model with the average post-intervention outcomes among controls as the dependent variable and scaled X_0 as predictors. The hyper-parameter λ in the penalty term determines the extent of extrapolation. The best λ is selected by a temporal leave-one-out cross-validation approach to avoid over-fitting brought about by extensive extrapolation [11]. The estimator for the average rate ratio during the post-intervention period is calculated by dividing the average outcome values of the treated unit by the average of the synthetic control across the post-intervention periods. A three-fold cross-fitting method mentioned in the main text is used to obtain the final debiased estimate. Specifically, weights estimated from the training sample are used to predict the outcome in the holdout period. Then the bias is measured by differences between the observed and predicted values in the holdout period. To make the cross-fitting practical, each covariate needs to have values in each training period. We included covariates with values across thirds of pre-intervention period.

Table S2. Sources, definitions, and available years of covariates included in the analysis

Variables	South Korea		Japan		Sources
	Available years for SC	Available years for SDID	Available years for SC	Available years for SDID[a]	
Economic indicators					
Gross national income per capita (2017 PPP$)	1990-2001	1990-2017[a]	NA	NA	UNDP: Human Development Report [12]
Demographic attributes					
Rural population (% of the total population)	1985-2001	1985-2017	1965-2017	1965-2017	World Bank: HNPS [13]
Average total years of schooling for the adult population	1985,1990-2001	1985,1990-2017[a]	1965,1970, 1975,1980	1965,1970,1975,1980,1985,1990-2017[a]	Our World in Data [14]
Health sanitation services					
Unsafe water sanitation and handwashing	1990-2001	1990-2017[a]	NA	NA	GBD Study 2019 [15]
Potential gastric cancer risk factors					
Diet high in sodium (summary exposure value, %)	1990-2001	1990-2017[a]	NA	NA	GBD Study 2019 [15]
Tobacco (summary exposure value, %)	1990-2001	1990-2017[a]	NA	NA	GBD Study 2019 [15]
Alcohol use (summary exposure value, %)	1990-2001	1990-2017[a]	NA	NA	GBD Study 2019 [15]
Prevalence of overweight (% of adults)	1985-2001	1985-2016[a]	NA	NA	World Bank: HNPS [13]

PPP: purchasing power parity; NA: not available; UNDP: United Nations Development Program, HNPS: Health Nutrition and Population Statistics; GBD: Global Burden of Disease; SC: synthetic control; SDID: synthetic difference in differences
a: missing data during the study period (South Korea 1985-2017, Japan 1965-2017) were imputed with spline interpolation.

Synthetic control group selection

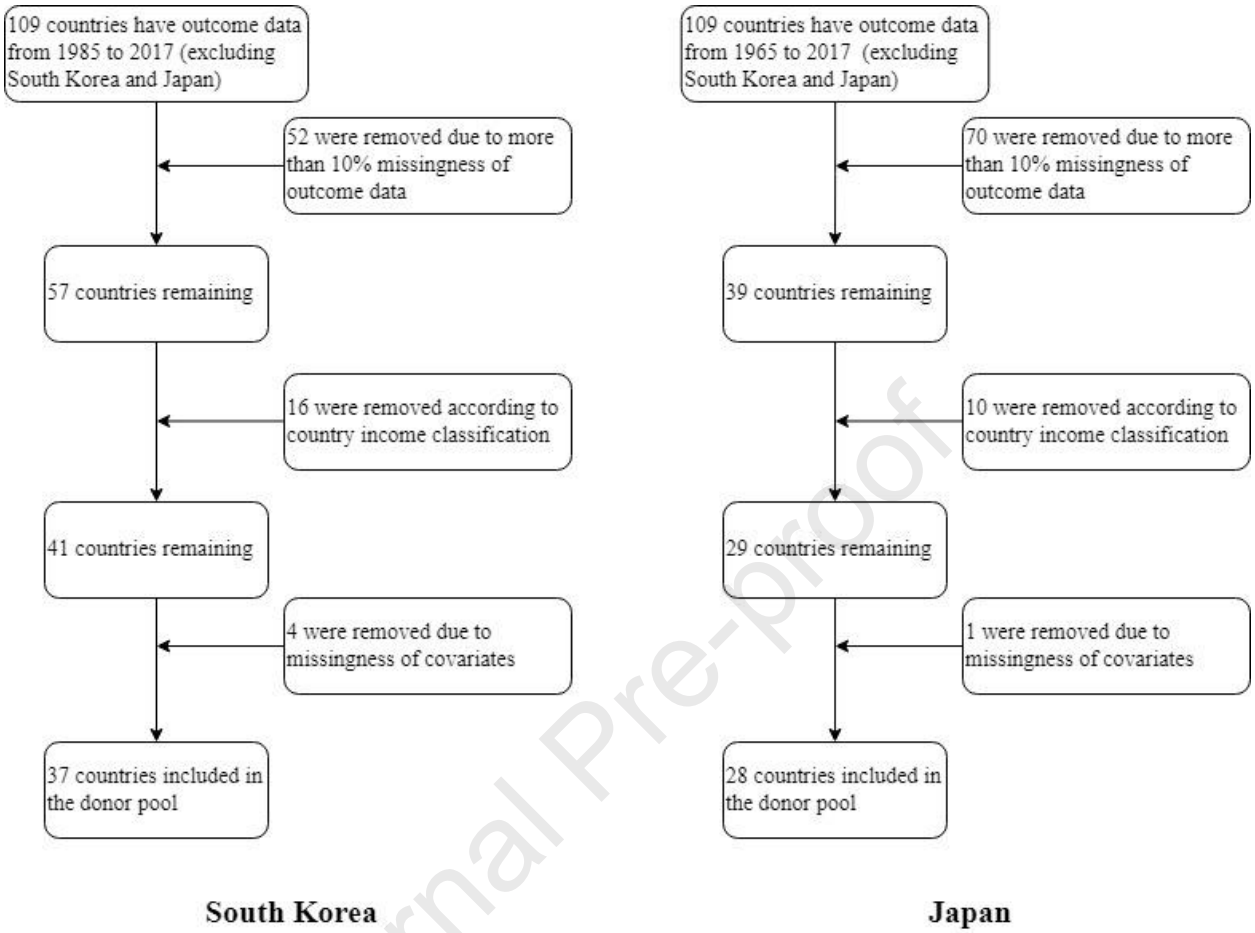


Figure S1. The selection procedure for the donor pools of South Korea and Japan

This flowchart can be applied to the analyses for all outcomes in the study (i.e., gastric cancer mortality, other upper gastrointestinal disease mortality, chronic obstructive pulmonary disease mortality, mortality for people aged between 35 and 39).

Synthetic difference-in-differences

The methodology paper by Arkhangelsky et al. described the synthetic difference-in-differences (SDID) in detail [16]. Briefly, the SDID method is like the synthetic control method (SCM) assigning weight (ω^{sdid}) to control unit so that the weighted average outcome for the control group is parallel to the outcome for the treated unit. In contrast to the SCM, SDID adds an intercept term (ω_0). This change means that the synthetic control's pre-intervention outcome trend does not have to perfectly match the treated unit as with the SCM. Instead, the parallel trend during pre-intervention period is sufficient. In addition, the SDID method also considers the time weights (λ^{sdid}) so that the weighted average of pre-intervention outcomes for each control unit is different from the average post-intervention outcome for the same control unit by a constant. The unit weights (ω^{sdid}) define a synthetic control unit using pre-intervention data, and the time weights (λ^{sdid}) define a synthetic pre-intervention period using control data. In this way, the difference-in-difference becomes more plausible with the unit and time weights. Finally, the SDID estimator is calculated using a weighted difference-in-difference regression. To adjust time-varying covariates, the SDID is applied to the residuals $Y_{it}^{resid} = Y_{it} - X_{it}\beta$ for the regression of outcome Y_{it} on covariates X_{it}. The adjustment method proposed by Arkhangelsky et al. is however sensitive to the scale of covariates. We thus used the modified adjustment method developed by Kranz [17]. We imputed missing covariate data using spline interpolation since the SDID method requires a balanced panel for covariates and outcome for both the pre- and post-intervention period. The R packages “synthdid” (version 0.0.9) and “xsynthdid” (version 0.1.0) were used for the SDID estimation.

Supplementary Results

Supplementary Tables

Table S3. Average covariate values for South Korea, Japan, synthetic controls, and donor pools before the intervention year

Covariates	South Korea				Japan			
	Treated country (South Korea)	Synthetic control for gastric cancer mortality	Synthetic control for other UGI disease mortality	Donor pool (average)	Treated country (Japan)	Synthetic control for gastric cancer mortality	Synthetic control for other UGI disease mortality	Donor pool (average)
Economic indicators								
Gross national income per capita (2017 PPP$)	18093	17994	21976	29290	—	—	—	—
Demographic attributes								
Rural population (% of total population)	24.98	28.53	27.21	24.26	26.48	28.03	28.20	26.76
Average total years of schooling for the adult population	9.82	10.15	9.19	9.15	7.66	7.86	7.56	6.92
Health sanitation services								
Unsafe water sanitation and handwashing (summary exposure value, %)	24.81	43.77	24.22	19.55	—	—	—	—
Gastric cancer risk factors								
Diet high in sodium (summary exposure value, %)	60.92	44.66	54.94	31.56	—	—	—	—
Tobacco (summary exposure value, %)	40.52	36.30	35.53	33.31	—	—	—	—
Alcohol use (summary exposure value, %)	13.23	13.44	11.88	12.70	—	—	—	—
Prevalence of overweight (% of adults)	21.26	59.59	44.96	48.13	—	—	—	—

UGI disease: upper gastrointestinal disease; PPP: purchasing power parity

Table S4. Estimates of annual effects for nationwide gastric cancer screening in South Korea

Year	Gastric cancer mortality		Other UGI disease mortality	
	Rate difference (per 100,000)	Rate ratio	Rate difference (per 100,000)	Rate ratio
2002	2.37	1.03	-0.33	0.98
2003	0.60	1.01	-2.40	0.84
2004	-2.71	0.96	-3.29	0.78
2005	-2.85	0.95	-3.63	0.74
2006	-3.40	0.94	-3.61	0.74
2007	-1.94	0.96	-3.29	0.76
2008	-5.45	0.90	-4.03	0.70
2009	-10.67	0.81	-3.75	0.70
2010	-11.76	0.79	-3.57	0.71
2011	-11.47	0.78	-2.86	0.76
2012	-15.57	0.71	-5.01	0.61
2013	-13.95	0.71	-5.58	0.57
2014	-14.83	0.69	-4.77	0.62
2015	-17.72	0.63	-5.13	0.59
2016	-16.79	0.62	-5.65	0.55
2017	-17.85	0.59	-7.03	0.47

UGI disease: upper gastrointestinal disease

Table S5. Estimates of annual effects for nationwide gastric cancer screening in Japan

Year	Gastric cancer mortality		Other UGI disease mortality	
	Rate difference (per 100,000)	Rate ratio	Rate difference (per 100,000)	Rate ratio
1983	-3.12	0.97	-3.06	0.88
1984	1.75	1.02	-3.31	0.87
1985	2.99	1.03	-2.53	0.89
1986	2.42	1.03	-1.59	0.93
1987	-3.48	0.96	-2.26	0.89
1988	-0.67	0.99	-0.77	0.96
1989	-2.69	0.97	-3.33	0.84
1990	-8.47	0.90	-2.01	0.90
1991	-7.47	0.91	-1.85	0.90
1992	-2.00	0.97	-0.56	0.97
1993	-5.16	0.93	-1.78	0.91
1994	-6.41	0.91	-1.23	0.93
1995	-1.69	0.98	0.04	1.00
1996	-1.75	0.97	-0.46	0.97
1997	-2.39	0.96	0.83	1.05
1998	1.12	1.02	1.02	1.06
1999	-0.38	0.99	1.30	1.08
2000	0.97	1.02	1.29	1.08
2001	1.61	1.03	1.88	1.13
2002	-1.10	0.98	2.38	1.17
2003	1.98	1.04	2.39	1.18
2004	3.15	1.07	2.93	1.24
2005	2.82	1.06	3.04	1.26
2006	2.60	1.06	2.91	1.25
2007	-1.55	0.97	2.75	1.23
2008	0.69	1.02	2.50	1.21
2009	0.34	1.01	2.49	1.22
2010	0.71	1.02	2.52	1.23
2011	-2.09	0.95	2.76	1.26
2012	-0.84	0.98	2.33	1.23
2013	-0.23	0.99	1.25	1.12
2014	0.73	1.02	2.05	1.21
2015	0.14	1.00	1.92	1.20
2016	-1.90	0.94	1.45	1.15
2017	-0.02	1.00	2.24	1.26

UGI disease: upper gastrointestinal disease

Supplementary Figures

Measure	Types of analysis	Gastric cancer mortality				Other UGI disease mortality			
		South Korea		Japan		South Korea		Japan	
RMSPE of outcomes during the pre-intervention period	Base-case analysis	0.0610		0.0289		0.0889		0.0893	
	Sensitivity analysis								
	Negative control age band (35-39) analysis	0.2217		0.1345		0.4697		0.1521	
	Negative control outcome analysis (COPD)	0.3371		0.2455		0.3371		0.2455	
	Leave-out-one examination	Exclude CHL	0.0994	Exclude ITA	0.0289	Exclude CZE	0.0909	Exclude CHE	0.0909
		Exclude CRI	0.0842	Exclude MLT	0.0308	Exclude GRC	0.0914	Exclude ISL	0.0900
		Exclude ITA	0.0619	Exclude SGP	0.0313	Exclude MUS	0.0876	Exclude PRT	0.0894
		Exclude SGP	0.0626	Exclude URY	0.0404	Exclude POL	0.0827	Exclude SGP	0.0996
				Exclude VEN	0.0402	Exclude SGP	0.0769	Exclude USA	0.0925
average RMSPE of outcomes during holdout periods in the cross-fitting	Base-case analysis	0.0739		0.0468		0.0901		0.1385	
	Sensitivity analysis								
	Negative control age band (35-39) analysis	0.3364		0.1941		0.5918		0.4358	
	Negative control outcome analysis (COPD)	0.5737		0.3319		0.5737		0.3319	
	Leave-out-one examination	Exclude CHL	0.1268	Exclude ITA	0.0467	Exclude CZE	0.0939	Exclude CHE	0.1522
		Exclude CRI	0.0948	Exclude MLT	0.0392	Exclude GRC	0.1019	Exclude ISL	0.1445
		Exclude ITA	0.0774	Exclude SGP	0.0532	Exclude MUS	0.1036	Exclude PRT	0.1444
		Exclude SGP	0.0786	Exclude URY	0.0491	Exclude POL	0.0772	Exclude SGP	0.1678
				Exclude VEN	0.0459	Exclude SGP	0.0840	Exclude USA	0.1264

Figure S2. Root mean squared percentage errors (RMSPE) between predicted and observed outcomes during the pre-intervention period, as well as the average RMSPE of outcomes during holdout periods in the cross-fitting procedure

The spectrum of colors, ranging from light to dark red, illustrates a progression from smaller to larger RMSPEs. Separate color labeling systems were applied to the RMSPE and the average RMSPE from the cross-fitting. RMSPE indicates the root mean squared percentage errors. UGI indicates upper gastrointestinal. COPD indicates chronic obstructive pulmonary disease. CHL: Chile; CRI: Costa Rica; CZE: Czechia; GRC: Greece; ISL: Iceland; ITA: Italy; MLT: Malta; MUS: Mauritius; POL: Poland; PRT: Portugal; SGP: Singapore; CHE: Switzerland; URY: Uruguay; USA: United States; VEN: Venezuela.

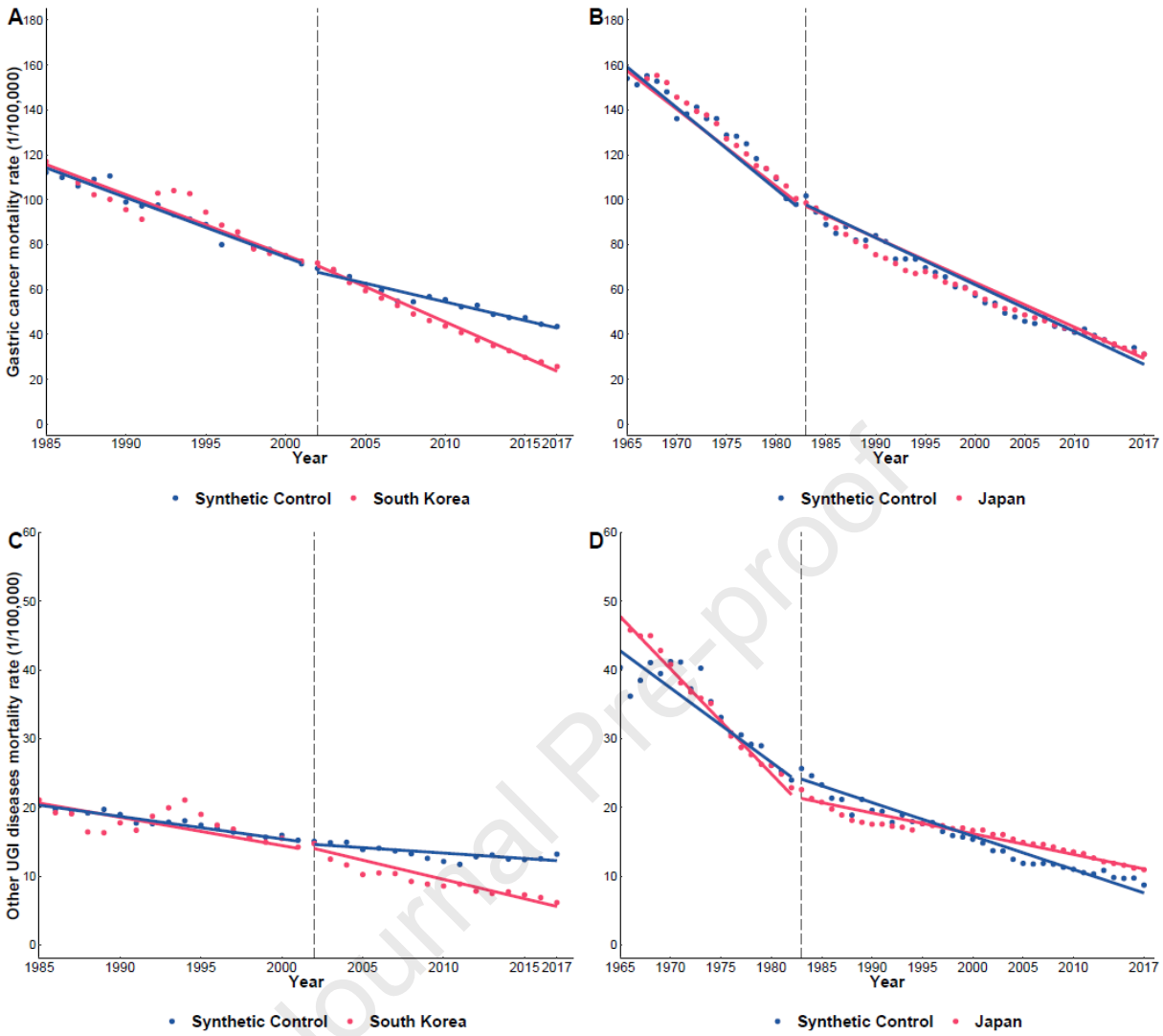


Figure S3. Trends before/after nationwide screening for gastric cancer mortality (A, B) and other upper gastrointestinal disease mortality (C, D) in South Korea (A, C) and Japan (B, D)

The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening.

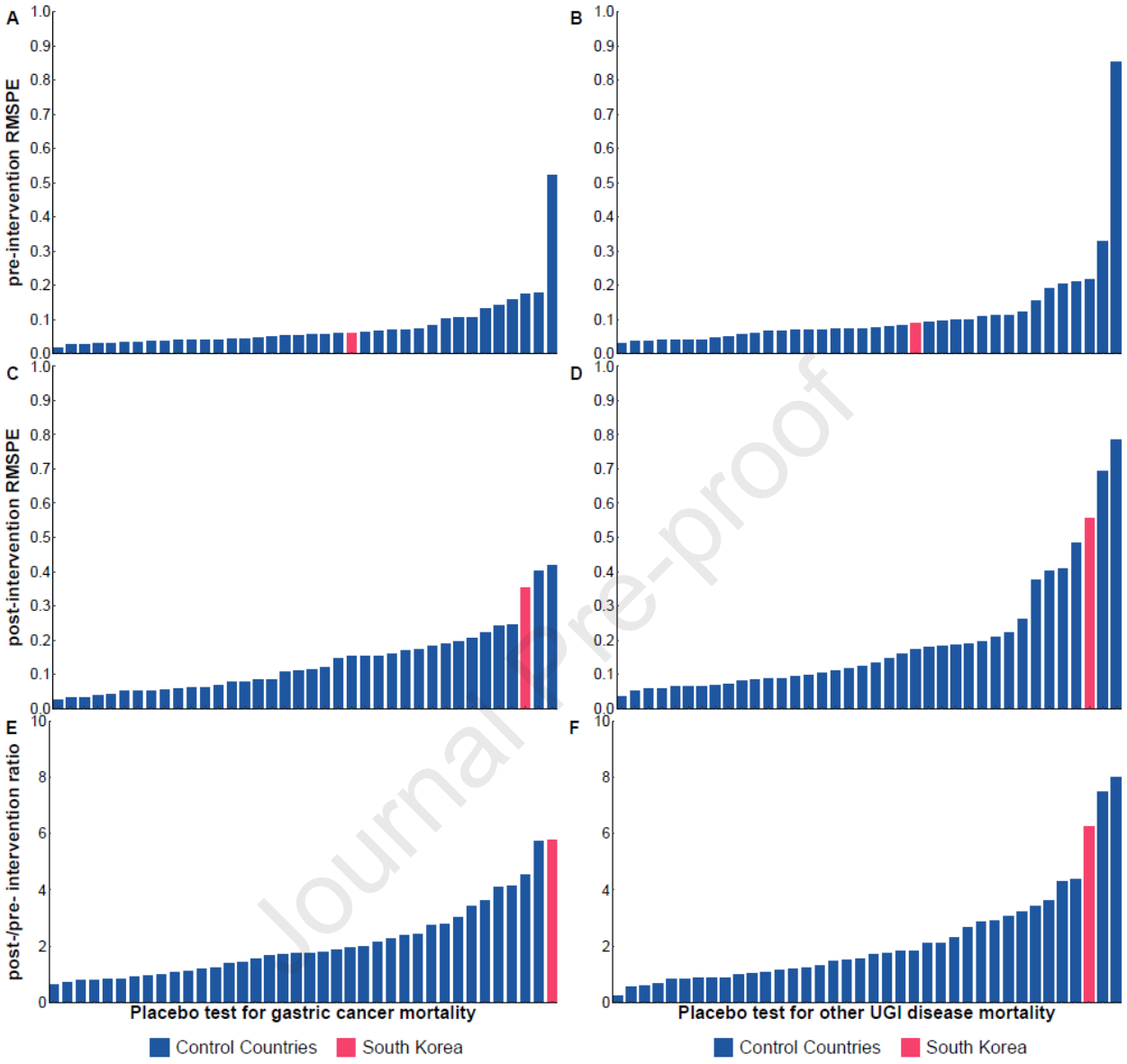


Figure S4. Pre- (A, B) and post- (C, D) intervention root mean squared percentage errors and their ratios (E, F) in placebo tests for gastric cancer mortality (A, C, E) and other upper gastrointestinal disease mortality (B, D, F) in South Korea

RMSPE indicates the root mean squared percentage errors. UGI indicates upper gastrointestinal.

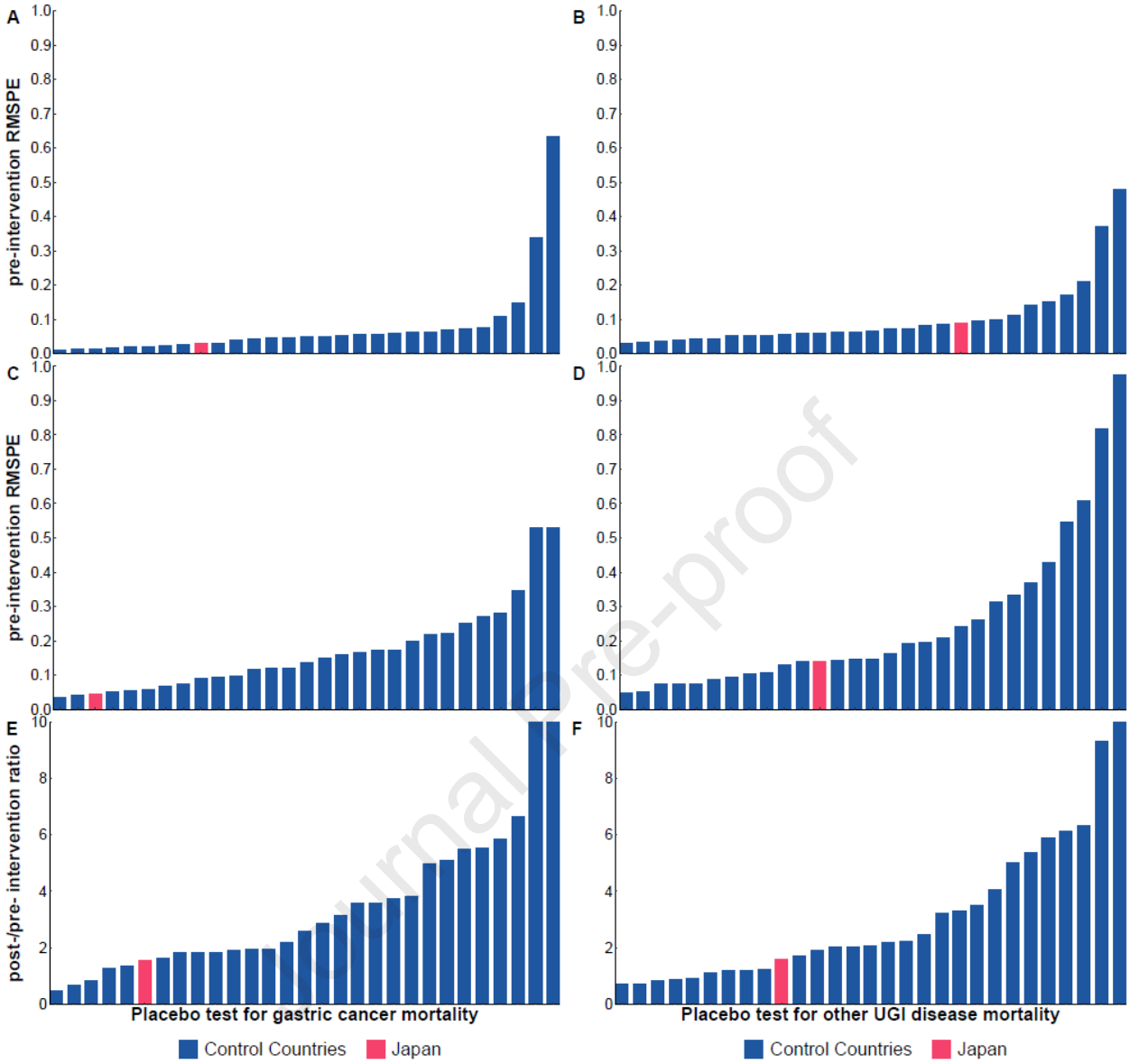


Figure S5. Pre- (A, B) and post- (C, D) intervention root mean squared percentage errors and their ratios (E, F) in placebo tests for gastric cancer mortality (A, C, E) and other upper gastrointestinal disease mortality (B, D, F) in Japan

RMSPE indicates the root mean squared percentage errors. UGI indicates upper gastrointestinal.

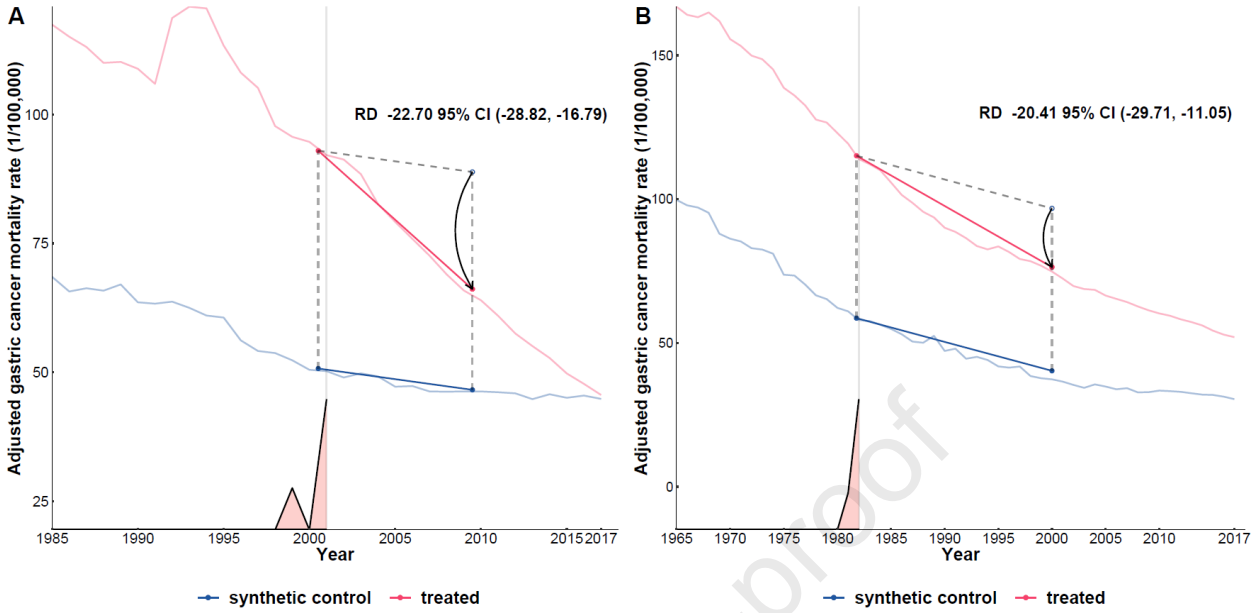


Figure S6. Sensitivity test: synthetic difference-in-difference analyses for effect on gastric cancer mortality in South Korea (A) and Japan (B).

The straight blue line represents the change in outcome from a pre-intervention period through the post-intervention period for the control group, while the solid red line represents the change for the treated group and the parallel dashed line shows the counterfactual change. The estimated effect is indicated by the arrow. The pink curve above x-axis indicates the size of time weight for different years. RD indicates rate difference. CI indicates confidence interval. Please note that the y-axis values indicate the adjusted outcome value after subtracting the effects of covariates instead of the original value of gastric cancer mortality.

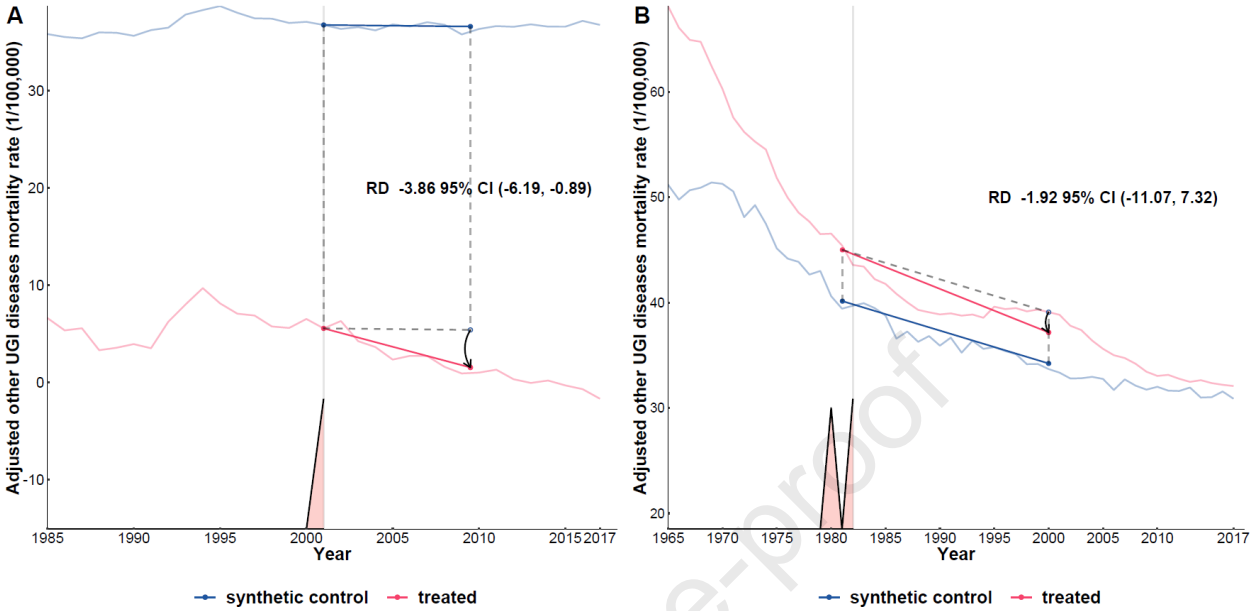


Figure S7. Sensitivity test: synthetic difference-in-difference analyses for effect on other upper gastrointestinal disease mortality in South Korea (A) and Japan (B)

The straight blue line represents the change in outcome from a pre-intervention period through the post-intervention period for the control group, while the solid red line represents the change for the treated group and the parallel dashed line shows the counterfactual change. The estimated effect is indicated by the arrow. The pink curve above x-axis indicates the size of time weight for different years. RD indicates rate difference. CI indicates confidence interval. Please note that the y-axis values indicate the adjusted outcome value after subtracting the effects of covariates instead of the original value of other upper gastrointestinal disease mortality.

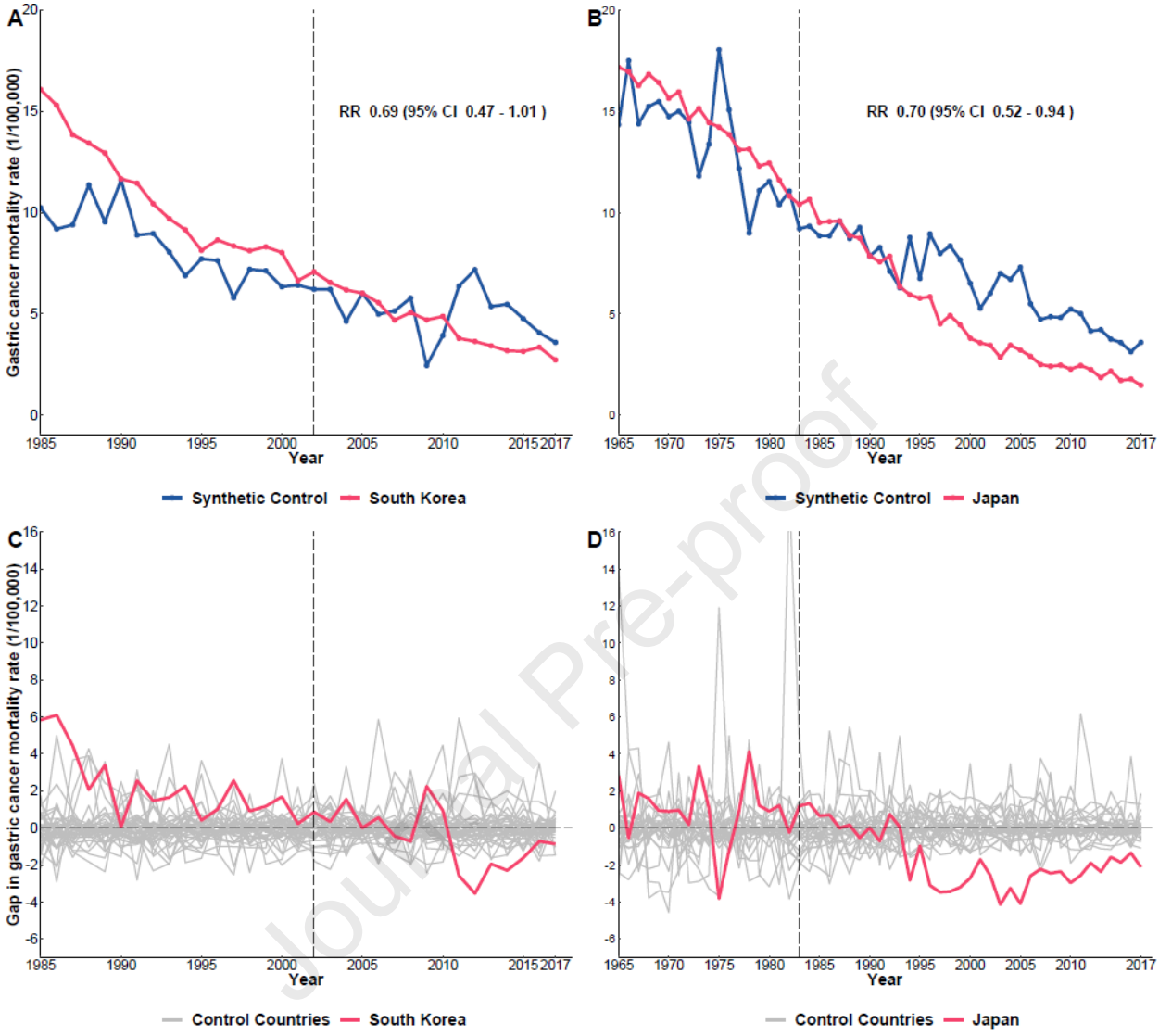


Figure S8. Sensitivity test: trends in gastric cancer mortality in South Korea (A) and Japan (B) versus their synthetic control countries as well as the intervention effects and placebo tests (C, D) among the population aged between 35 and 39

The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening. In panels C and D, the red lines denote the estimated intervention effects in South Korea and Japan respectively. The grey lines correspond to the effects estimated in placebo tests, where each control country was treated as the unit of interest. RR indicates rate ratio. CI indicates confidence interval.

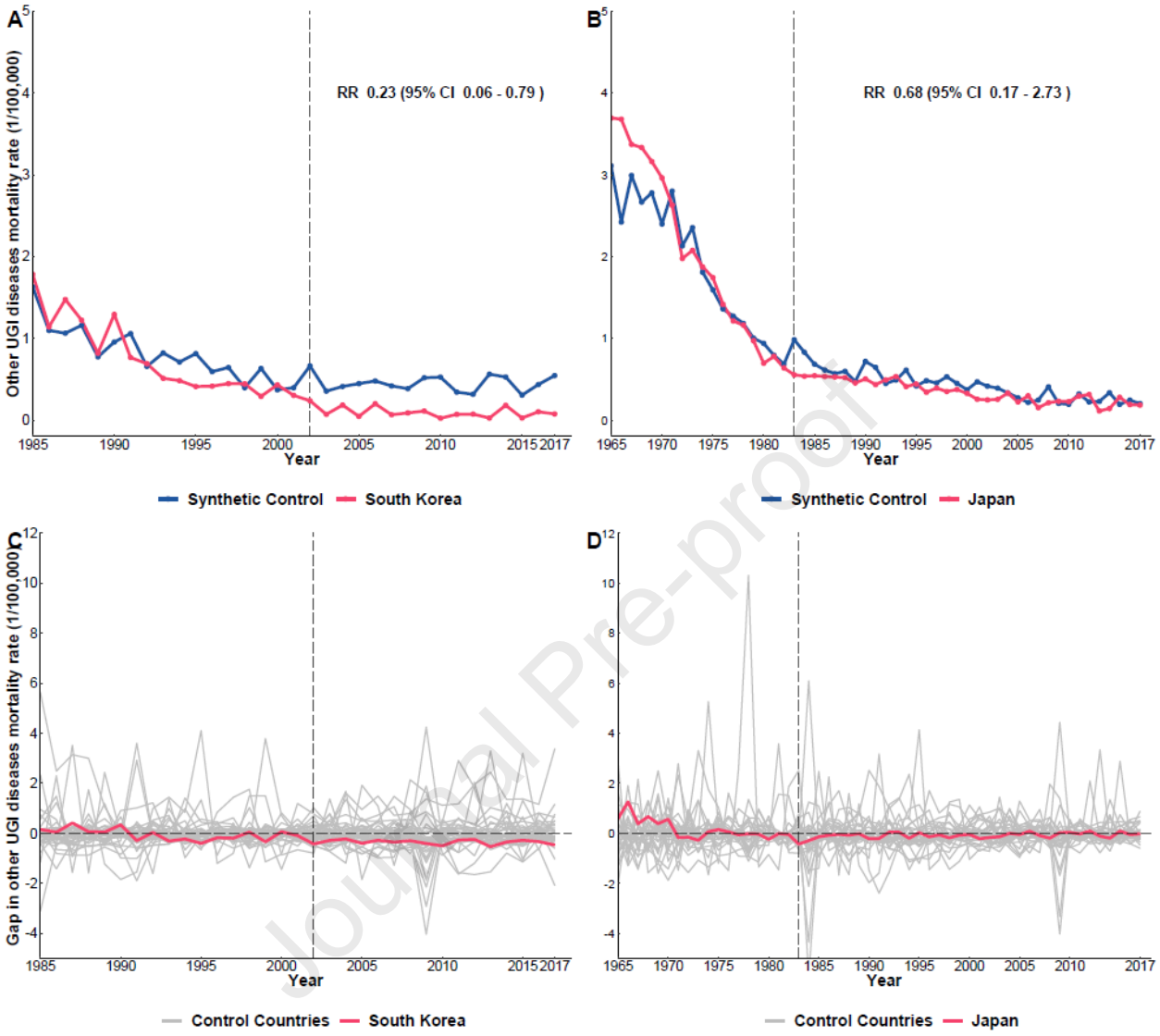


Figure S9. Sensitivity test: trends in other upper gastrointestinal disease mortality in South Korea (A) and Japan (B) versus their synthetic control countries as well as the intervention effects and placebo tests (C, D) among the population aged between 35 and 39

The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening. In panels C and D, the red lines denote the estimated intervention effects in South Korea and Japan, respectively. Meanwhile, the grey lines correspond to the effects estimated in placebo tests, where each control country was treated as the unit of interest. UGI disease indicates upper gastrointestinal disease. RR indicates rate ratio. CI indicates confidence interval.

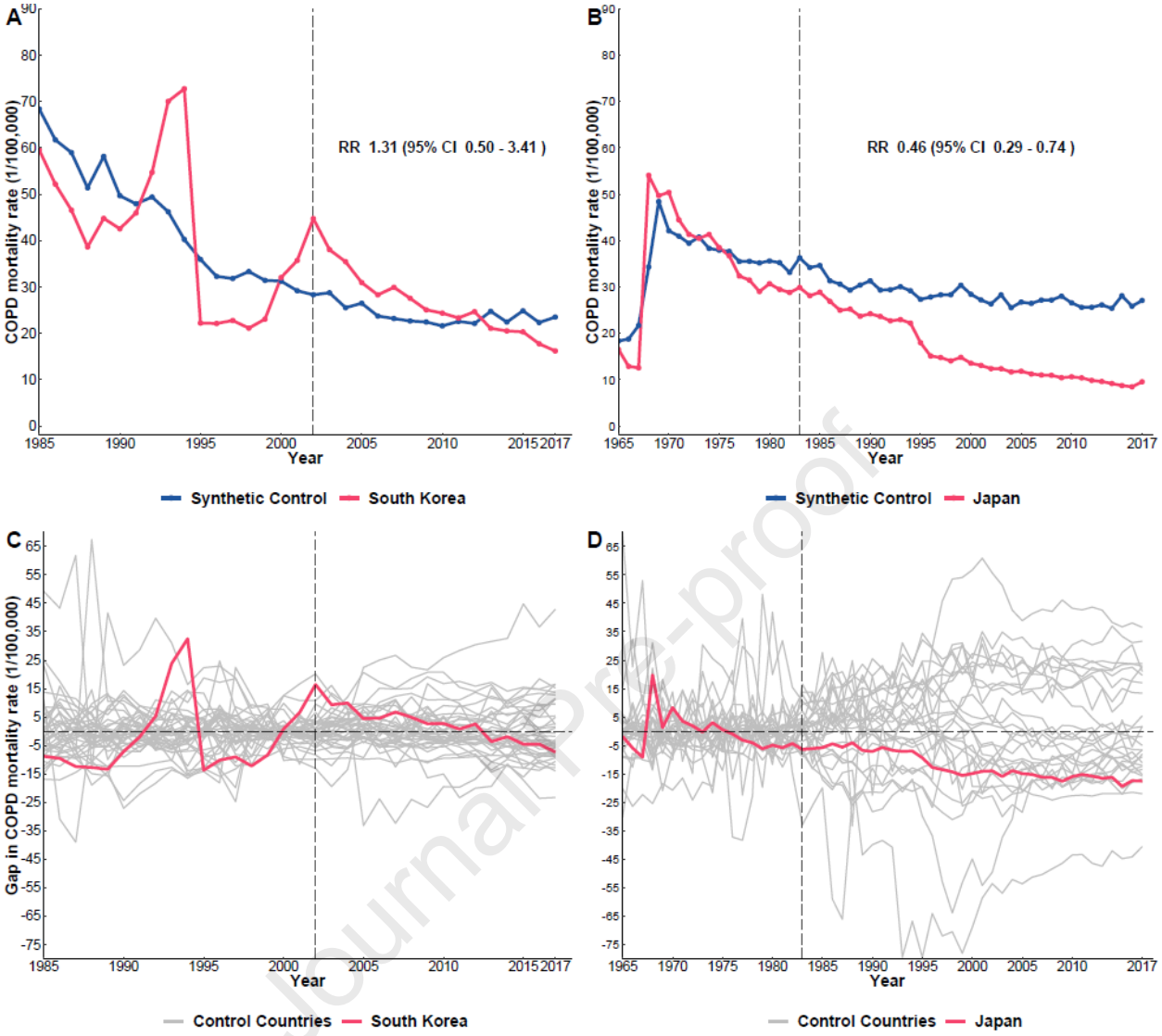


Figure S10. Sensitivity test: trends in chronic obstructive pulmonary disease mortality in South Korea (A) and Japan (B) versus their synthetic control countries as well as the intervention effects and placebo tests (C, D).

The dashed vertical line indicates the introduction of nationwide organized gastric cancer screening. In panels C and D, the red lines denote the estimated intervention effects in South Korea and Japan, respectively. Meanwhile, the grey lines correspond to the effects estimated in placebo tests, where each control country was treated as the unit of interest. COPD indicates chronic obstructive pulmonary disease. RR indicates rate ratio. CI indicates confidence interval.

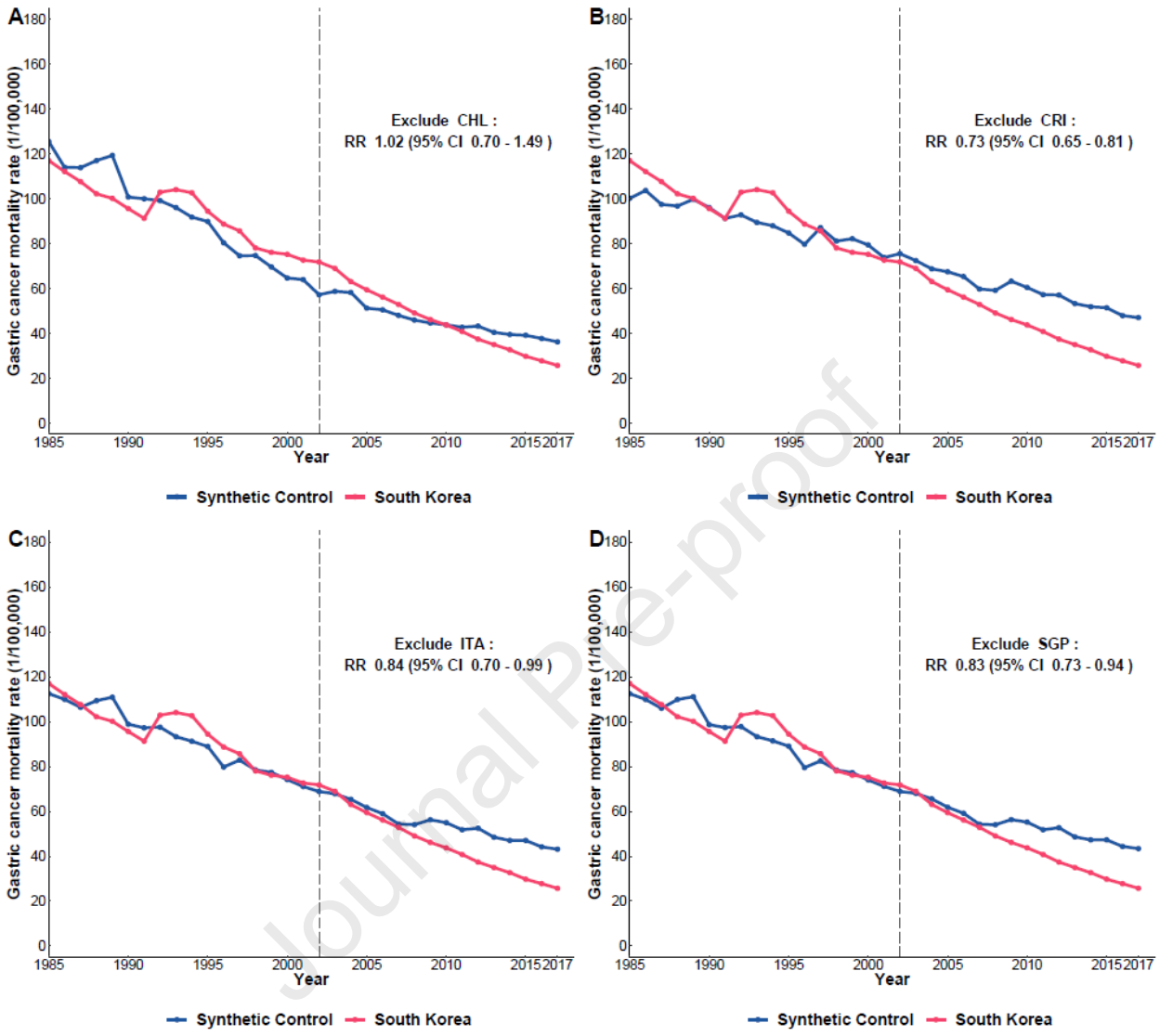


Figure S11. Sensitivity test: leave-out-one examination for the effect on gastric cancer mortality in South Korea

CHL: Chile; CRI: Costa Rica; ITA: Italy; SGP: Singapore; RR: rate ratio; CI: confidence interval.

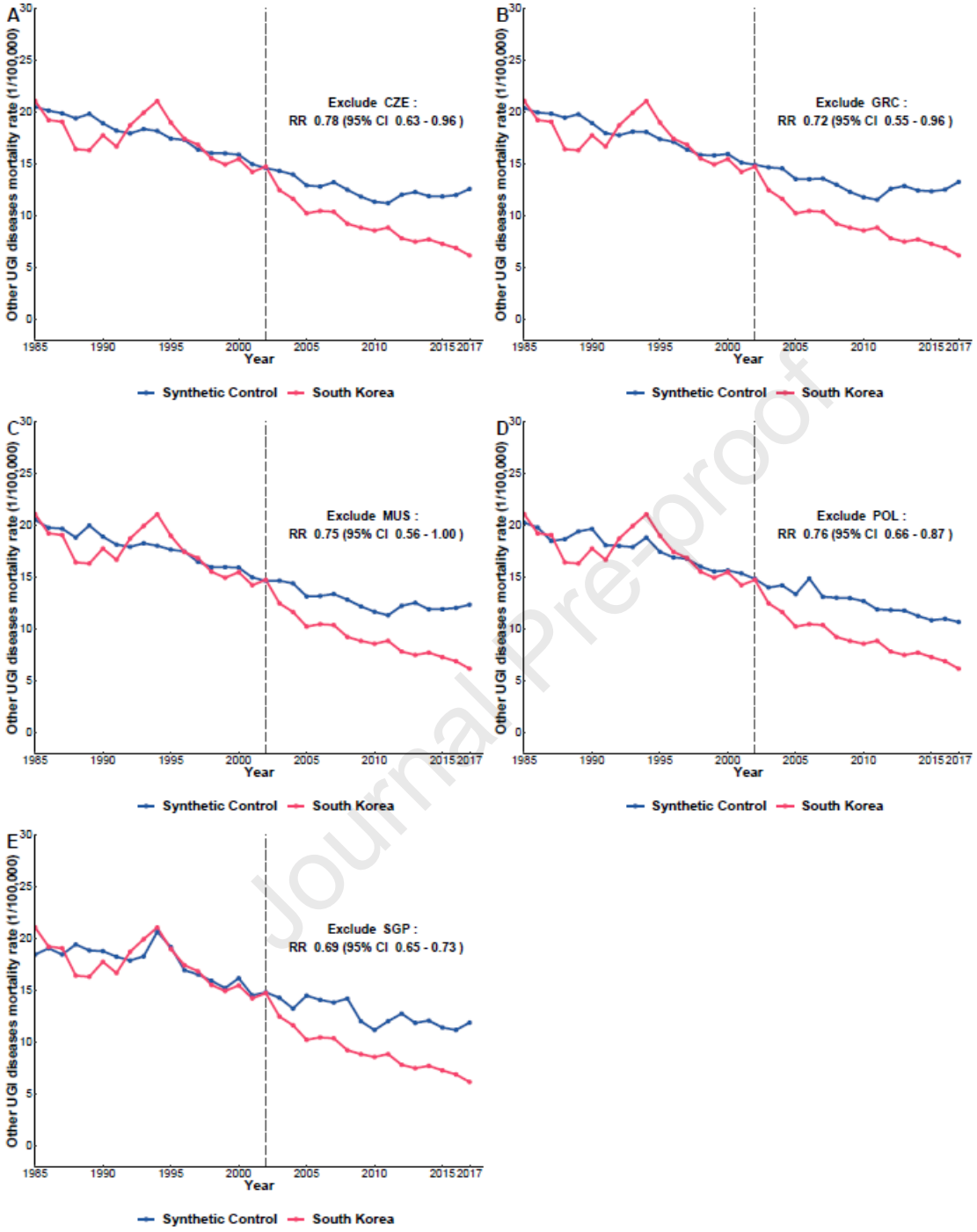


Figure S12. Sensitivity test: leave-out-one examination for the effect on other upper gastrointestinal disease mortality in South Korea

UGI disease: upper gastrointestinal disease; CZE: Czechia; GRC: Greece; MUS: Mauritius; POL: Poland; SGP: Singapore; RR: rate ratio; CI: confidence interval

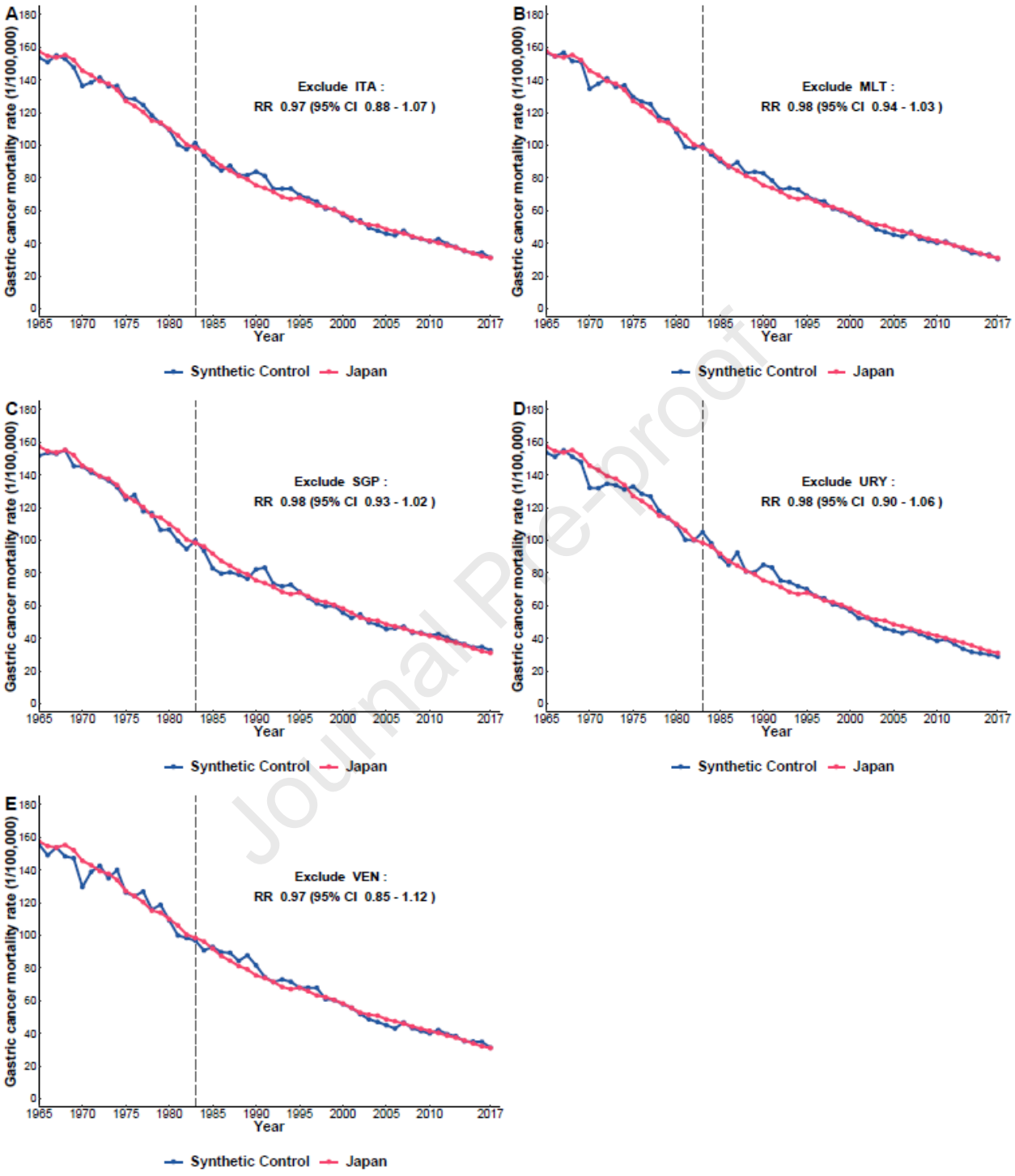


Figure S13. Sensitivity test: leave-out-one examination for the effect on gastric cancer mortality in Japan

ITA: Italy; MLT: Malta; SGP: Singapore; URY: Uruguay; VEN: Venezuela; RR: rate ratio; CI: confidence interval

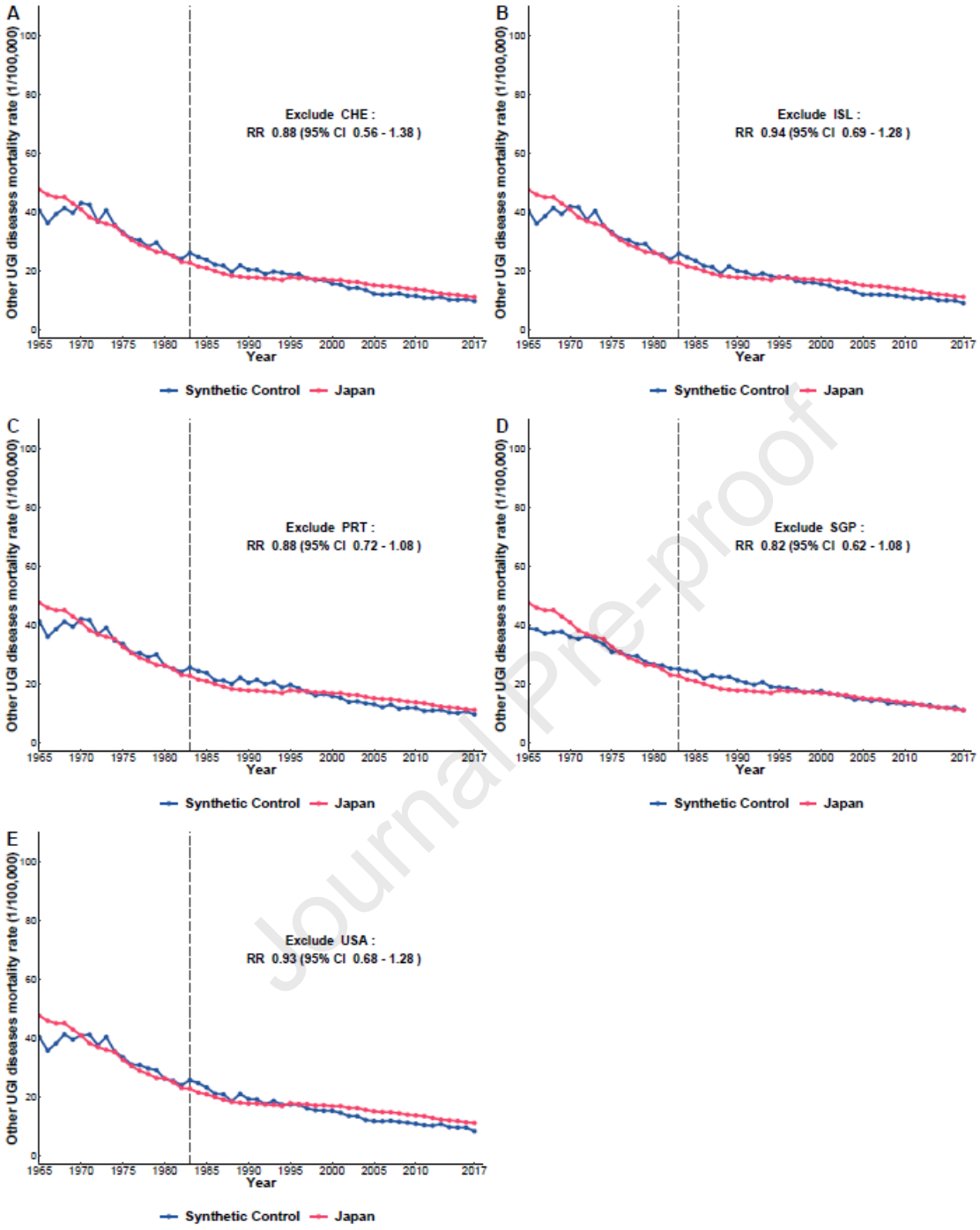


Figure S14. Sensitivity test: leave-out-one examination for the effect on other upper gastrointestinal disease mortality in Japan

UGI disease: upper gastrointestinal disease; CHE: Switzerland; ISL: Iceland; PRT: Portugal; SGP: Singapore; USA: United States; RR: rate ratio; CI: confidence interval

References

1. Lee S, Jun JK, Suh M, et al. Gastric cancer screening uptake trends in Korea: results for the National Cancer Screening Program from 2002 to 2011: a prospective cross-sectional study. Medicine (Baltimore) 2015;94:e533.
2. Sano H, Goto R, Hamashima C. What is the most effective strategy for improving the cancer screening rate in Japan? Asian Pac J Cancer Prev 2014;15:2607-12.
3. Goto R, Hamashima C, Mun S, et al. Why screening rates vary between Korea and Japan--differences between two national healthcare systems. Asian Pac J Cancer Prev 2015;16:395-400.
4. Hamashima C, Systematic Review G, Guideline Development Group for Gastric Cancer Screening G. Update version of the Japanese Guidelines for Gastric Cancer Screening. Jpn J Clin Oncol 2018;48:673-683.
5. Mabe K, Inoue K, Kamada T, et al. Endoscopic screening for gastric cancer in Japan: Current status and future perspectives. Dig Endosc 2022;34:412-419.
6. Ministry of Health Labour and Welfare. Report on Regional Public Health Services and Health Promotion Services, 2011.
7. Ministry of Health Labour and Welfare. Report on Regional Public Health Services and Health Promotion Services, 2017.
8. Ministry of Health Labour and Welfare. Report on Regional Public Health Services and Health Promotion Services, 2002.
9. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. J Am Stat Assoc 2010;105:493-505.
10. Bonander C. A (Flexible) Synthetic Control Method for Count Data and Other Nonnegative Outcomes. Epidemiology 2021;32:653-660.
11. Ben-Michael E, Feller A, Rothstein J. The Augmented Synthetic Control Method. J Am Stat Assoc 2021;116:1789-1803.

[dataset] 12. UNDP (United Nations Development Programme). Human Development Report 2021/2022, 2022. https://hdr.undp.org/data-center/documentation-and-downloads

[dataset] 13. The World Bank. Health Nutrition and Population Statistics, The World Bank DataBank, 2022. https://databank.worldbank.org/source/health-nutrition-and-population-statistics

[dataset] 14. Our World in Data. Average years of schooling, Our World in Data Dabase, 2022. https://ourworldindata.org/grapher/mean-years-of-schooling-long-run

[dataset] 15. Institute for Health Metrics and Evaluation (IHME), Global Burden of Disease Study 2019 (GBD 2019) Results, 2020. https://vizhub.healthdata.org/gbd-results/

16. Arkhangelsky D, Athey S, Hirshberg DA, et al. Synthetic Difference-in-Differences. Am Econ Rev 2021;111:4088-4118.
17. Kranz S. Synthetic Difference-in-Differences with Time-Varying Covariates. mimeo 2021. https://github.com/skranz/xsynthdid/blob/main/paper/synthdid_with_covariates.pdf