Table 2. Study characteristics.

Study ID (Study location)	Design; Recruitment period (*y*/*m*)	Setting	Exclusions	Subjects, *n*; (Male, %)	Mean age (range), *y*	Mean follow up, *y*	Cancer incidence rate, */100,000 person-year*	Ascertainment of gastric cancer development
Katsushika study [31,32] (Katsushika, Tokyo, Japan)	Prospective; 2000	Population-based health checkup	ND	4,490 (37)	47 (40–55)	3.9	46	Gastric cancer screening program registry and hospital record. Endoscopy recommended if PG test positive[a] or barium X-ray if negative.
Wakayama study [27,33] (Wakayama, Wakayama, Japan)	Prospective; 1994/4-1995/3	Workplace health checkup	Women; symptomatic patients; previous gastric resection; users of H2RAs or NSAIDs; gastric cancer diagnosed <1 year after surveillance (n = 8)	4,655[b] (100)	50 (40–59)	11.6[b]	161	Annual double-contrast barium X-ray and PG test followed by endoscopy +/− biopsy if either test positive
Watase 2004 [34] (Adachi, Tokyo, Japan)	Prospective; 1996	Population-based health checkup	Symptomatic patients; previous gastric resection; users of PPIs, patients with renal failure	5,449 (37)	51 (40–60)	4.8	58	Review of health checkup database and gastric cancer screening program registry. Endoscopy recommended if positive for PG test[c]
Watabe 2005 [35] (Chiba, Japan)	Prospective; 1995/3-1997/2	Opportunistic health checkup	Gastric cancer; peptic ulcer; and past history of gastrectomy	6,983 (68)	49 (ND)	4.7	130	Annual endoscopy (mean 5.1 times during the follow-up period)
Hisayama study [36,37] (Hisayama, Fukuoka, Japan)	Retrospective analysis of a prospective cohort; 1988	Population-based health checkup	Previous gastrectomy or gastric cancer; unavailable serum sample	2,446 (42)	57 (40-)	14[d]	260[e]	Records on annual health checkup and screening barium radiography; contact by mail or telephone; use of a daily monitoring system; hospital or clinic records on barium radiography, upper endoscopy, and histologic diagnosis; autopsies of subjects who died during the study period[f].
Kim 2008 [38] (Seoul, South Korea)	Prospective; 1992–1998	Opportunistic health checkup	ND	975 (90)	45 (ND)	9.9	21	Endoscopy every 1 to 3 years
Mizuno 2010 [39] (Kyoto, Japan)	Retrospective analysis of a prospective cohort; 1987	Population-based health checkup	ND	2,859 (35)	ND (35-)[g]	9.3[h]	229[i]	Cancer registry based on notification by local hospitals, gastric cancer screening, activities of public health nurses, and death certificates.
Zhang 2012 [40] (Zanhuang, Hebei, China)	Prospective; 1996–1997	Population-based health checkup	Gastric cancer; peptic ulcer; other severe diseases; and subjects with questionable *H. pylori* antibody results	1,501 (37)	45 (30-)	14	124[j]	Annual home visits and review of histology and X-rays from the local clinics and hospitals.

Table 2. Cont.

Study ID (Study location)	Design; Recruitment period (*y*/*n*)	Setting	Exclusions	Subjects, *n*; (Male, %)	Mean age (range), *y*	Mean follow up, *y*	Cancer incidence rate, *I 100,000 person-year*	Ascertainment of gastric cancer development
Okuno 2012 [41] (Kurobe, Toyama, Japan)	Prospective; 1995	Workplace health checkup	Age ≥60; Previous gastric cancer; gastric cancer <6 months after PG test (n = 3); no PG test results	4,383 (65)	45 (35–60)	12.3	111	Annual screening x-ray gastrography and/or endoscopy[k]. Self-report or physicians' report of gastric cancer confirmed through the correspondences with the testing institutions.

[a]Recommendation of endoscopy with biannual follow-up contact was offered if PG test positive.
[b]5,209 subjects with a mean follow-up of 9.7 years for the analysis of PG test only.
[c]Recommendation of endoscopy was offered annually for two years.
[d]10 years for the analysis of a 4-group risk model based on both PG test and H. pylori infection status.
[e]Approximately estimated based on 89 gastric cancer cases identified during the follow-up period of 14 years.
[f]Autopsy was performed 75% of all deaths from any causes.
[g]83% of participants were 74 years of age or younger.
[h]Median.
[i]Approximately estimated based on 61 gastric cancer cases identified during the median follow-up period of 9.3 years.
[j]Approximately estimated based on 26 gastric cancer cases identified during the followup period of 14 years.
[k]Total screening rates by x-ray gastrography and/or endoscopy were 78% in 1995, 71% in 1999, 75% in 2004, and 82% in 2009.
FY = fiscal year; H2RAs = histamine receptor 2 antagonist; ND = no data; NSAIDs = non-steroidal anti-inflammatory drugs; PG = pepsinogen; PPI = proton pomp inhibitor.
doi:10.1371/journal.pone.0109783.t002

Study Risk of Bias Assessment tool (PROBAST) [17]. One reviewer assessed study quality, and the rating was confirmed by at least one other reviewer. Three out of 64 (5%) quality ratings by the second reviewer were inconsistent. Any discrepant results were resolved by consensus.

Data synthesis and analysis

The predictive ability of the pepsinogen test and *H. pylori* antibodies as standalone tests were analyzed using the DerSimonian-Laird random effects model meta-analysis to obtain summary HRs with their corresponding 95% confidence intervals (CIs) for studies that reported time-to-event data in the main analysis and the Mantel-Haenszel fixed-effects model meta-analysis for sensitivity analyses. For studies that reported cumulative count data, we performed the Mantel-Haenszel fixed-effects meta-analysis to obtain summary odds ratios (ORs) with their corresponding 95% CIs in the main analysis because studies in general reported the incidence rates of gastric cancer in the test-negative group to be less than 1% with substantial imbalances between the test-positive and -negative groups [18]. The Peto OR method and the Mantel-Haenszel fixed-effects model for combining summary risk differences were also used in sensitivity analyses. To supplement the measures of predictive ability, we also obtained summary estimates of sensitivity and specificity with their corresponding 95% CIs using bivariate random effects meta-analysis with the exact binomial likelihood [19] and constructed summary receiver-operating characteristic (ROC) curves and confidence regions for summary sensitivity and specificity [20].

Studies that assessed the risk prediction model based on the pepsinogen test and *H. pylori* serology consistently defined four risk groups (**Table 1**). Suboptimal methodology and reporting of model performance are common in prognostic model studies using time-to-event data [21,22]. After perusal of the reported measures of model performance, we determined to quantitatively synthesize HRs across risk groups; no studies reported the recommended standard measures of discrimination or calibration [22]. From four risk strata, it is possible to form six pairwise comparisons. None of the studies, however, assessed and reported all the logically comparable contrasts but typically reported only three HRs of gastric cancer development, comparing Groups B, C, and D with Group A only. Therefore, in addition to conventional meta-analysis of direct evidence on the reported contrasts, we performed multivariate meta-analysis for predictive tests with three or more risk strata with a Bayesian framework to combine the totality of direct and indirect evidence in a single analysis, taking correlations between the risk strata into account [23,24]. We calculated the summary HRs and ORs (for cumulative count data) with their corresponding 95% credible intervals (CrIs) using the fixed-effects model in the main analysis and the random-effect model in sensitivity analysis. Additionally, we calculated the probability for each risk group that it would be ranked from best to worst among the four risk strata. Finally, we repeated the multivariate meta-analysis in a *post-hoc* set of sensitivity analyses by combining Group C and Group D to form a 3-risk group model (**Table 1**).

To quantitatively explore model performance with reported cumulative count data, we performed "descriptive" meta-analysis of the discrimination and calibration using the DerSimonian-Laird random-effects model [25], acknowledging not taking account of potential effects of censoring. For each study, as the measure of discrimination, we estimated the *c*-statistic and its corresponding 95% CIs [26]. To assess the calibration of the model, for each study we calculated the expected over observed event ratio (E/O) and its Poisson exact 95% CIs for each risk group and for all the risk groups combined. Expected events were calculated by

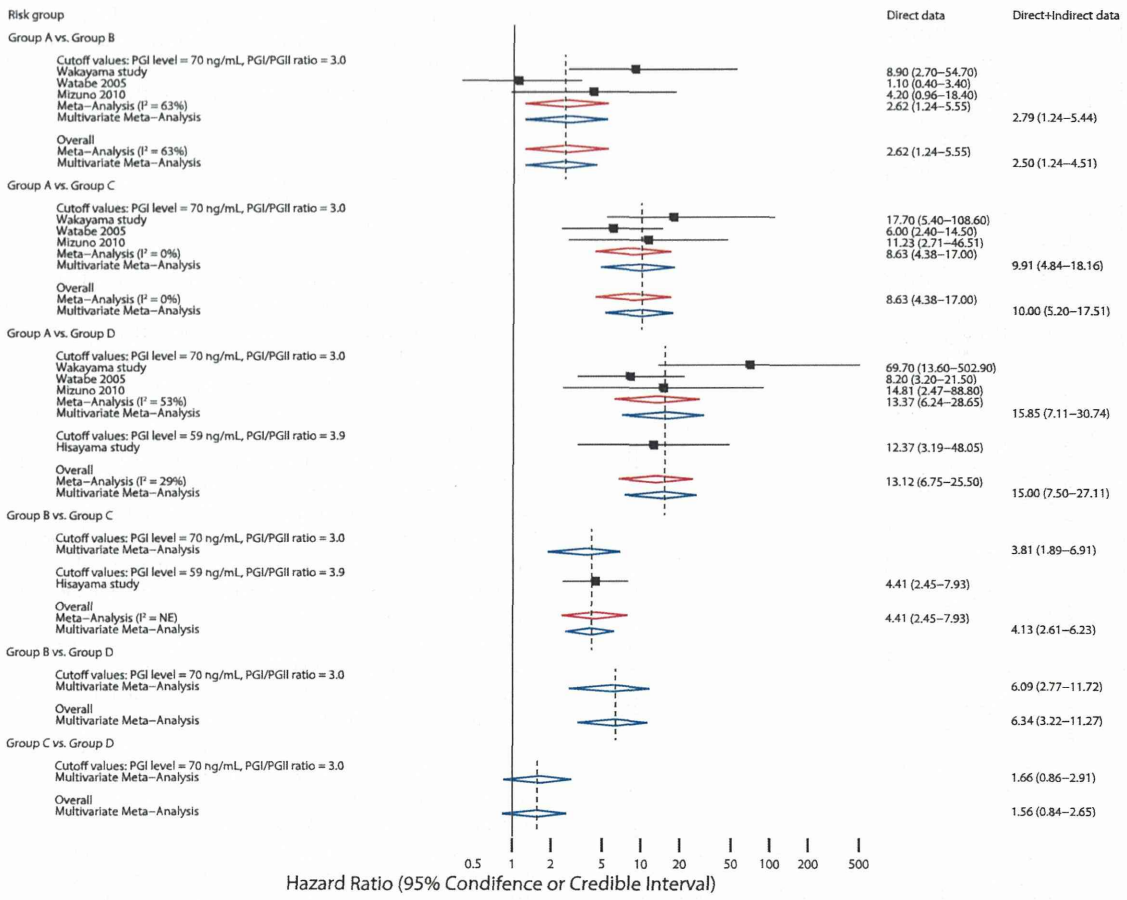

Figure 2. Meta-analysis of hazard ratio for four-risk-group prediction model to predict gastric cancer development. The red and blue diamonds depict a summary hazard ratio with extending 95% confidence interval (CI) or 95% credible interval (CrI), estimated from direct meta-analysis or multivariate meta-analysis, respectively. Each square and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study. NE = not estimable.
doi:10.1371/journal.pone.0109783.g002

applying the proportionate cumulative gastric cancer incidence estimates from long-term follow-up results of the first reported study [27] to the corresponding four risk groups of the subsequent studies assuming a constant incidence rate as reported [27]. E/O statistics less than, equal to, and more than 1 respectively suggest an under-, perfect-, and over-prediction of the model.

We quantified between-study heterogeneity with the I^2 statistic and considered I^2 to be suggestive of intermediate or high heterogeneity when >50% or >75%, respectively [28]. For each model in the Bayesian multivariate meta-analysis we based results on 3 different chains and 200,000 iterations after a burn-in of 10,000 iterations, and model convergence was assessed by Brooks-Gelman-Rubin criteria [29]. We did not perform tests for funnel plot asymmetry because there were fewer than ten eligible studies [30]. Also, we did not perform subgroup or meta-regression analyses due to the small number of studies. All analyses were conducted using Stata SE, version 12.1 (Stata Corp, College Station, TX, USA) and WinBUGS 1.4.3 (MRC Biostatistics Unit, Cambridge, UK). P-values for all comparisons were 2-tailed, and statistical significance was defined as a p-value less than 0.05.

Results

Literature flow and eligible studies

Our main literature searches identified 2843 citations, of which 154 were considered potentially eligible and reviewed in full (**Figure 1**). Six additional citations were identified through supplementary searches. We excluded 76 studies that did not meet our inclusion criteria. The updated search found three additional eligible studies. In the end, 9 independent cohorts reported in 12 publications [27,31–41] were considered eligible.

Study and clinical characteristics

The 9 eligible cohort studies (7 from Japan, 1 from Korea, and 1 from China) included 33,741 asymptomatic participants of gastric cancer screening programs (**Table 2**). Five studies [32,34,37,39,40] were conducted in communities, whereas two [35,38] were opportunistic screening in clinical settings, and another two [27,41] were workplace health checkups. Although all studies prospectively enrolled participants, two studies [37,39] reported that data were analyzed retrospectively. The mean age at enrollment ranged between 45 and 57 years, and the mean follow-up ranged between 3.9 and 14 years. During the study period, only between 2 and 89 gastric cancer cases were detected per cohort,

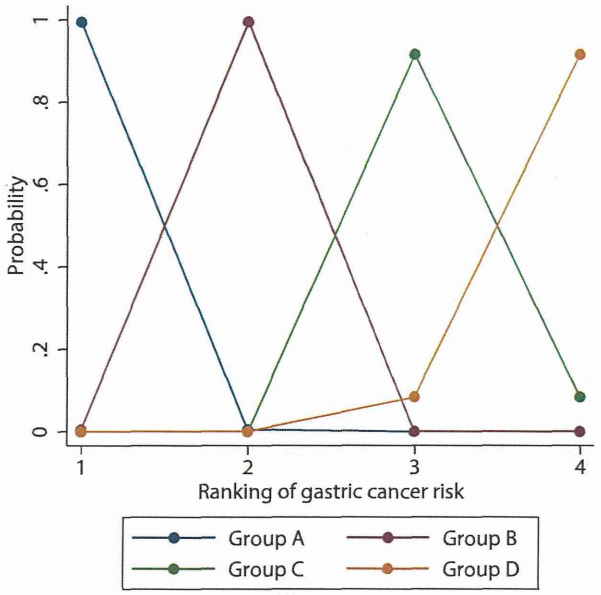

Figure 3. Rankogram of risk of gastric cancer development based on four-risk-group prediction model. Ranking probability of gastric cancer risk for each group, estimated from direct multivariate meta-analysis is shown. The 4 rankings show the risk of gastric cancer development: rank 1, lowest risk; rank 2, second lowest risk; rank 3, second highest risk; and rank 4, highest risk.
doi:10.1371/journal.pone.0109783.g003

which corresponded to heterogeneous cancer incidence rates of between 21 and 260 cases per 100,000 person-years. Only did 2 cohorts [27,35] analyze gastric cancer incidence by histological subtype (i.e., intestinal type or diffuse type). Two studies excluded from the analysis cases of gastric cancer diagnosed early after enrollment: 8 cases diagnosed within 1 year in one [27,33] and 3 cases diagnosed within 6 months in the other [41]. Review of the registry data on annual health checkups with radiographic screening and medical records was the most commonly adopted method to ascertain gastric cancer cases. Only in two studies [35,38] was periodic endoscopic screening performed to detect gastric cancer.

Three studies [32,34,41] evaluated the serum pepsinogen test alone, while a single study [38] exclusively assessed *H. pylori* antibodies as a standalone risk factor (**Table S1**). Five studies [27,35,37,39,40] evaluated both tests and the risk-prediction model, consisting of four risk strata based on the two tests. Of the seven studies that reported when samples were assayed, two analyzed the stored serum 7 to 14 years after blood collection. All seven studies that reported the method used to measure pepsinogen concentrations used an identical assay with a set of recommended cutoff values to diagnose CAG (pepsinogen I≤ 70 ng/mL and pepsinogen I/II≤3.0) [42]. Two studies adopted additional sets of cutoffs (**Table S1**). Various assays were used for *H. pylori* antibodies and heterogeneous estimates of sensitivity and specificity were reported (**Table S1**).

Assessment of study quality

Figure S1 shows the results of validity rating. No study adequately reported all seven items relevant to study validity that we assessed, that is, study design, selection of participants, participant characteristics, start of follow-up, test characteristics, methods of ascertainment of gastric cancer development, and methods of data analysis (**Table S2**). Reporting was particularly poor regarding blinding of interpreters of the two tests to clinical outcomes, and blinding of outcome assessors to the test results. Three studies [31,32,34,39] excluded more than 50% of all potentially eligible participants, and a retrospective design was adopted in 2 studies [36,37,39]. The follow-up period is shorter than 5 years in three studies [31,32,34,35]. Four studies [31,32,34,38,40] failed to adjust for any potential confounders in analyzing risk estimates.

Pepsinogen test and *H. pylori* antibodies

Four studies, including 14,343 subjects [33,37,39,41], reported HRs for the pepsinogen test to predict gastric cancer development. All studies but one [37] adopted the recommended cutoff values for this analysis. The random-effects meta-analysis showed that subjects with a positive test had a higher risk of gastric cancer than those with a negative test (summary HR, 3.5; 95% CI, 2.7–4.7; $p<0.001$; $I^2=0\%$) (**Figure S2-A**). Cumulative count data were available in 8 studies including 32,766 subjects [27,32,34–36,39–41]: a positive test result was similarly significantly associated with a higher risk of gastric cancer compared with a negative result (fixed-effects OR, 3.9; 95% CI, 3.2–4.8; $p<0.001$; $I^2=37\%$) (**Figure S2-B**). These studies had a summary sensitivity of 0.57 (95% CI, 0.49–0.65) and a summary specificity of 0.76 (95% CI, 0.69–0.81) (**Figure S2-C**).

For *H. pylori* antibodies, HR estimates were available from 3 studies including 9960 subjects [33,36,39]. The random-effects meta-analysis showed that subjects positive for *H. pylori* antibodies had a higher risk of gastric cancer than those with a negative test (summary HR, 3.2; 95% CI, 2.0–5.2; $p<0.001$; $I^2=0\%$) (**Figure S3-A**). Six studies including 19,419 subjects [27,35,37–40] reported cumulative count data for OR estimation, and the fixed-effects meta-analysis found a similarly significant association between positive *H. pylori* antibodies and a higher incidence of gastric cancer (summary OR, 2.7; 95% CI, 2.0–3.8; $p<0.001$; $I^2=10\%$) (**Figure S3-B**). Summary estimates of prognostic accuracy were 0.87 (95% CI, 0.76–0.94) for sensitivity and 0.30 (95% CI, 0.23–0.39) for specificity (**Figure S3-C**).

In the preplanned sensitivity analyses for these two tests, the summary estimates of the alternative models were not materially different from those in the main analysis (data not shown).

Risk prediction model

Predictive ability of the risk-prediction model based on the pepsinogen test and *H. pylori* antibodies was first reported in the Wakayama study of 2004 [33], where the baseline gastric cancer risk was estimated in a male population from a workplace health checkup. Four subsequent studies evaluated the model in three community-dwelling populations [35,37,40] and in a cohort of participants in opportunistic health checkups [39], which we considered validation cohorts.

Four studies (a total of 16,943 subjects) that reported HRs [27,35,37,39] were included in the meta-analysis of predictive ability. For predicting gastric cancer development, the 95% CrI of the summary HRs for 5 out of 6 possible contrasts did not include 1, suggesting that in the pairwise contrasts, other than the comparison between Group C and Group D, there was more than 95% probability that one of the two comparators had a higher risk of gastric cancer than the other (**Figure 2**). Specifically, multivariate meta-analyses suggested that Group A had a lower risk than Group B and Group C, and that compared with Group C and Group D, Group B had a lower risk. There was no significant difference in the risk of gastric cancer between Group C and Group D (summary HR, 1.49; 95% CrI: 0.84–2.65). The ranking

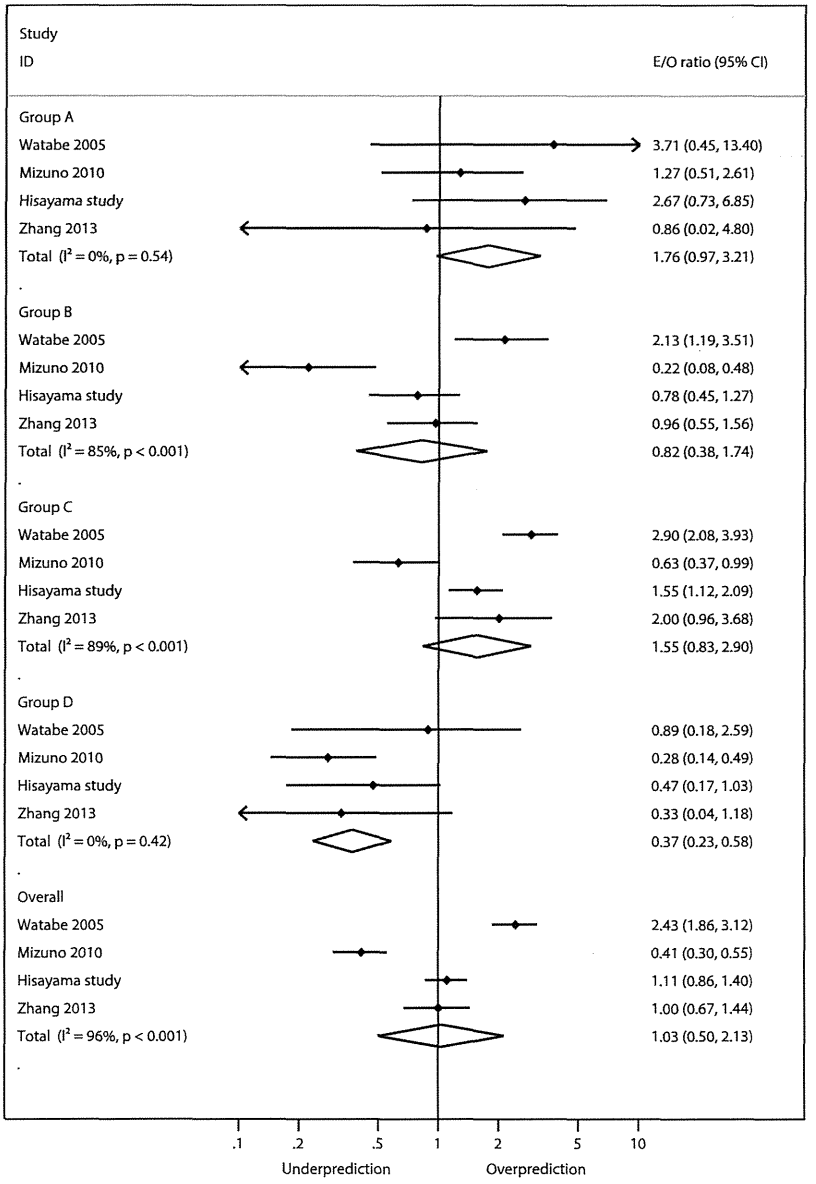

Figure 4. Meta-analysis of the expected over observed (E/O) ratios. The diamonds depict a summary E/O ratio and extending 95% confidence interval (CI). Each closed circle and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study. Studies are ordered by publication year.
doi:10.1371/journal.pone.0109783.g004

analysis showed that Groups A and B, respectively, had the lowest and second-lowest risk of gastric cancer development (posterior cumulative probability to rank the lowest and the second-lowest risk groups was both >99%), whereas Groups C and D could be the highest or second-highest risk groups (92% and 8%, respectively, for being ranked as the second-highest group, and 8% and 92%, respectively, for the highest risk group) (**Figure 3**). In sensitivity analyses using alternative models, and subgroup analyses of only studies that adopted the recommended cutoff values for the pepsinogen test, the summary HR estimates as well as the results of the ranking analysis were similar to those of the main analysis (**Figure S4**).

Five studies (a total of 18,444 subjects) with cumulative count data [27,35,37,39,40] were included in the multivariate meta-analysis of OR. The summary estimates were similar to the findings in the meta-analysis of HR, and again, there was no evidence of difference between Group C and Group D (summary OR, 1.64; 95% CrI: 0.84–2.88) (**Figure S5**). The summary estimates for sensitivity analyses were stable and the results were not materially different from the main analysis (**Figure S6**). In the *post-hoc* sensitivity analysis of 3-risk-strata model, the multivariate meta-analysis and the ranking analysis showed that Group A had a lower risk than Group B and combined Group C and Group D, and compared with combined Group C and Group D, Group B had a lower risk (**Fig. S7–S9**).

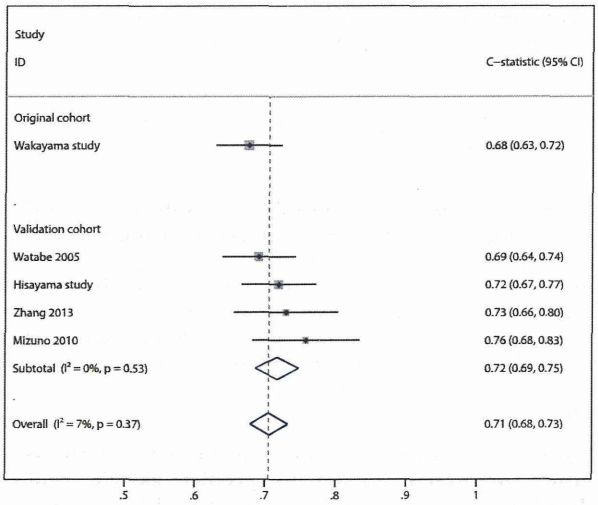

Figure 5. Meta-analysis of *c*-statistics. The diamonds depict a summary *c*-statistic and extending 95% confidence interval (CI). Each square and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study. The size of each square is proportional to the weight of each study in the meta-analysis. Studies are ordered by sample size.
doi:10.1371/journal.pone.0109783.g005

While two studies presented Kaplan-Meier plots of cumulative gastric cancer incidence by risk group [27,35] and four studies calculated p-values for differences in gastric cancer incidence between the risk strata by Log-rank test [27,35,39] or Chi-squared test [40], none reported recommended statistical measures or graphical displays for assessing model performance of time-to-event data [22]. Although the meta-analysis for overall study population suggested that the calibration was generally good across all risk strata (summary E/O ratio, 1.03; 95% CI: 0.50–2.13; p = 0.94), high between-study heterogeneity was found ($I^2 = 96\%$), suggesting that there were variations in the populations assessed in the validation studies (**Figure 4**). Specifically, the E/O ratio of one study showed an over-prediction (E/O, 2.43; 95% CI: 1.86–3.12; p<0.001), whereas an under-prediction was suggested for another study (E/O, 0.41; 95% CI: 0.30–0.55; p<0.001). In contrast, meta-analyses of the *c*-statistic suggested that the discrimination was in general fair with low evidence of between-study heterogeneity (summary *c*-statistic, 0.71; 95% CI: 0.68–0.73; $I^2 = 7\%$) (**Figure 5**).

Discussion

In this meta-analysis based on 9 prospective cohorts from Eastern Asia, we found that adults with a positive pepsinogen test, as a standalone test, had an approximately fourfold higher risk of gastric cancer than those with a negative test. Likewise, the risk of gastric cancer for those with positive *H. pylori* antibodies was about threefold higher than for those with a negative result. The performance of these tests did not seem to be different across the cohorts regardless of country or gastric cancer incidence. These findings are in general agreement with previous meta-analyses [9,10,43–46] based mostly on case-control and nested case-control studies, or cross-sectional studies.

In our multivariate meta-analysis, the prediction model seemed to be moderately accurate in separating asymptomatic adults into four risk groups. Although our results failed to show a significant difference between Group C and Group D, this should not be viewed as evidence that the risk of the two groups is equal because the lack of statistical significance may be due to small number of subjects categorized as Group D or events thereof.

Regarding the model performance, our descriptive meta-analysis found that the fair discriminatory performance reported from the first cohort seemed to be retained across the subsequent studies, whereas the calibration was not consistently validated, suggesting clinical heterogeneity across studies. One explanation could be that different screening settings enrolled different populations. Another might be variability in study design including different methodologies for diagnosing gastric cancer, follow-up time, and exclusion criteria adopted in the original studies.

Our study has several limitations. First, our meta-analysis is based on a small number of studies exclusively from Eastern Asia. Thus, our findings may not be generalizable to the populations in other regions. Second, our descriptive assessment of model performance is exploratory, based on the available cumulative count data with inconsistent follow-up periods and heterogeneous methods adopted to verify gastric cancer cases. Assessing how these affect the model performance would need data at the level of the individual. Third, the small number of eligible studies precluded subgroup analyses or meta-regression for *H. pylori* antibody assays. Therefore, how each different assay affects the results is unclear. Fourth, *H. pylori* and gastric atrophy are generally believed to be more relevant in the pathogenesis of intestinal type gastric cancer [7]. Few studies with the pertinent information precluded the subgroup analyses by histological subtype. Lastly, publication bias is still of concern because our searches were limited to the English- and Japanese-language literature.

Despite its development without formal statistical modeling and the paucity of rigorous external validation, the four-risk-group model has already been implemented in several screening programs including both private and public organizations in Japan. Given that the model is simple and both tests are easy to administer with minimal discomfort, the rapid acceptance is not surprising. A risk-stratified two-stage screening program incorporating the four-risk-group model may hold the promise of remedying the current low cancer screening rates; the risk model could efficiently select "high-risk" populations that would need a conventional screening modality while allowing those identified with a lower risk to omit the painful conventional tests. Notwithstanding these theoretical advantages, comparative evidence on clinically important outcomes such as improvements in gastric-cancer-specific mortality regarding the model-incorporated "stepwise" screening strategy compared with conventional strategies is still lacking and the consequences of withholding conventional screening tests from those labeled "low risk" by the model are unclear.

In summary, the serum pepsinogen test, *H. pylori* antibodies, and the four-risk-group prediction model seem to have the potential to stratify middle-aged presumptively healthy adults in Eastern Asia for predicting the risk of gastric cancer. Before wider implementation in daily practice, to understand how these two tests and the risk model in particular will affect clinically important outcomes of screened populations, future research needs to focus on comparative studies to evaluate the impact of screening programs adopting the risk model. Given the challenges in conducting randomized trials, a decision modeling analysis incorporating information on the risk model as well as data on effectiveness of therapeutic interventions would be a realistic first step to take [47]. However, even if the modeling analysis is positive, we should not automatically discard the possibility of generating randomized comparative evidence as in other cancer

screening fields [48]. In addition, given the variable prevalence of *H. pylori* infection across different generations and different countries [49], and also the recent introduction of eradication therapies, both of which are expected to affect the test results, validation of the current model performance is still necessary [47].

Supporting Information

Figure S1 Quality assessment of studies included in the meta-analysis. The stacked bar charts illustrate quality rating for risk of bias for predictive factor studies by the Quality In Prognosis Studies (QUIPS-2) tool (A) [16], and risk of bias (B) and concerns about applicability (C) for studies of both predictive factor and risk prediction model by the Prediction Study Risk of Bias Assessment tool (PROBAST) [17]. The percentages of studies that met the given ratings for each domain are shown.
(EPS)

Figure S2 Meta-analysis of hazard ratio (A), odds ratio (B), and sensitivity and specificity (C) for the pepsinogen test to predict gastric cancer development. The diamonds depict the summary hazard ratio (A) or odds ratio (B) and extending 95% confidence interval (CI). Each square and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study. The size of the square is proportional to the weight of each study in the meta-analysis. Studies are ordered by sample size. Individual study estimates of sensitivity and specificity are plotted in the receiver operating characteristic (ROC) space (C). The size of each circle is proportional to the sample size for each study (all study participants). The dashed crescent boundary represents the 95% confidence region for the summary sensitivity and specificity (shown as the square). The solid line represents the summary ROC curve.
(EPS)

Figure S3 Meta-analysis of hazard ratio (A), odds ratio (B), and sensitivity and specificity (C) for *H. pylori* antibodies to predict gastric cancer development. The diamonds depict the summary hazard ratio (A) or odds ratio (B) and extending 95% confidence interval (CI). Each square and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study. The size of the square is proportional to the weight of each study in the meta-analysis. Studies are ordered by sample size. Individual study estimates of sensitivity and specificity are plotted in the receiver operating characteristic (ROC) space (C). The size of each circle is proportional to the sample size for each study (all study participants). The dashed crescent boundary represents the 95% confidence region for the summary sensitivity and specificity (shown as the square). The solid line represents the summary ROC curve.
(EPS)

Figure S4 Sensitivity analysis for multivariate meta-analysis of hazard ratio for the four-risk-group prediction model. The red and blue diamonds and horizontal lines depict a summary hazard ratio and corresponding 95% credible interval (CrI), estimated from the fixed- or random-effects multivariate meta-analysis, respectively. Subgroup results for studies that adopted the conventional cutoff for pepsinogen levels are also shown.
(EPS)

Figure S5 Meta-analysis of odds ratio for four-risk-group prediction model to predict gastric cancer development. The red and blue diamonds depict a summary odds ratio with extending 95% confidence interval (CI) or 95% credible interval (CrI), estimated from direct meta-analysis or multivariate meta-analysis, respectively. Each square and horizontal line indicates the odds ratio and corresponding 95% CI, respectively, for each study.
(EPS)

Figure S6 Sensitivity analysis for multivariate meta-analysis of odds ratio for the four-risk-group prediction model. The red and blue diamonds and horizontal lines depict a summary odds ratio and corresponding 95% credible interval (CrI), estimated from the fixed- or random-effects multivariate meta-analysis, respectively. Subgroup results for studies that adopted the conventional cutoff for pepsinogen levels are also shown.
(EPS)

Figure S7 Meta-analysis of odds ratio for three-risk-group prediction model to predict gastric cancer development. The red and blue diamonds depict a summary odds ratio with extending 95% confidence interval (CI) or 95% credible interval (CrI), estimated from direct meta-analysis or multivariate meta-analysis, respectively. Each square and horizontal line indicates the hazard ratio and corresponding 95% CI, respectively, for each study.
(EPS)

Figure S8 Sensitivity analysis for multivariate meta-analysis of odds ratio for the three-risk-group prediction model. The red and blue diamonds and horizontal lines depict a summary odds ratio and corresponding 95% credible interval (CrI), estimated from the fixed- or random-effects multivariate meta-analysis, respectively. Subgroup results for studies that adopted the conventional cutoff for pepsinogen levels are also shown.
(EPS)

Figure S9 Rankogram of risk of gastric cancer development based on three-risk-group prediction model. Ranking probability of gastric cancer risk for each group, estimated from direct multivariate meta-analysis is shown. The 3 rankings show the risk of gastric cancer development: rank 1, lowest risk; rank 2, second lowest risk; rank 3, highest risk.
(EPS)

Table S1 Test characteristics.
(DOCX)

Table S2 Quality assessment of included studies.
(DOCX)

Checklist S1 PRISMA Checklist.
(DOC)

Acknowledgments

We thank Drs. Fumiko Ito and Yoshiyuki Watanabe for providing unpublished data on their original work, and Dr. Robert Wolff of the PROBAST steering group for providing an earlier version of the recently developed PROBAST tool.

Author Contributions

Conceived and designed the experiments: TT CH. Performed the experiments: TT HN KK IM TY RT CH. Analyzed the data: TT HN KK CH. Contributed reagents/materials/analysis tools: TT. Contributed to the writing of the manuscript: TT HN KK IM TY RT CH. Collection of data: TT HN KK IM TY RT CH.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, et al. (2013) GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. Lyon, France: International Agency for Research on Cancer. Available: http://globocan.iarc.fr/Default.aspx. Accessed 2013 Dec 29.
2. Bray F, Ren JS, Masuyer E, Ferlay J (2013) Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer 132: 1133–1145.
3. The Steering Committee for the Development of Gastric Cancer Screening Guidelines (2013) The Japanese Guidelines for Gastric Cancer Screening. Available: http://canscreen.ncc.go.jp/. Accessed 2013 Dec 29.
4. Leung WK, Wu MS, Kakugawa Y, Kim JJ, Yeoh KG, et al. (2008) Screening for gastric cancer in Asia: current evidence and practice. Lancet Oncol 9: 279–287.
5. Foundation for Promotion of Cancer Research. Cancer Statistics in Japan - 2013. Available: http://ganjoho.jp/data/professional/statistics/backnumber/2013/cancer_statistics_2013.pdf. Accessed 2013 Dec 29.
6. World Health Organization (2012) Biological agents. Volume 100 B. A review of human carcinogens. IARC monographs on the evaluation of carcinogenic risks to humans/World Health Organization, International Agency for Research on Cancer 100: 1–441. Available: http://monographs.iarc.fr/ENG/Monographs/vol100B/mono100B-1.pdf. Accessed 2013 Dec 29.
7. Correa P, Piazuelo MB (2012) The gastric precancerous cascade. J Dig Dis 13: 2–9.
8. Polk DB, Peek RM Jr (2010) Helicobacter pylori: gastric cancer and beyond. Nat Rev Cancer 10: 403–414.
9. Islami F, Sheikhattari P, Ren JS, Kamangar F (2011) Gastric atrophy and risk of oesophageal cancer and gastric cardia adenocarcinoma–a systematic review and meta-analysis. Ann Oncol 22: 754–760.
10. Dinis-Ribeiro M, Yamaki G, Miki K, Costa-Pereira A, Matsukawa M, et al. (2004) Meta-analysis on the validity of pepsinogen test for gastric carcinoma, dysplasia or chronic atrophic gastritis screening. J Med Screen 11: 141–147.
11. Weck MN, Brenner H (2008) Association of Helicobacter pylori infection with chronic atrophic gastritis: Meta-analyses according to type of disease definition. Int J Cancer 123: 874–881.
12. Adamu MA, Weck MN, Gao L, Brenner H (2010) Incidence of chronic atrophic gastritis: systematic review and meta-analysis of follow-up studies. Eur J Epidemiol 25: 439–448.
13. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG (2009) Prognosis and prognostic research: what, why, and how? BMJ 338: b375.
14. Hamashima C, Saito H, Nakayama T, Sobue T (2008) The standardized development method of the Japanese guidelines for cancer screening. Jpn J Clin Oncol 38: 288–295.
15. Rector TS, Taylor BC, Wilt TJ (2012) Chapter 12: systematic review of prognostic tests. J Gen Intern Med 27 Suppl 1: S94–101.
16. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C (2013) Assessing bias in studies of prognostic factors. Ann Intern Med 158: 280–286.
17. PROBAST steering committee (2013) Prediction risk of bias assessment tool (PROBAST). Available: http://colloquium.cochrane.org/fr/abstracts/prediction-study-risk-bias-assessment-tool-probast. Accessed 2013 Dec 29.
18. The Agency for Healthcare Research and Quality (2013) Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD). Available: http://www.ncbi.nlm.nih.gov/pubmed/21433403. Accessed 2013 Dec 29.
19. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, et al. (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 58: 982–990.
20. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA (2007) A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 8: 239–251.
21. Mallett S, Royston P, Waters R, Dutton S, Altman DG (2010) Reporting performance of prognostic models in cancer: a review. BMC Med 8: 21.
22. Royston P, Altman DG (2013) External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 13: 33.
23. Lu G, Ades AE (2004) Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 23: 3105–3124.
24. Woods BS, Hawkins N, Scott DA (2010) Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. BMC Med Res Methodol 10: 54.
25. Meads C, Ahmed I, Riley RD (2012) A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. Breast Cancer Res Treat 132: 365–377.
26. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44: 837–845.
27. Yoshida T, Kato J, Inoue I, Yoshimura N, Deguchi H, et al. (2013) Cancer development based on chronic active gastritis and resulting gastric atrophy as assessed by serum levels of pepsinogen and Helicobacter pylori antibody titer. Int J Cancer 134: 1445–1457.
28. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327: 557–560.
29. Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 7: 434–455.
30. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, et al. (2011) Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 343: d4002.
31. Sugishita Y., Ito F, Seiko A (2003) [Evaluation of Two Steps Gastric Cancer Detection Program Using Serum Pepsinogen Test Applied for Urban Residents Living in Katsushika Ward of Tokyo from 2000 to 2001]. Nihon Gan Kenshin Shindan Gakkaishi: 10: 161–166. Japanese.
32. Ito F, Watanabe Y, Miki K (2007) [Effect of the Two-step Serum Pepsinogen Test Method on Reducing Stomach Cancer Mortality among the Urban Residents]. Nihon Gan Kenshin Shindan Gakkaishi 14: 156–160. Japanese.
33. Ohata H, Kitauchi S, Yoshimura N, Mugitani K, Iwane M, et al. (2004) Progression of chronic atrophic gastritis associated with Helicobacter pylori infection increases risk of gastric cancer. Int J Cancer 109: 138–143.
34. Watase H, Inagaki T, Yoshikawa I, Furihata S, Watanabe Y, et al. (2004) [Five years follow up study of gastric cancer screening using the pepsinogen test method in Adachi city]. Nihon Gan Kenshin Shindan Gakkaishi 11: 77–81. Japanese.
35. Watabe H, Mitsushima T, Yamaji Y, Okamoto M, Wada R, et al. (2005) Predicting the development of gastric cancer from combining Helicobacter pylori antibodies and serum pepsinogen status: a prospective endoscopic cohort study. Gut 54: 764–768.
36. Oishi Y, Kiyohara Y, Kubo M, Tanaka K, Tanizaki Y, et al. (2006) The serum pepsinogen test as a predictor of gastric cancer: the Hisayama study. Am J Epidemiol 163: 629–637.
37. Shikata K, Kiyohara Y, Kubo M, Yonemoto K, Ninomiya T, et al. (2006) A prospective study of dietary salt intake and gastric cancer incidence in a defined Japanese population: the Hisayama study. Int J Cancer 119: 196–201.
38. Kim N, Park RY, Cho SI, Lim SH, Lee KH, et al. (2008) Helicobacter pylori infection and development of gastric cancer in Korea: long-term follow-up. J Clin Gastroenterol 42: 448–454.
39. Mizuno S, Miki I, Ishida T, Yoshida M, Onoyama M, et al. (2010) Prescreening of a high-risk group for gastric cancer by serologically determined Helicobacter pylori infection and atrophic gastritis. Dig Dis Sci 55: 3132–3137.
40. Zhang X, Xue L, Xing L, Wang J, Cui J, et al. (2012) Low serum pepsinogen I and pepsinogen I/II ratio and Helicobacter pylori infection are associated with increased risk of gastric cancer: 14-year follow up result in a rural Chinese community. Int J Cancer 130: 1614–1619.
41. Okuno T, Kido T, Sakurai M, Nakamura K, Morikawa Y, et al. (2012) A 15-year cohort study on the incidence of gastric cancer and the validity of testing based on serum pepsinogen screening test. Journal of the Tsuruma Health Science Society Kanazawa University 36: 15–23.
42. Miki K (2006) Gastric cancer screening using the serum pepsinogen test method. Gastric Cancer 9: 245–253.
43. Xue FB, Xu YY, Wan Y, Pan BR, Ren J, et al. (2001) Association of H. pylori infection with gastric carcinoma: a Meta analysis. World J Gastroenterol 7: 801–804.
44. Helicobacter and Cancer Collaborative Group (2001) Gastric cancer and Helicobacter pylori: a combined analysis of 12 case control studies nested within prospective cohorts. Gut 49: 347–353.
45. Eslick GD, Lim LL, Byles JE, Xia HH, Talley NJ (1999) Association of Helicobacter pylori infection with gastric carcinoma: a meta-analysis. Am J - Gastroenterol 94: 2373–2379.
46. Huang JQ, Sridhar S, Chen Y, Hunt RH (1998) Meta-analysis of the relationship between Helicobacter pylori seropositivity and gastric cancer. Gastroenterol 114: 1169–1179.
47. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, et al. (2012) Risk prediction models: II. External validation, model updating, and impact assessment. Heart 98: 691–698.
48. Bretthauer M, Hoff G (2012) Comparative effectiveness research in cancer screening programmes. BMJ 344: e2864.
49. Brown LM (2000) Helicobacter pylori: epidemiology and routes of transmission. Epidemiol Rev 22: 283–297.

RESEARCH ARTICLE

What is the Most Effective Strategy for Improving the Cancer Screening Rate in Japan?

Hiroshi Sano[1]*, Rei Goto[2], Chisato Hamashima[3]

Abstract

Background: Cancer screening rates in Japan are much lower than those in Western countries. This study evaluated the relationship between cancer screening rates and strategies used to improve screening rates, and determined which strategy is the most effective. Materials and Methods: All municipalities are responsible for conducting gastric, lung, colorectal, cervical, and breast cancer screenings in Japan. Of the 1,746 municipalities in total, 92-99% were included in the analyses for each cancer screening. Using national data in 2009, the correlations between cancer screening rates and strategies for improving screening rates of all municipalities, both large (populations of over 30,000) and small (populations of under 30,000), were determined. The strategies used were as follows: sending personal invitation letters, personal visits by community health workers, use of a clinical setting for screening, and free screening. Results: Of all four strategies used to improve cancer screening rates, sending personal invitation letters had the highest correlations with all screening rates, with the exception of breast cancer screening. The partial correlation coefficients linking this strategy with the screening rates in all municipalities were 0.28, 0.32, 0.30, and 0.26 for gastric, lung, colorectal, and cervical cancer screening, respectively. In large municipalities, the correlations between the number of examinees in a clinical setting and the screening rates were also relatively high, particularly for cervical cancer screening (r=0.41). Conclusions: Sending personal invitation letters appears to be particularly effective in improving cancer screening rates in all municipalities. All municipalities should implement a system that sends personal invitation letters for cancer screening. In large municipalities, increasing the availability of screening in a clinical setting is also effective in improving cancer screening rates.

Keywords: Cancer screening - screening rate - strategy to increase participation- correlation - Japan

Asian Pac J Cancer Prev, **15 (6)**, 2607-2612

Introduction

Since population-based screening for cancer was introduced under the Health and Medical Service Act for the elderly in 1983, municipalities have been responsible for conducting cancer screenings in Japan. Screening programs for five kinds of cancers (gastric, lung, colorectal, cervical, and breast cancers) have become continuously conducted by all municipalities. However, cancer screening rates in Japan are much lower than those in Western countries and Korea, including examinations other than population-based screening that are conducted as part of a public policy to reduce mortality rates. While the screening rates for breast and cervical cancer in 2010 were 80.4% and 85.0%, respectively, in the United States, 70.9% and 67.9% in Korea, and 73.4% and 78.5%, in the United Kingdom, both screening rates were 24.3% in Japan (OECD, 2011; Suh et al., 2013).

To improve cancer screening rates, effective strategies that motivate people to be screened need to be successfully implemented. The U.S. Center for Disease Control and Prevention (CDC) conducted systematic reviews on the effectiveness of various interventions in increasing the screening rates for breast, cervical, and colorectal cancers, and published guidelines based on their findings, which recommend certain interventions for improving the screening rates for these cancers (Baron et al., 2008a; Sabatino et al., 2012; Community Preventive Services Task Force (CPSTF), 2013). The guidelines also aid decision makers in choosing an appropriate intervention (Townsend et al., 2009; Blumenthal et al., 2010; Lobb et al., 2011; Hannon et al., 2012).

In Japan, there are no guidelines on the types of strategies that improve cancer screening rates. Some studies have previously evaluated the effectiveness of various strategies (Hisamichi et al., 1991; Watanabe, 2003; Shimada et al., 2010a; Shimada et al., 2010b; Matsuda et al., 2011; Takaku, 2011; Kuroki, 2012; Yoshida et al., 2012), but it was difficult to compare the effectiveness of these strategies, as each study focused on

*[1]Faculty of Economics, Shiga University, Shiga, [2]The Hakubi Center of Advanced Research, Graduate School of Economics, Kyoto University, Kyoto, [3]Cancer Screening Assessment and Management Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan *For correspondence: hi-sano@biwako.shiga-u.ac.jp*

the effectiveness of an individual strategy using different subjects and methodologies. As the most effective strategy in improving cancer screening rates differs depending on the country and region (McAvoy and Raza, 1991; King et al., 1994; Saywell et al., 1999; Champion et al., 2003; Saywell et al., 2003; Saywell et al., 2004; Blumenthal at al., 2010; Lee et al., 2012; Frie et al., 2013), it remains unclear which strategy would be the most effective in Japan. Therefore, a study comparing the effectiveness of different strategies used to improve cancer screening rates in Japan is warranted and poised to be very useful for decision makers.

The aim of the present study was to quantitatively evaluate the relationships between cancer screening rates and strategies used to improve screening rates, as well as to determine which strategy is the most effective in Japan.

Materials and Methods

Subjects

The subjects were selected from a total of 1,746 municipalities that conducted gastric, lung, colorectal, cervical, and breast cancer screening in Japan. Cancer screening rates of municipalities were determined from data in the Report on Regional Public Health Services and Health Promotion Services between April 2009 and March 2010, which was prepared by the Ministry of Health, Labour, and Welfare (MHLW, 2010). In this report, the number of participants and persons eligible for the cancer screenings was tallied by sex and age in 1,746 municipalities. Persons eligible for the cancer screenings conducted by municipalities included women aged ≥20 years for cervical cancer screening, women aged ≥40 years for breast cancer screening, and both men and women aged ≥40 years for other cancer screenings. Using this report, the following characteristics of municipalities were determined: the number of eligible persons, the ratio of males to females, and percentage of those aged ≥65 years.

Data on strategies implemented by each municipality for cancer screening were obtained from a survey on the implementation of cancer screening among the different municipalities, which was conducted by the MHLW in January 2010. In this survey, the MHLW collected data on the content of examinations, strategies, and out-of-pocket costs for cancer screening among the different municipalities. 1,740 of all municipalities (99.7%) had responded to this survey. The CDC recommends interventions that use client reminders and small media, and interventions that include one-on-one education by telephone or via face-to-face encounters for colorectal, cervical, and breast cancer screening (Sabatino et al., 2012). It also recommends interventions that make screening accessible and easier for colorectal and breast cancer, and reduce out-of-pocket costs for breast cancer screening (Baron et al., 2008a; Sabatino et al., 2012; CPSTF, 2013). Based on these recommendations, similar strategies were assessed, in particular: sending personal invitation letters, personal visitations by community health workers, number of individuals screened in a clinical setting, and free screening. The use of newsletters in place of small media was not evaluated because about 90% of municipalities already implemented this strategy.

Table 1. Characteristics of Cancer Screening in Japan between April 2009 and March 2010

Variable	Gastric	Lung	Colorectal	Cervical	Breast
			All municipalities		
Number of municipalities	1,718	1,610	1,726	1,717	1,693
Screening rate(%); mean (S.D.)	15.8 (12.0)	27.4 (18.9)	21.4 (13.6)	16.9 (10.4)	13.2 (10.8)
Strategies					
Sending personal invitation letters; n (%)	946 (55.1)	889 (55.2)	947 (54.9)	966 (56.3)	933 (55.1)
Personal visitations by community health workers; n (%)	105 (6.1)	99 (6.2)	107 (6.2)	104 (6.1)	102 (6.0)
Number of individuals screened in clinical settings; mean (S.D.)	613 (2,891)	1,369 (6,582)	1,947 (6,967)	1,691 (5,091)	759 (2,706)
Free screening; n (%)	143 (8.3)	362 (22.5)	167 (9.7)	161 (9.4)	119 (7.0)
Characteristics of eligible persons					
Number of eligible persons; mean (S.D.)	22,315 (47,190)	22,821 (43,855)	22,946 (49,320)	18,701 (41,438)	13,747 (29,503)
Ratio of males to females; mean (S.D.)	0.73 (0.16)	0.72 (0.17)	0.73 (0.17)	-	-
Percentage of those aged ≥65 years; mean (S.D.)	52.9 (12.1)	53.4 (12.3)	53.0 (12.1)	42.6 (13.0)	53.2 (11.9)
			Large municipalities (population ≥30,000)		
Number of municipalities	809	767	812	808	800
Screening rate(%); mean (S.D.)	12.3 (9.0)	22.1 (15.7)	18.3 (10.9)	15.2 (8.1)	12.2 (8.6)
Strategies					
Sending personal invitation letters; n (%)	407 (50.3)	396 (51.6)	416 (51.2)	422 (52.2)	406 (50.8)
Personal visitations by community health workers; n (%)	24 (3.0)	21 (2.7)	24 (3.0)	22 (2.7)	22 (2.8)
Number of individuals screened in clinical settings; mean (S.D.)	1,255 (4,116)	2,806 (9,326)	4,052 (9,736)	3,455 (7,011)	1,538 (3,786)
Free screening; n (%)	70 (8.7)	162 (21.1)	79 (9.7)	75 (9.3)	45 (5.6)
Characteristics of eligible persons					
Number of eligible persons; mean (S.D.)	42,401 (62,911)	43,018 (57,003)	43,709 (65,937)	36,094 (55,436)	26,134 (39,346)
Ratio of males to females; mean (S.D.)	0.69 (0.17)	0.68 (0.17)	0.69 (0.17)	-	-
Percentage of those aged ≥65 years; mean (S.D.)	52.0 (12.4)	52.3 (12.5)	52.0 (12.4)	39.8 (12.2)	51.5 (11.4)
			Small municipalities (population <30,000)		
Number of municipalities	909	843	914	909	893
Screening rate(%); mean (S.D.)	18.9 (13.4)	32.2 (20.2)	24.1 (15.0)	18.4 (11.9)	14.2 (12.4)
Strategies					
Sending personal invitation letters; n (%)	539 (59.3)	493 (58.5)	531 (58.1)	544 (59.9)	527 (59.0)
Personal visitations by community health workers; n (%)	81 (8.9)	78 (9.3)	83 (9.1)	82 (9.0)	80 (9.0)
Number of individuals screened in clinical settings; mean (S.D.)	41 (175)	62 (263)	78 (285)	122 (228)	62 (127)
Free screening; n (%)	73 (8.0)	200 (23.7)	88 (9.6)	86 (9.5)	74 (8.3)
Characteristics of eligible persons					
Number of eligible persons; mean (S.D.)	4,439 (3,126)	4,445 (3,084)	4,500 (3,124)	3,241 (2,321)	2,650 (1,864)
Ratio of males to females; mean (S.D.)	0.77 (0.15)	0.76 (0.16)	0.77 (0.16)	-	-
Percentage of those aged ≥65 years; mean (S.D.)	53.8 (11.8)	54.5 (12.1)	53.9 (11.8)	45.0 (13.1)	54.8 (12.1)

Municipalities were excluded from the study if there were missing values in these variables or <10 eligible persons. Furthermore, municipalities were also excluded if they did not perform the following examinations: gastric X-ray for gastric cancer, chest X-ray for lung cancer, fecal occult blood tests for colorectal cancer, Pap smear for cervical cancer, and mammography for breast cancer. These examinations are recommended for population-based screening as there is sufficient evidence to suggest that these tests reduce the cancer mortality rate in Japan (Hamashima et al., 2008; Hamashima et al., 2010; National cancer center, 2013). Of all municipalities, 1,718 (98.4%), 1,610 (92.2%), 1,726 (98.9%), 1,717 (98.3%), and 1,693 (97.0%) municipalities were included in the analyses for gastric, lung, colorectal, cervical, and breast cancer screening, respectively.

Statistical analysis

Partial correlation coefficients were calculated to quantitatively evaluate the relationships between cancer screening rates and the strategies used to improve screening rates in various municipalities. The coefficients indicate how closely each strategy is related to the cancer screening rate after excluding the effects of confounding factors, including the other three strategies, the number of eligible persons, the ratio of males to females, and the percentage of elderly.

The relationships between cancer screening rates and the strategies may vary with the population size of the municipalities. Therefore, partial correlation coefficients were also separately calculated for large municipalities (with populations of over 30,000) and small municipalities (with populations of under 30,000). In 2009, a municipality was seen as a city if the population was over 30,000, but seen as a town or village if not. All analyses were performed using STATA version 12 (StataCorp, College Station, TX, USA).

Results

The characteristics of cancer screening in Japan are presented in Table 1. The average screening rates for gastric, lung, colorectal, cervical, and breast cancer were 15.8%, 27.4%, 21.4%, 16.9%, and 13.2%, respectively. The strategy of sending invitation letters was implemented at about 55% of the municipalities, whereas personal visitations by community health workers were implemented at only 6% of all municipalities. Free screening was implemented at 23% of all municipalities for lung cancer screening and at 7-10% of all municipalities

Table 2. Partial Correlations between Cancer Screening Rates and Strategies Used to Improve Screening Rates in Japan

Variable	Gastric	Lung	Colorectal	Cervical	Breast
			All municipalities		
Number of municipalities	1,718	1,610	1,726	1,717	1,693
Strategies					
Sending personal invitation letters d	0.28[a]	0.32[a]	0.30[a]	0.26[a]	0.13[a]
Personal visitations by community health workers d	0.23[a]	0.15[a]	0.22[a]	0.18[a]	0.12[a]
Number of individuals screened in clinical settings	0.17[a]	0.19[a]	0.21[a]	0.25[a]	0.18[a]
Free screening d	0.03	0.13[a]	0.06[b]	0.08[a]	0.01
Characteristics of eligible persons					
Number of eligible persons	-0.24[a]	-0.26[a]	-0.23[a]	-0.28[a]	-0.21[a]
Ratio of males to females	0.07[a]	0.09[a]	0.03	-	-
Percentage of those aged ≥65 years	0.07[a]	0.20[a]	0.12[a]	-0.01	-0.13[a]
			Large municipalities (population ≥30,000)		
Number of municipalities	809	767	812	808	800
Strategies					
Sending personal invitation letters d	0.39[a]	0.39[a]	0.36[a]	0.30[a]	0.17[a]
Personal visitations by community health workers d	0.15[a]	0.11[a]	0.15[a]	0.06[c]	0.07[b]
Number of individuals screened in clinical settings	0.28[a]	0.31[a]	0.35[a]	0.41[a]	0.28[a]
Free screening d	0.05	0.14[a]	0.08[b]	0.04	0.03
Characteristics of eligible persons					
Number of eligible persons	-0.32[a]	-0.32[a]	-0.33[a]	-0.43[a]	-0.31[a]
Ratio of males to females	-0.02	0.02	-0.07[b]	-	-
Percentage of those aged ≥65 years	0.05	0.17[a]	0.03	0.01	-0.06[c]
			Small municipalities (population <30,000)		
Number of municipalities	909	843	914	909	893
Strategies					
Sending personal invitation letters d	0.25[a]	0.30[a]	0.29[a]	0.24[a]	0.11[a]
Personal visitations by community health workers d	0.22[a]	0.13[a]	0.22[a]	0.20[a]	0.12[a]
Number of individuals screened in clinical settings	0.15[a]	0.09[a]	0.11[a]	0.20[a]	0.19[a]
Free screening d	0.02	0.13[a]	0.02	0.09[a]	0.00
Characteristics of eligible persons					
Number of eligible persons	-0.36[a]	-0.29[a]	-0.34[a]	-0.36[a]	-0.28[a]
Ratio of males to females	-0.01	0.02	-0.02	-	-
Percentage of those aged ≥65 years	0.01	0.16[a]	0.10[a]	-0.05	-0.15[a]

* [a]p values ≤0.01; [b]p values ≤0.05; [c]p values ≤0.1; [d]Dummy variables

for other types of cancer screening. The average number of individuals that had been screened in the clinical setting was the largest for colorectal cancer screening, and the smallest for gastric cancer screening.

The average cancer screening rates were higher in small municipalities than large municipalities for all cancer screening. Personal visitations by health workers were implemented in about 9% of all small municipalities, which was about 6% higher than that of large municipalities for all screenings. The average number of individuals screened in the clinical setting of large municipalities was more than 20-fold greater than that of small municipalities for all cancer screening. This may be because many small municipalities did not implement cancer screening in the clinical setting (i.e., about 80% for gastric, lung, and colorectal cancers, 27% for cervical cancers, and 43% for breast cancers).

The partial correlation coefficients for the relationships between cancer screening rates and the strategies used in Japan are presented in Table 2. In all municipalities, there were positive correlations between the screening rates for all cancers and the strategies used, with the exception of free screening ($p<0.01$). Of the four strategies, sending personal invitation letters had the highest correlation coefficients with cancer screening rates. They were as follows: 0.28 for gastric cancer screening, 0.32 for lung cancer screening, 0.30 for colorectal cancer screening, and 0.26 for cervical cancer screening. For cervical cancer screening, the correlation between the number of individuals screened in the clinical setting and the screening rates was similar to that of sending invitation letters. For breast cancer screening, all strategies had a low or no correlation with the screening rates in all municipalities.

In large municipalities, the correlation coefficients between sending invitation letters and the screening rates were relatively high. Specifically, they were as follows: 0.39 for gastric cancer screening, 0.39 for lung cancer screening, 0.36 for colorectal cancer screening, and 0.30 for cervical cancer screening. In large municipalities, the correlation coefficients between the number of individuals screened in the clinical setting and the screening rates were also relatively high, particularly for cervical cancer screening ($r=0.41$). For breast cancer screening, the correlation coefficient rose to 0.28 in large municipalities. In small municipalities, the correlation coefficients between cancer screening rates and the strategies used were similar to those of all municipalities, with the exception of the number of individuals screened in the clinical setting.

Discussion

In Japan, the National Cancer Control Plan was published in 2007 with the aim of increasing cancer screening rates above 50% within 5 years (MHLW, 2012a). To achieve this goal, municipalities had to implement effective strategies that would increase screening for various types of cancer. Previous studies have shown that sending personal invitation letters (Watanabe, 2003; Shimada et al., 2010a; Shimada et al., 2010b; Matsuda et al., 2011), distributing leaflets and pamphlets (Hisamichi et al., 1991; Yoshida et al., 2012), and increasing the availability of cancer screening in clinical settings (Takaku, 2011) were effective in improving cancer screening rates in Japan. However, it was unclear which strategy was the most effective. In the present study, after excluding the effects of confounding factors, correlations between four different strategies and cancer screening rates were evaluated.

Of all strategies, sending personal invitation letters had the highest positive correlations with screening rates for gastric, lung, colorectal, and cervical cancers. This strategy appears to be particularly effective in improving cancer screening rates in large municipalities. In most Western countries, the importance of a national call-recall system, which gives call and recall notifications by mail or telephone, is well recognized by the government for the purposes of increasing cancer screening (Quinn et al., 1998; Baron et al., 2008b). In Japan, municipalities are responsible for implementing strategies to improve cancer screening rates. However, nearly half of the municipalities did not implement this strategy. To improve cancer screening rates, all municipalities need to prioritize establishing a system that sends personal invitation letters for cancer screening.

The number of individuals that had been screened in the clinical setting also demonstrated positive correlations with all cancer screening in large municipalities. The correlation was particularly high for cervical cancer screening. Previous studies reported on the effectiveness of making access to screening easier by reducing the time or distance between the service delivery settings and the examinees in increasing colorectal and breast cancer screening in Western countries (Dolan et al., 1999; Baron et al., 2008c). Thus, increasing the availability of screening in the clinical setting should be effective in improving the screening rates for not only colorectal and breast cancer, but also for cervical cancer, in Japan. However, the quality assurance of cancer screening in the clinical settings was insufficient compared to that of mass screening in Japan (Arisue et al., 2007; Osaka City, 2010). Additionally, many small municipalities did not implement cancer screening in the clinical setting (i.e., about 80% for gastric, lung, and colorectal cancers, 27% for cervical cancers, and 43% for breast cancers). This may be due to several reasons. For example, small municipalities have been under more severe fiscal constraints than large municipalities (Ministry of Internal Affairs and Communications, 2011), and consequently are more likely not to have any incentives for increasing cancer screening in the clinical settings (Takaku, 2011). Thus, to increase cancer screening in the clinical setting, particularly in small municipalities, these problems need to be resolved.

Personal visitations by community health workers had low, but positive, correlations with cancer screening rates compared to sending invitation letters and the number of individuals that were screened in the clinical setting. However, these correlations were higher in small municipalities versus large municipalities. This may be because this strategy was implemented better in small municipalities than large municipalities. This strategy

is unlikely to be implemented in large municipalities because it is difficult to employ many community health workers for the number of eligible persons. It was previously reported that the cost effectiveness of one-on-one education per additional mammogram increased substantially if the cost of labor increased (Stockdale et al., 2000). Thus, each municipality needs to pay sufficient attention to fiscal constraints and decide whether to implement this strategy.

Free screening had a weak correlation with cancer screening rates. To improve screening rates, the MHLW had initiated a strategy that distributed free coupons to some individuals for breast and cervical cancer screening beginning in 2009 and for colorectal cancer screening beginning in 2011. The distribution of free coupons improved the screening rates for women who had not been screened for cervical cancer in Fukuoka Prefecture (Kuroki, 2012). However, changes in price for cancer screenings had little influence on demand for screenings in Hokkaido Prefecture (Takemura et al., 2001). The CDC recommends interventions that reduce out-of-pocket costs for breast cancer screening, but does not recommend such interventions for cervical and colorectal cancer screening due to insufficient evidence (Baron et al., 2008a; Sabatino et al., 2012; CPSTF, 2013). Therefore, reducing the out-of-pocket costs alone appears to be insufficient for improving the cancer screening rates.

However, it should be mentioned that none of the strategies had strong (or very high) correlations with cancer screening rates. For breast cancer screening, even sending personal invitations had a very low correlation with the screening rates. Therefore, just sending personal invitation letters and increasing the availability of screening in the clinical settings does not appear to greatly improve the cancer screening rates. The CDC recommends provider-oriented interventions, which evaluate the providers' performance and present the providers with the results, to increase cancer screening (Sabatino et al, 2008). The MHLW reported that implementing cancer screening and specific health checkups simultaneously improved cancer screening rates in some municipalities (MHLW, 2012b). In addition to these strategies, future studies that determine other effective strategies for improving cancer screening rates are warranted, including where screening occurs and the medical personnel involved (Tsunematsu et al., 2013).

This study has several limitations that need to be discussed. First, while using partial correlation analysis to determine the relationships between cancer screening rates and the strategies has provided some foundational knowledge on the topic, the causality of these relationships is still unclear. It is also necessary to consider that these findings may be a result of reverse causality, meaning that the implementation of strategies is influenced by cancer screening rates. Second, data on costs of the strategies implemented by each municipality could not be used in the analyses (Saywell et al., 1999; Stockdale et al., 2000; Saywell et al., 2003; Saywell et al., 2004). The cost-effectiveness of these strategies should be evaluated. Third, some municipalities might have conducted cancer screening not for all eligible persons but for very limited persons, such as those who sought to receive screening or had received a year before. Such municipalities should be excluded in the analyses. Fourth, it is necessary to further evaluate which strategies are more effective than those studied herein.

In conclusion, of the strategies used to improve cancer screening rates, sending personal invitation letters had the greatest positive correlations with screening rates for gastric, lung, colorectal, and cervical cancers. This strategy appears to be particularly effective in improving cancer screening rates in large municipalities. All municipalities should predominantly focus on establishing a system that sends personal invitation letters for cancer screening. In large municipalities, increasing the availability of screening in the clinical setting may also be effective in improving cancer screening rates.

Acknowledgements

This study was supported by National Cancer Center, Tokyo, Japan (Grant number: 23-A-41). The authors have no competing interests.

References

Arisue T, Yasuda I, Eizuka T, et al (2007). Current status and issues for individual cancer screening. *J Gastroenterol Cancer Screen*, **45**, 172-82.

Baron RC, Rimer BK, Coates RJ, et al (2008a). Methods for conducting systematic reviews of evidence on effectiveness and economic efficiency of interventions to increase screening for breast, cervical, colorectal cancers. *Am J Prev Med*, **35**, 26-33.

Baron RC, Rimer BK, Breslow RA, et al (2008b). Client-directed interventions to increase community demand to breast, cervical, colorectal cancer screening: a systematic review. *Am J Prev Med*, **35**, 34-55.

Baron RC, Rimer BK, Coates RJ, et al (2008c). Client-directed interventions to increase community access to breast, cervical, colorectal cancer screening: a systematic review. *Am J Prev Med*, **35**, 56-66.

Blumenthal DS, Smith SA, Majett CD, Alema-Mensah E (2010). A trial of 3 interventions to promote colorectal cancer screening in African Americans. *Cancer*, **116**, 922-9.

Champion V, Maraj M, Hui S, et al (2003) Comparison of tailored interventions to increase mammography screening in nonadherent older women. *Prev Med*, **36**, 150-8.

Community Preventive Services Task Force (2013). Cancer prevention and control: client-oriented interventions to increase breast, cervical, colorectal cancer screening. [http://www.thecommunityguide.org/cancer/screening/client-oriented/index.html]

Dolan N, McDermott M, Morrow M, Venta L, Martin G (1999). Impact of same-day screening mammography availability: results of a controlled clinical trial. *Arch Intern Med*, **159**, 393-8.

Frie KG, Ramadas K, Anju G, et al (2013). Determinants of participation in a breast cancer screening trial in Trivandrum District, India. *Asian Pac J Cancer Prev*, **14**, 7301-7.

Hamashima C, Shibuya D, Yamazaki H, et al (2008). The Japanese guidelines for gastric cancer screening. *Jpn J Clin Oncol*, **38**, 259-67.

Hamashima C, Aoki D, Miyagi E, et al (2010). The Japanese guideline for cervical cancer screening. *Jpn J Clin Oncol*, **40**, 485-504.

Hannon PA, Vu T, Ogdon S, et al (2012). Implementation and Process Evaluation of a Workplace Colorectal Cancer Screening Program in Eastern Washington. Health Promot Pract, first published on Jul 13.

Hisamichi S, Fukao A, Sugawara N, et al (1991). Evaluation of mass screening programme for stomach cancer in Japan. Cancer Screening, UICC, Cambridge University Press, Cambridge, 357-72.

King ES, Rimer BK, Seay J, Balshem A, Engstrom PF (1994). Promoting mammography use through progressive interventions: Is it effective? *Am J Public Health*, **84**, 104-6.

Kuroki H (2012). Survey on the trends in uterine cervical cancer screening in Japanese women: The efficacy of free coupons in the screening. *J Obstet Gynaecol Res*, **38**, 35-9.

Lee MH, Lee YY, Jung DW, et al (2012). Effectiveness of interventions to increase the participation rate of gastric cancer screening in the republic of Korea: a pilot study. *Asian Pac J Cancer Prev*, **13**, 861-6.

Lobb R, Opdyke KM, McDonnell CJ, et al (2011). Use of evidence-based strategies to promote mammography among medically underserved women. *Am J Prev Med*, **40**, 561-5.

Matsuda T, Monnma T, Oizumi H, et al (2011). Measures to improve the cancer screening rate in Yamagata prefecture. *J Gastroenterol Cancer Screen*, **49**, 252-9.

McAvoy BR, Raza R (1991). Can health education increase uptake of cervical smear testing among Asian women? *BM J*, **302**, 833-6.

Ministry of Health, Labour, Welfare (2010). Report on regional public health services and health promotion services.

Ministry of Health, Labour, Welfare (2012a). The National Cancer Control Plan.

Ministry of Health, Labour, Welfare (2012b). Assessment on progress of the medical care expenditure regulation plan (mid-year/ FY 2010).

Ministry of Internal Affairs and Communications (2011). Annual report on municipal finances FY 2009.

National cancer center (2013). Recommendation for cancer screenings.

OECD (2011). StatExtracts.

Osaka City (2010). Cancer screenings in the future in Osaka City.

Quinn M, Babb P, Jones J, Allen E (1998). Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. *BMJ*, **318**, 904-8.

Sabatino SA, Lawlence B, Elder R, et al (2012). Effectiveness interventions to increase screening for breast, cervical, colorectal cancers: nine updated systematic reviews for the guide to community preventive services. *Am J Prev Med*, **43**, 97-118.

Sabatino SA, Habarta N, Baron RC (2008) Interventions to increase recommendation and delivery of screening for breast, cervical, colorectal cancers by healthcare providers: systematic reviews of provider assessment and feedback and provider incentives. *Am J Prev Med*, **35**, 67-74.

Saywell RM Jr, Champion VL, Skinner CS, Menon U, Daggy J (2004). A cost-effectiveness comparison of three tailored interventions to increase mammography screening. *J Womens Health*, **13**, 909-18.

Saywell RM Jr, Champion VL, Zollinger TM, et al (2003). The cost effectiveness of 5 interventions to increase mammography adherence in a managed care population. *Am J Manag Care*, **9**, 33-44.

Saywell RM Jr, Champion VL, Skinner CS, et al (1999). Cost-effectiveness of five interventions to increase mammography screening. *Prev Med*, **29**, 374-82.

Shimada T, Kato K, Inomata Y, Kikuchi R, Shibuya D (2010a). Evaluation of recall letters and application methods in view of increasing gastric cancer screening rates. *J Gastroenterol Cancer Screen*, **48**, 647-54.

Shimada T, Kato K, Inomata Y, Kikuchi R, Shibuya D (2010b). Evaluation of recall letters and application methods in view of increasing colorectal cancer screening rates. *J Gastroenterol Cancer Screen*, **48**, 655-62.

Stockdale S, Keeler E, Duan N, Derose K, Fox S (2000). Cost and cost-effectiveness of a church-based intervention to promote mammography screening. *Health Serv Res*, **35**, 1037-57.

Suh M, Choi KS, Lee YY, Park B, Jun JK (2013). Cancer screening in Korea, 2012: results from the Korean National Cancer Screening Survey. *Asian Pac J Cancer Prev*, **14**, 6459-63.

Takaku R (2011). Do municipalities want to increase checkup rates of cancer screening tests?: an empirical analysis on the implementation of individual cancer checkup system. *Iryo To Shakai*, **21**, 121-36.

Takemura S, Sone T, Ohida T, et al (2001). Estimating price elasticities of demand for cancer screening programs: using the willingness to pay the user fee measured by the contingent valuation method. *Byoin Kanri*, **38**, 119-28.

Tsunematsu M, Kawasaki H, Masuoka Y, Kakehashi M (2013). Factors affecting breast cancer screening behavior in Japan - assessment using the health belief model and conjoint analysis. *Asian Pac J Cancer Prev*, **14**, 6041-8.

Townsend JS, Richardson LC, Steele CB, White DE (2009). Evidence-based interventions and screening recommendations for colorectal cancer in comprehensive cancer control plans: a content analysis. *Prev Chronic Dis*, **6**, 1-9.

Watanabe R (2003). An analysis of participation in cancer screening in Japan. *Iryo To Shakai*, **13**, 113-32.

Yoshida M, Kondo K, Nakanishi C, Tada T (2012). Interventional study for improvement of lung cancer screening rate. *J Med Invest*, **59**, 127-35.