Taylor & Francis Group

Acta Oncologica

Challenges in evaluation of screening for gastric cancer among men based on nonrandomized design

Ilkka Vohlonen, Matti Härkönen, Nea Malila, Eero Pukkala, Pentti Sipponen, Veli Koistinen & Matti Hakama

Published online: 23 Feb 2017.

Submit your article to this journal ⤤

Article views: 159

View related articles ⤤

View Crossmark data ⤤

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Challenges in evaluation of screening for gastric cancer among men based on nonrandomized design

Ilkka Vohlonen[a]*, Matti Härkönen[b], Nea Malila[c,d], Eero Pukkala[e,f], Pentti Sipponen[g], Veli Koistinen[h] and Matti Hakama[i]

[a]Department of Public Health, Health Policy, University of Eastern Finland, Kuopio, Finland; [b]Department of Clinical Chemistry, Clinical Chemistry, University of Helsinki, Helsinki, Finland; [c]Cancer Epidemiology, Finnish Cancer Registry, Helsinki, Finland; [d]School of Health Sciences, University of Tampere, Tampere, Finland; [e]Epidemiology, Finnish Cancer Registry, Helsinki, Finland; [f]School of Health Sciences, University of Tampere, Tampere, Finland; [g]Department of Pathology, Patolab Oy, Espoo, Finland; [h]Department of Biostatistics, Finnish Consulting Group, Helsinki, Finland; [i]Department of Epidemiology, Finnish Cancer Registry, Helsinki, Finland

ABSTRACT

Background: Objective was to quantify biases in screening for gastric cancer when comparing attenders to nonattenders using serum pepsinogen I (SPGI) level as primary test.

Methods: In mid 1990s, all men aged 51–65 years from two Finnish cities were invited to SPGI screening. Mortality and premature mortality in attenders were compared to nonattenders. Efficacy of screening was studied by 15 years' follow-up of standardized mortality ratio (SMR) and potential years of life lost (PYLL) due to gastric cancer. Bias due to selective attendance was quantified using corrective coefficients based on total cancer incidence and mortality, and gastric cancer-specific incidence and mortality for total population and nonattenders.

Results: In 1994–1996, men aged 51–65 years (16,872) were invited to SPGI assay and 12,175 men (72%) attended. SPGI was 25 microg/l or less in 610 (5%) men, indicating severe atrophic gastritis (AG). Post-screening gastroscopy was performed to 435 men with low SPGI. Of these, 168 men were referred for treatment due to abnormal focal lesions. Attributable proportions in reductions of SMR and PYLL from gastric cancer due to screening were 59% and 67%. After correcting for selective participation, attributable proportions were reduced to 23% and 39%.

Conclusions: Biomarker screening by low SPGI among middle-aged men followed by upper gastrointestinal endoscopy decreased long-term and premature mortality due to gastric cancer. However, in spite of methodological corrections done, the results do not justify any firm conclusions or recommend general screening programs. Randomized trials are warranted for this purpose.

Abbreviations: SPGI: serum pepsinogen I; AG: atrophic gastritis; SMR: standardized mortality ratio; SIR: standardized incidence ratio; PYLL: potential years of life lost

Introduction

In Finland, a population-based sample of middle-aged men from two cities was invited in 1994–1996, to investigate whether a serum pepsinogen I (SPGI) screening, followed by upper gastrointestinal (GI) endoscopy in those with low SPGI, can be applied as a public health program. It was known that SPGI levels of 25 microg/l or less indicate advanced (moderate or severe) atrophic gastritis (AG) in the gastric corpus and fundus with high reliability [1–4]. However, the indicators of structure, process, and outcome related to screening were completely unknown.

The objective of this study was to find out whether the recent laboratory evidence of the etiological role of SPGI was sufficient to start screening and to quantify challenges due to bias in the evaluative design. The screening was based on personal invitation of the target population but it did not contain any controls. We studied gastric cancer mortality by comparing the attenders in screening with the nonattenders over more than 15 years (from 1994 to 2011) after screening in 1994–1996. Since the design did not allow for a non-biased assessment of efficacy, statistical corrections were performed to adjust for the effects of selective participation.

Material and methods

Study population and study cohorts

All 16,872 men born in 1929–1943 and living in 1994–1996 in two Finnish cities (Kotka and Vantaa) were identified by population registry, and were invited to give a blood sample (serum) for the SPGI test: half of them in autumn 1994 and the remainder in autumn 1996 (Table 1). Altogether, 12,175

men (72%) attended the SPGI screening and 4697 did not. They form the attenders (screened) and nonattenders (not screened) in the present investigation, respectively (Figure 1).

Blood sampling and SPGI test

The men were invited by mail to visit municipal health centers for blood sampling drawn by experienced laboratory nurses. Fasting sera were collected in aprotinin (Trasylol, Bayer Germany, 200 KIE/ml) containing Venoject tubes and stored at −70 °C until analyzed. SPGI was analyzed using the specific ELISA tests provided by Biohit Oyj, Helsinki, Finland. The assay has been calibrated to correspond to results obtained by radioimmunoassay (RIA) used by Samloff in 1982 using 238 serum samples with serum pepsinogen concentrations between 1.5 and 120 microg/L. The clinical sensitivity and specificity for advanced AG are 92% and 90% for the test, respectively, according to the manufacturer's kit instructions for use. The screening was approved by the Ethics Committee of the University Hospital of Kuopio in 1994.

Low SPGI levels (25 microg/l or lower), indicating the presence of advanced (moderate or severe) AG in gastric corpus and fundus, were found in 610 men (5% of the screened men). Men with low SPGI levels were contacted by telephone and invited to gastroscopy unless they had medical contraindications.

Screening endoscopy in 1994–1996

A diagnostic upper GI endoscopy (screening gastroscopy) was performed in 435 SPGI-test positive men; i.e., 71% of the men with a low SPGI (Figure 1). All these endoscopies were performed by one experienced clinician. In endoscopy, biopsies from the stomach were collected: two from the antrum and two from the corpus, according to the guidelines of the Sydney System. Specimens for histopathology were also taken from all focal abnormalities (e.g., erosions, ulcers, nodules, polyps, masses, unusual color spots, etc.) seen by an endoscopist in the stomach mucosa. In all, 168 men (1.4% of all men in the screened cohort) were referred for further clinical treatment or surveillance. Of these, 56 men had focal gastric lesions with abnormal and atypical morphology in biopsy pathology, and all these were considered potentially cancerous or precancerous lesions (Figure 1).

Follow-up and collection of mortality data in 1994–2011

Long-term follow-up of all men invited for screening in 1994–1996 was run for cancer incidence, deaths, or emigration. These men formed the attenders (screened) and nonattenders (nonscreened) cohorts as explained above. The follow-up started the month after start of screening (on 1 December 1994 for the half of the men and on 1 November

Table 1. Number and percentage of men, number of accumulated person years, and mean time per person in follow-up in 1994–2011 by screening status.

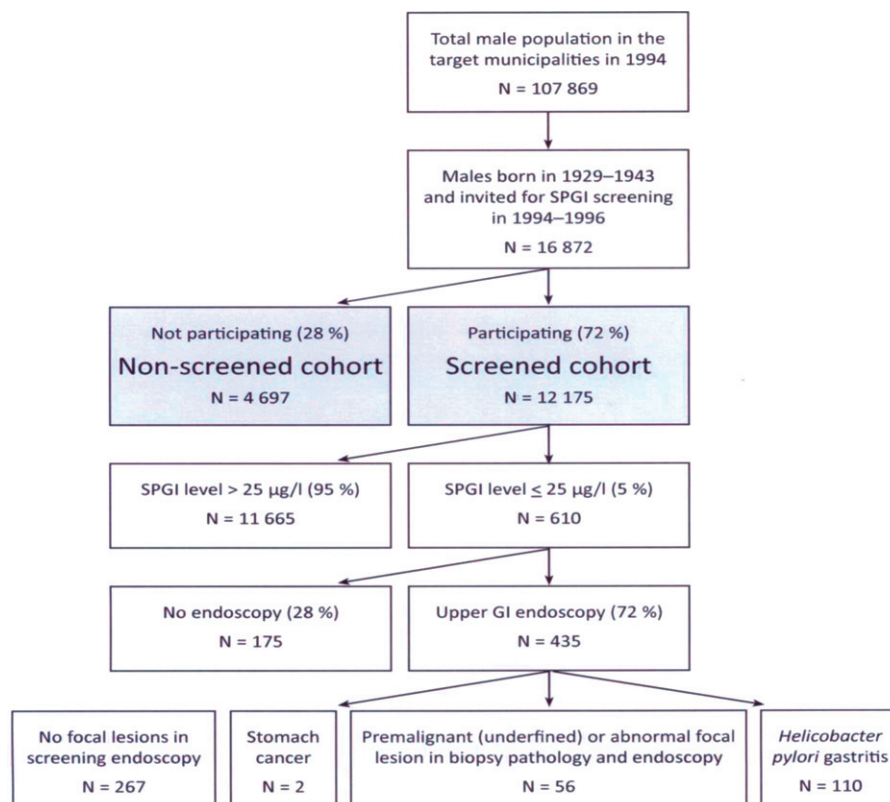| Screening status	Men		Person years	
	Number	Percent	Number	Mean/person
Non-screened	4697	27.8	54,955	11.7
Screened	12,175	72.2	166,798	13.7
Total invited	16,872	100.0	221,023	13.1

Figure 1. Design of the feasibility study on screening for gastric cancer in middle-aged men in two Finnish cities in 1994–1996 using a biomarker.

1996 for the other half) and ended at death, emigration or on 31 December 2011 (Table 1). Information on gastric cancer cases and deaths in the cohorts was received from the Finnish Cancer Registry (FCR) and data on emigration and date of death was from the population register. The cause of death information originates from Statistics Finland [5]. The individuals in the study were linked to registers using the national personal identity code of all Finnish residents.

Due to mandatory reporting of all cancer diagnoses in Finland, the FCR has a good national coverage of over 99% of solid cancer cases in Finland [6–8]. The registry includes data of all gastric cancer cases (topography code C16 in ICD-O-3) but it does not include data of cases classified as noninvasive premalignant lesions (dysplasia or intramucosal neoplasia). The cancer information used in this study included the date of diagnosis, topography, morphology, and data indicating whether the cancer patient died from the cancer in question or from any other cause.

Statistical methods

The standardized mortality ratios (SMRs) and potential years of life lost (PYLL) were calculated for gastric cancer and for all cancers combined (cancer at any site) [6,9] for attenders and nonattenders separately. In order to calculate the SMR, the observed numbers of death were compared with the expected numbers among the attenders and nonattenders. The expected numbers for each age group (5-year age categories) and 4-year calendar period were estimated by multiplying the number of person years in the category accumulated among the attenders and the nonattenders cohorts with the respective mortality rate in the total male population in Finland. The 95% confidence intervals for SMRs were estimated assuming a Poisson distribution for the numbers of observed deaths. Standardized Incidence Ratios (SIRs) were calculated according to the same principles as SMRs.

PYLL is an indicator of premature mortality. It represents the total number of life years lost (not lived) by an individual who was not assumed to die by a given age. The PYLL-value was calculated for each dead person as the difference between the observed date of death and the expected length of life. In the calculation of PYLL-values, the expected length of life was set at the Finnish average standard (80 years). Those men who died after their 80-year-old birthday did not contribute to the PYLL value in any way. PYLL-values were classified according to the cause of death (stomach cancer or all cancers) and reported per 1000 persons and per death among the screened and the non-screened.

Statistical analyses

Due to selective attendance, the effect of screening, i.e., efficacy, was measured by two types of attributable proportions in reductions of SMR and PYLL, the crude rates and the corrected rates. In analyses of the attributable proportions in reductions of SMR and PYLL, the differences in the geographic (e.g., supply and use of health services), social (e.g., demographic and etiological), and health (e.g., earlier diagnoses) characteristics of the men between the two cohorts were corrected for, since the potential self-selection correlates with the probability of death from gastric cancer. The regional differences (health care supply) and the general risks of any cancer were corrected for by taking into account the standardized incidence ratios and SMRs for all cancers of the total population and of the nonattenders, as described below.

In order to adjust for self-selection, we denote:
M = SMR or PYLL due to gastric cancer,
I = SIR of gastric cancer,
α = those screened,
αn = those not screened,
m = SMR or PYLL due to all cancers,
i = SIR of all cancers,
E1 = crude effect (attributable proportion),
E2 = effect with correction for selection and confounding factors (corrected attributable proportion).

The crude estimate of the effect (E1) was estimated as: $E1 = 1 - (M_{\alpha n}/M_n)$.

The corrected estimate of the effect (E2) was estimated as: $E2 = 1 - (I_{\alpha n}/I_\alpha)(M_{\alpha n}/M_n)(ma_{\alpha n}/m_\alpha)(i_{\alpha n}/i_n)$.

Selection bias was corrected by the relative risk of background incidence of gastric cancer in screened and non-screened cohorts ($I_{\alpha n}/I_\alpha$). Furthermore, those attending screening may have a better general health and a more favorable prognosis than those not attending. We assumed that these biases are seen in the survival and in the ratio of cancer mortality to cancer incidence (m/i) in the cohorts, justifying the correction factor ($m_{\alpha n} i_\alpha$)/($m_\alpha i_{\alpha n}$). The same formulae were used for correcting the differences in reductions of PYLL values due to stomach cancer between the attenders and nonattenders cohorts by adjusting for PYLL values for all cancers.

The mortality analyses of gastric and all cancers were first performed with refined SMRs, i.e., excluding deaths from gastric cancer diagnosed before the beginning of the follow-up. There were only four such deaths in the present series. The refined and the crude SMR rates of all gastric cancer deaths turned out to be similar independently of time of the diagnosis. Only the crude SMRs and PYLL values were presented including all gastric cancers without considering the timing of diagnosis.

Results

The SMRs and SIRs for all cancers combined (cancers at all sites) and for gastric cancer separately in both study cohorts are presented in Table 2. SMR for gastric cancer (0.53) was significantly below unity in the attender cohort and it was higher in the nonattenders cohort (SMR 1.28). The SMR for all cancers in the attender cohort was 0.91 and in the nonattenders cohort 1.45.

The SIRs for all cancers were roughly similar in both study cohorts. However, in the attender cohort, the SIR for stomach cancer was lower than expected (0.75, 95% CI 0.57–0.95) and

Table 2. Standardized mortality ratios (SMR) and standardized incidence ratios (SIR) among the screened and non-screened men by the site of cancer 1994–2011.

Cancer site	Event	Screened (N = 12,175)				Non-screened (n = 4697)			
		OBS	EXP	SMR or SIR	95%CI	OBS	EXP	SMR or SIR	95%CI
Stomach cancer	Deaths	31	58.1	0.53[a]	0.36–0.75	24	18.7	1.28	0.82–1.90
	Cases	60	80.5	0.75	0.57–0.95	23	26.1	0.88	0.56–1.32
All cancers	Deaths	1 020	1 116.7	0.91	0.86–0.97	514	354.8	1.45	1.33–1.57
	Cases	3 088	2 747.6	1.12	1.08–1.16	976	881.4	1.11	1.04–1.17

[a]$p < 0.001$.

Table 3. Potential years of life lost (PYLL values) per 1000 persons and mean PYLL values per death before age of 80 among the screened and non-screened men by site of cancer 1994–2011.

Cancer site	Screened (n = 12,175)			Non-screened (n = 4697)		
	PYLL	95%CI	PYLL/death	PYLL	95%CI	PYLL/death
Gastric cancer	25.7	15.5–36.0	10.4	77.3	43.1–111.4	15.1
All cancers	875	815–935	11.0	1388	1259–1518	13.2

Table 4. Crude and corrected proportions (%) in reduction of mortality and of potential years of life lost (PYLL) among men attributable to SPGI screening.

Outcome indicator	Attributable proportion (%)	
	Crude (E1)	Corrected (E2)
Mortality	59	23
PYLL	67	39

it was lower in the attenders than in the nonattenders cohort (0.88, 95% CI 0.56–1.32) (Table 2).

The PYLL value for gastric cancer was significantly lower among the attenders (27.7, 95% CI 15.5–36.0) than among the nonattenders (77.3 95% CI 43.1–111.4) cohort (Table 3). Again, the PYLL value for all cancers was lower in the attender cohort than in the nonattenders cohort but this difference was smaller than the difference in PYLL for gastric cancer.

The crude and corrected effects of screening on mortality and prolongation of life are given in Table 4. The relative decrease in SMR attributable to SPGI screening for gastric cancer resulted in a value 0.59 for the crude proportion (E1). The corresponding crude attributable proportion for PYLL was 0.67. After correcting for selection bias, the corrected attributable proportions (E2) for SMR and PYLL were 0.23 and 0.39, respectively (Table 4).

Discussion

Gastric cancer is one of the common cancers with high mortality worldwide. The age-adjusted incidence rate (world standard) in 2013 among Finnish men was 6 per 100 000 person years, which is one-tenth of the respective rate in the 1950s [10]. The life expectancy of patients with gastric cancer is not long, unless the cancer is diagnosed in its early stage. In 2009–2013, the five-year relative survival ratio in males with stomach cancer in Finland was 23% [10]. About a half of gastric cancers, being most often of the so-called intestinal type, are considered to develop in AG and an acid-free stomach via the 'Correa cascade' [11–14]. Therefore, early identification and endoscopy of subjects with AG may facilitate the

diagnosis of gastric cancer at an early stage, and may enable treatment of premalignant gastric lesions (intramucosal neoplasias) in an asymptomatic phase [6–9]. A low serum level of pepsinogen I (SPGI) is a reliable biomarker of AG and is, therefore, a tool to noninvasively delineate subjects with advanced AG who need a prompt diagnostic upper GI endoscopy because of increased cancer risk [1–4]. Correspondingly, it is conceivable that an early diagnosis of AG with a biomarker test followed with a diagnostic upper GI endoscopy will improve the cancer survival, resulting also in decrease of premature mortality.

In order to develop a program of screening for stomach cancer by a biomarker, this study comprised a large population-based sample of men born in 1929–1943 who were invited to screening by SPGI/endoscopy in 1994–1996 in two Finnish cities. This screening program applying a simple SPGI biomarker blood test was well accepted with a high participation percentage (72%) and the screening process was adequate. Here we evaluate whether the design was adequate for evaluation of the efficacy of the screening program by gastric cancer mortality.

Biases between the attenders and nonattenders cohorts constitute a dilemma in assessment of the efficacy of the screening. Because the screening trial could not be implemented to allow comparison of invited and non-invited groups of men, the only option was to try to correct for the effects of selective attendance afterwards using general indicators of cancer frequency and mortality in the categories to be compared.

In general, the reports of screening programs on cancers do not include data or attempts to correct for selection biases even though they are certainly noteworthy in all screening projects [8,15]. The men who did not want to participate in the SPGI test may have been less interested in health issues in general than those who participated and were, therefore, more liable to early death from gastric cancer and to high mortality from any cause of death [16–20]. The men in the present nonattender cohort had a clearly higher incidence (SIR) of gastric cancer and of all cancers than those in the attender cohort. In spite of the corrections afterwards, there may still remain biases that could not be taken into account.

The present study suggests that biomarker screening by SPGI may reduce mortality from gastric cancer by one-fifth over 15 years. Correspondingly, screening was estimated to reduce the potential years of life-lost (PYLL) due to gastric cancer by almost 40%. These estimates were, however, inconclusive due to methodological issues as described above.

In our study, two cases of gastric cancer were found by screening in men with low SPGI in 1994-1996, and both patients died from gastric cancer within 5 years. Therefore, the long term decrease in mortality between 1994 and 2011 was likely due to treatment of pre-cancerous lesions observed in screening endoscopy or were due to eradication of the on-going *Helicobacter pylori* infection [16–20]. It can be assumed that the precancerous lesions in an atrophic stomach mucosa develop to invasive cancers during life-time in up to one third of cases if the lesions are not properly treated [20].

In our study, the expected number of deaths due to gastric cancer was 58 and this number was based on a forecast made on the basis of the Finnish Cancer Registry. In the screening for gastric cancer by a biomarker, we found 56 lesions and 2 cases of gastric cancer. In this view, expected and observed cases of gastric cancer were similar.

This observation indicates, that at least some 20 men would be at risk of gastric cancer among the 56 men who were diagnosed with a precancerous lesion. Based on the observed gastric cancer deaths, the crude attributable proportion in reduction of deaths due to gastric cancer was 59% and therefore 16 deaths due to gastric cancer were prevented if attendance was not selective. Because of selection, we corrected it in the analyzes and the corrected attributable proportion was 23% with the observed reduction of 6 deaths due to gastric cancer. We cannot regard this figure as an unbiased one, because the unknown potential of residual confounding after the methodological corrections. Hence we do not know whether the screening for gastric cancer was efficacious or not.

Analyzes of the effects of SPGI screening on premature gastric cancer mortality using PYLL gave results consistent with those obtained by cancer mortality. The corrected attributable proportion of SPGI screening in reduction of PYLL for gastric cancer was 39%. PYLL is dependent on the value for life expectancy selected for the calculations. In the present investigation, a life expectancy of 80 years was selected. Deaths before this age are considered premature and most deaths due to gastric cancer in Finland occur before age of 80.

The greater reduction in corrected attributable proportion in reduction of PYLL than of SMR for gastric cancer (39% vs. 23%) may indicate that some cancerous lesions among the attending were observed and treated in early and curable stages.

In spite of the association between SPGI screening and decreased gastric cancer mortality and PYLL, the present results may still be biased although methodological corrections were done to correct for confounding. Therefore, the present results do not justify definitive conclusions or implementation of general screening programs. Information on efficacy of a new and potential screening technology for cancer needs to be based on well-designed study protocols which result in unbiased evidence. Therefore, controlled and randomized screening studies are an ultimate prerequisite. In the developed countries, incidence of gastric cancer is decreasing along with *H. pylori* infection and AG, but in the developing countries these are still major public health problems.

Acknowledgments

Disclosure statement

References

[1]   Storskrubb T, Aro P, Ronkainen J, et al. Serum biomarkers provide an accurate method for diagnosis of atrophic gastritis in a general population: the Kalixanda study. Scand J Gastroenterol. 2008;43:1448–1455.

[2]   Miki K, Morita M, Sasajima M, et al. Usefulness of gastric cancer screening using the serum pepsinogen test method. Am J Gastroenterol. 2003;98:735–739.

[3]   Dinis-Ribeiro M, Yamaki G, Miki K, et al. Meta-analysis on the validity of pepsinogen test for gastric carcinoma, dysplasia or chronic atrophic gastritis screening. J Med Screen. 2004;11: 141–147.

[4]   Sackett DL, Holland WW. Controversy in the detection of disease. Lancet. 1975;2:357–359.

[5]   Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. Acta Oncol. 1994;33:365–369.

[6]   Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.

[7]   Cambell DT, Stanley JC. Experimental and quasi-experimental designs for research. Chicago (IL): Rand McNally; 1966.

[8]   Cochrane AL, Holland WW. Validation of screening procedures. Br Med Bull. 1971;27:3.

[9]   Holland WW, Steward S. Screening in health care. London: The Nuffield Provincial Hospitals Trust; 1990.

[10]   Engholm G, Ferlay J, Christensen N, et al. Gastric precancerous process in a high risk population: cohort follow-up. Cancer Res. 1990;50:4737–4740.

[11]   Correa P, Piazuelo MB. The gastric precancerous cascade. J Dig Dis. 2012;13:2–9.

[12]   Haenszel W, Correa P, Cuello C, et al. Gastric cancer in Colombia. II. Case-control epidemiologic study of precursor lesions. J Natl Cancer Inst. 1976;57:1021–1026.

[13]   Sipponen P, Kekki M, Haapakoski J, et al. Gastric cancer risk in chronic atrophic gastritis: statistical calculations of cross-sectional data. Int J Cancer. 1985;35:173–177.

[14]   Varis K, Sipponen P, Laxen F, et al. The Helsinki Gastritis Study Group. Implications of serum pepsinogen I in early endoscopic diagnosis of gastric cancer and dysplasia. Scand J Gastroenterol. 2000;35:950–956.

[15]   Jang JS, Choi SR, Qureshi W, et al. Long-term outcomes of endoscopic submucosal dissection in gastric neoplastic lesions at a single institution in South Korea. Scand J Gastroenterol. 2009;44:1315–1322.

[16]   De Vries AC, Kuipers EJ. Epidemiology of premalignant gastric lesions: implications for the development of screening and surveillance strategies. Helicobacter. 2007;12:22–31.

[17] Crumley AB, Going JJ, McEwan K, et al. Endoscopic mucosal resection for gastroesophageal cancer in a U.K. population. Long-term follow-up of a consecutive series. Surg Endosc. 2011;25:543–548.

[18] Ichinose M, Yahagi N, Oka M, et al. Screening for gastric cancer in Japan. In: Wu GY, Aziz K, editors. Cancer screening. A practical guide for physicians. Potowa (NJ): Humana Press; 2001. p. 87–102.

[19] Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p. 8–146.

[20] Wilson JMG, Junger G. Principles and practice of screening for disease. Public Health Paper Number 34. Geneva: WHO; 1968.