Sociodemographic disparities in gastric cancer and the gastric precancerous cascade: A population-based study

Dianqin Sun,[a,1] Lin Lei,[b,1] Changfa Xia,[a] He Li,[a] Maomao Cao,[a] Siyi He,[a] Zhiyi Zhang,[c] Guizhou Guo,[d] Guohui Song,[e] Ji Peng,[c] and Wanqing Chen[a]*

[a]Office of Cancer Screening, National Cancer Center / National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No.17 Pan-jia-yuan South Lane, Chaoyang District, Beijing 100021, China
[b]Department of Cancer Prevention and Control, Shenzhen Center for Chronic Disease Control, Shenzhen 518020, China
[c]Department of Gastroenterology, Gansu Wuwei Tumor Hospital, Wuwei, China
[d]Linzhou Cancer Hospital, Anyang 456500, China
[e]Cixian Cancer Institute, Handan 056500, China

The Lancet Regional Health - Western Pacific 2022;23: 100437
Published online xxx
https://doi.org/10.1016/j.lanwpc.2022.100437

Summary

Background Gastric carcinogenesis is a multistep process initiating with chronic gastritis and progressing through atrophy, intestinal metaplasia, and dysplasia to carcinoma. This study aims to comprehensively investigate sociodemographic disparities in each stage of gastric carcinogenesis and estimate to what extent the inequalities could be ascribed to risk factors of gastric cancer (GC).

Methods We used the baseline data from a community-based study in China's high-risk areas, totalling 27094 participants. Gastric mucosa status was ascertained by endoscopy and biopsies. An overall socioeconomic status (SES) variable was generated by latent class analysis. We calculated relative risks (RRs) and 95% confidence intervals (CIs) using modified Poisson regression to assess associations of sociodemographic factors with each cascade stage. We estimated the percentage of the excess risk for neoplastic lesions among vulnerable populations that can be explained by established risk factors.

Results Age and sex showed associations with all gastric lesions, whose RRs increased with lesion progressing. Compared with individuals without schooling, the RRs of neoplastic lesions for people with primary, secondary, and post-secondary education were 0·86 (95% CI 0·76–0·97), 1·00 (95% CI 0·88–1·13), and 0·70 (95% CI 0·47–1·03), respectively. Participants with medium SES had a lower risk of neoplastic lesions than people in the low SES group (RR 0·83, 95% CI 0·74–0·93). GC risk factors could explain 33·6% of the excess risk of neoplastic lesions among men and a small proportion of the disparities among SES groups.

Interpretation Age and sex were essential sociodemographic factors for GC and precursor diseases. Individuals with low educational levels or SES were more likely to have neoplastic lesions. About one-third of the sex difference and a slight fraction of the socioeconomic inequalities could be attributed to included risk factors.

Funding Sanming Project of Medicine in Shenzhen, National Natural Science Foundation, and Special Project of Bejing-Tianjin-Hebei Basic Research Cooperation.

Introduction

Gastric cancer (GC) is the fifth most frequent cancer and the fourth leading cause of cancer death globally, responsible for more than one million new cases and almost 768 thousand deaths in 2020.[1] Although recent decades have seen the decrease of GC incidence and mortality,[2] sociodemographic inequalities in GC have long been observed on a global scale and seen as a typical case of social disparities in cancer continuum. Men, the old, and people with low educational level or socioeconomic status (SES) tend to have a higher GC risk.[3–5] Addressing social gaps in cancer health is a global priority, which may boost cancer control progress and achieve health equity. However, few data can explain the persistent sociodemographic gaps in GC.

*Corresponding author.
E-mail address: chenwq@cicams.ac.cn (W. Chen).
[1] Dianqin Sun and Lin Lei contributed equally.

Research in context

Evidence before this study

We searched PubMed for studies published before January 1, 2022, with the search terms "disparities" or "inequalities", and ("gastric" or "stomach"), and ("cancer" or "precancerous" or "neoplasm" or "preneoplastic"), and ("sociodemographic" or "social" or "sex" or "education" or "socioeconomic"). We found the well-establishment and long existence of sociodemographic disparities in gastric cancer globally. Gastric carcinogenesis is a multistep process. Several studies estimated the sociodemographic inequalities in gastric cancer, but few gave the disparity pattern across the gastric precancerous cascade. Further, no study so far has quantified known gastric cancer risk factors' contributions to the sociodemographic differences.

Added value of this study

To the best of our knowledge, this is the first comprehensive analysis of the sociodemographic disparities in each stage of the gastric precancerous cascade. This study also evaluates to what extent the significant sociodemographic inequalities in the prevalence of neoplastic lesions, including gastric dysplasia and carcinoma, can be explained by differences in the main established gastric cancer risk factors. We found that age and sex showed significant associations with all gastric lesions, whose ORs increased with lesion progressing. Individuals with low educational levels or socioeconomic status were more likely to have neoplastic lesions. Main gastric cancer risk factors, including diet, alcohol drinking, smoking, obesity, and occupational exposure, could explain 33·6% of the excess risk of neoplastic lesions among men and a small proportion of the disparities among socioeconomically vulnerable populations.

Implications of all the available evidence

Recent studies have shown a decreased trend of gastric cancer incidence and mortality, but sociodemographic inequalities in gastric cancer have long been observed. Targeting main gastric cancer risk factors, including diet, alcohol drinking, smoking, obesity, and occupational exposure, may mitigate about one-third of the excess risk of neoplastic lesions in men but have little influence on the socioeconomic disparities. Future studies are needed to reveal further the mechanisms underlying the sociodemographic differences in gastric cancer.

According to the widely accepted multistep model of gastric carcinogenesis,[6] the intestinal type of gastric cancer is generally preceded by a sequence of precancerous lesions. The Correa's cascade indicated that non-atrophic gastritis could advance to atrophic gastritis, and over the years, progress through intestinal metaplasia, dysplasia to eventually GC. Comprehensive analysis of risk factors for GC and gastric precancerous cascade could potentially lead to a better understanding of the gastric carcinogenesis mechanisms. Even though associations between sociodemographic factors and GC have been well-established, few analyses give a complete picture of sociodemographic inequalities in GC and different premalignant gastric lesions. Furthermore, evidence suggests that men and individuals with low SES are more likely to have lower consumption of vegetables and fruits and high smoking rates,[7,8] which are considered risk factors for GC.[9] Thus, it is plausible that observed sociodemographic disparities are driven by main GC risk factors. Nevertheless, few published studies have quantified the contributions of GC risk factors to the disparities.

Therefore, in this study, we aimed to examine the associations of sociodemographic factors with GC and its precursor diseases using the cross-sectional data from China's high-risk areas. Besides, we assessed to what extent the major sociodemographic inequalities in the prevalence of neoplastic lesions, including gastric dysplasia and carcinoma, can be explained by differences in the main established GC risk factors.

Methods

Study population

We used the baseline data from a community-based prospective study in China, the protocol of which has been published and registered (ChiCTR-EOR-16008577).[10] The baseline survey was conducted between May 2015 and July 2017, targeting all residents aged 40-69 in 81 villages of China's high-risk areas (Cixian, Linzhou, and Liangzhou district in Wuwei). China's high-risk areas refer to a district with a GC incidence of more than 30/100,000.[11] And we summarized the GC incidence in the three study centres in Table S1 as the background information. We successfully recruited 27957 participants out of the 63969 residents. After the face-to-face interview, 846 individuals were excluded due to age range restriction, cancer history, endoscopy history in the past three years, and erroneous survey data. Of the remaining 27111 participants who received the endoscopic examination, 17 were excluded for missing information. Finally, 27094 participants were included in our analysis. The flow chart of participants' inclusion in this study is presented in Figure 1. The study was approved by the ethics committee of the National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (2015SQ00223). All participants provided written informed consents.

Assessment of sociodemographic factors

We collected information regarding educational level, marital status, number of family members, and annual

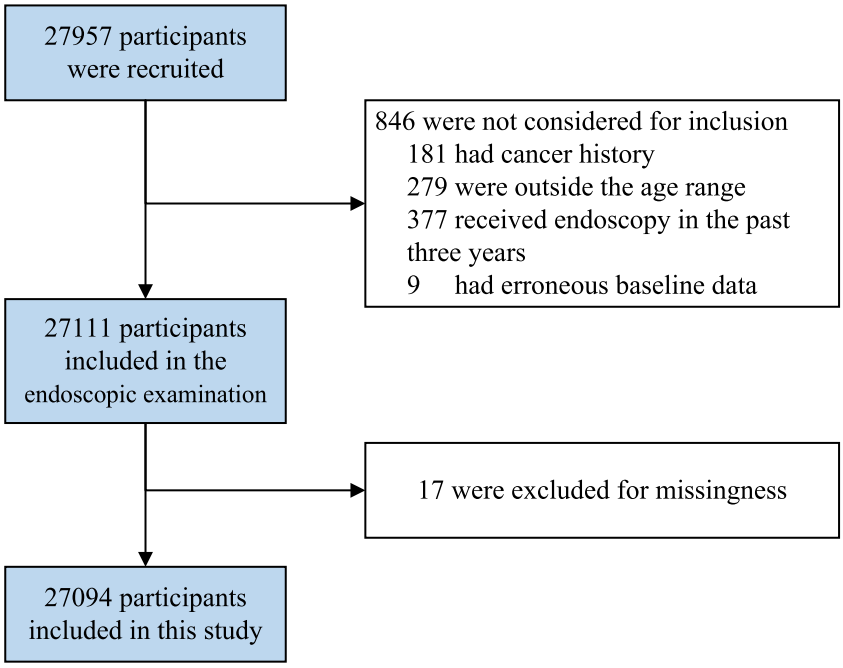

Figure 1. Flow diagram of participants excluded for various reasons and final inclusion.

family income by face-to-face interview. Family income level was measured by the ratio of annual family income per person to the national poverty level (2300 Yuan). It then was categorized into four groups: poor (≤1), low (1-2), medium (2-4), and high (>4). Similar to previous literature,[12] we used the latent class analysis (LCA) to construct an overall SES based on these three categorical variables: educational attainment (no formal schooling, primary school, secondary school, post-secondary school), marital status (married, single, divorced, widowed), and family income level (poor, low, medium, high). LCA can identify a set of mutually exclusive latent classes based on participants' responses to observed categorical indicators. Akaike information criterion (AIC), Bayesian information criterion (BIC), and likelihood ratio statistic G^2 were used to determine the number of classes. Item-response probability was used for defining latent classes. Table S2 shows the AIC, BIC, and G^2 in models with different latent classes. Table S3 presents item-response probabilities. Finally, we identified three latent classes, indicating low, medium, and high SES.

Assessment of other GC risk factors

According to the latest World Cancer Research Fund (WCRF) report[13] and the International Agency for Research on Cancer (IARC) Monograph series,[14] main GC risk factors include occupational exposure to rubber or radiation, obesity, and lifestyle factors (smoking, alcohol drinking, fruit consumption, vegetable consumption, and consumption of foods preserved by salting).

Interviewees were asked whether they were exposed to rubbery, radiation, or radioactive material in their working environment. Participants' height and weight were measured, and body mass index (BMI) was calculated as weight (kg)/height $(m)^2$. Based on the self-reported smoking amount, type of tobacco smoking, and time length, we calculated pack-years of cigarette smoking assuming that a factory cigarette is equivalent to one gram of pipes and hand-rolled cigarettes. Four categories were defined for smoking: no smoking, less than 20 pack-years of smoking, more than 20 and less than 40 pack-years, and more than 40 pack-years. For alcohol intake, we first defined infrequent and regular drinkers based on whether participants drank for more than one year. Regular drinkers were further subclassified as a low, medium, and high consumption according to sex: for men, the groupings were less than 20g, 20−40g, and 40g or more per day, and for women, less than 10g, 10−20g, and 20g or more per day. The alcohol intake was quantified as grams of pure alcohol per day, based on the beverage type and the amount drunk per day, assuming that the percentage of alcohol by volume was 15% for rice wine, 52% for spirits. We assumed that 750 mL of beer or 200 mL of grape wine is equalized to 50g of spirits. Frequencies of fruit consumption, vegetable consumption, and consumption of foods preserved by salting were categorized into three levels: every day, more than one time per week, and less than four times per month.

Outcome assessment

All participants included in this study received standard endoscopy from experienced physicians. Indigo carmine dye was applied to the stomach when necessary to aid the diagnosis of suspicious lesions. Biopsies of suspicious lesions were performed for further pathological diagnosis. Two experienced pathologists independently reviewed the biopsy slides, and consultation resolved discrepancies. Finally, endoscopic findings were reported as normal gastric mucosa, non-atrophic gastritis, atrophy, intestinal metaplasia, dysplasia (including indefinite for dysplasia), cancer, or others. According to whether the lesion is negative for dysplasia, the five types of lesions were further classified as (1) non-neoplastic lesions (non-atrophic gastritis, atrophy, and intestinal metaplasia) and (2) neoplastic lesions (dysplasia and carcinoma).

Statistical analyses

When the outcome is common (prevalence rate > 7%), odds ratio from logistic regression is no longer a good approximation to relative risk (RR).[15] Thus, we used the modified Poisson regression to measure associations of sociodemographic factors with each category of gastric lesions.[16] We applied Begg and Gray's method[17] to approximate the multinomial regressions by performing multiple separate modified Poisson regression models with normal gastric mucosa as a reference group. Adjusted RRs and 95% confidence intervals (CIs) were reported. Some sociodemographic factors might mediate associations of upstream factors with the outcome. For example, the educational level may influence the risk by affecting family income level. Thus, we applied a sequential method to include variables in regression models. Covariates included in each model were summarized in supplementary Table S4. Explicitly speaking, we estimated the effects of age and sex in the model, including age, sex, and study centre. Educational level was then added to the model to assess its RR. Marital status and family income level were additionally added. The effect of SES was measured in the model adjusted for age, sex, and study centre. To shed light on the relationship between sociodemographic factors and non-neoplastic and neoplastic lesions, we performed two separate modified Poisson regression models with normal gastric mucosa as the reference, which included the same covariates as the above models used. To assess the effect heterogeneity in non-neoplastic and neoplastic lesions, we conducted another model among cases with non-neoplastic or neoplastic lesions and calculated the P-value for heterogeneity.

To explore how much of the excess risk of neoplastic lesions was related to the established GC risk factors among men, people with low educational level, or low SES, we calculated the percentage excess risk explained (ERE%). GC risk factors adjusted in the model included smoking (nominal), alcohol drinking (nominal), fruit consumption (nominal), vegetable consumption (nominal), consumption of foods preserved by salting (nominal), occupational exposure to rubber or radiation (dichotomous), BMI (continuous). Estimated from RR, ERE% was calculated as (unadjusted RR − adjusted RR)/(unadjusted RR − 1) if RR ≥1.[18] Since we chose men, no formal schooling, and low SES as the reference group to keep consistent with above analyses, RRs of sex, education, SES were below one. Thus, we used the inverted RRs in the formula to calculate ERE%.[18] All analyses were performed using R version 4.0.1 (R Foundation for Statistical Computing, Vienna, Austria). Two-sided *P* values <0.05 were considered to be significant.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Among the 27094 participants included in the analysis, we identified 7741 cases of non-atrophic gastritis, 5129 cases of atrophic gastritis, 424 cases of intestinal metaplasia, 1938 cases of dysplasia and 143 cases of GC. A total of 307 cases were reported as others. As shown in Table 1, participants had a mean age of 53·24 ± 7·66 years, BMI of 24·73 ±3·95 kg/m^2; 57·01% were females, 94·03% were married, and 11·42% were classified as low SES.

Figure 2 shows the associations of sociodemographic factors with GC risk and different precursor lesions. Only age and sex showed significant associations with all gastric lesions. The RRs of age increase with the growing lesion severity, ranging from 1·01 to 1·13. Similar associations were found for sex, with the RRs comparing women and men ranging from 0·95 to 0·29. However, the consistent effect with a graded fashion was not observed for other sociodemographic factors. For educational level, we only found a statistically significant association with dysplasia (RR 0·86, 95% CI 0·76 −0·98, primary school *vs.* without formal schooling). Even though low-income people had a lower risk of non-atrophic gastritis (RR 0·84, 95% CI 0·76−0·94) than people in poverty, we did not observe statistically significant relationships between family income level and more severe lesions. In contrast, there were statistically significant associations of SES with intestinal metaplasia and dysplasia but not for less advanced lesions. We did not find any significant educational level, income level, and SES associations with the risk of GC. But the effect of educational level on GC in our analysis also showed a dose-response manner, with RRs decreasing from 0·74 to 0·36 (Figure 2).

	Total (*n*=27094)
Age, mean (SD)	53·24 (7·66)
Female, *n* (%)	15446 (57·01%)
Education, *n* (%)	
Not receiving formal schooling	4547 (16·78%)
Primary education	9295 (34·31%)
Secondary education	12920 (47·69%)
Post-secondary education	332 (1·22%)
Marital status, *n* (%)	
Married	25477 (94·03%)
Single	169 (0·62%)
Divorced	120 (0·44%)
Widowed	1328 (4·90%)
Family income level, *n* (%)	
Poor	750 (2·77%)
Low	3347 (12·35%)
Medium	12201 (45·03%)
High	10796 (39·85%)
Socioeconomic status, *n* (%)	
Low	3094 (11·42%)
Medium	14806 (54·65%)
High	9194 (33·93%)
Smoking, *n* (%)	
Never	20083 (74·12%)
Pack-years≤20	4020 (14·84%)
20< Pack-years ≤40	2260 (8·34%)
Pack-years >40	731 (2·70%)
Alcohol drinking, *n* (%)	
Infrequent	23903 (88·22%)
Low	326 (1·20%)
Medium	1117 (4·13%)
High	1748 (6·45%)
Fruit consumption, *n* (%)	
Less than four times per month	473 (1·75%)
At least one time per week	6115 (22·57%)
Every day	20506 (75·68%)
Vegetable consumption, *n* (%)	
Less than four times per month	14604 (53·90%)
At least one time per week	10786 (39·81%)
Every day	1704 (6·29%)
Foods preserved by salting, *n* (%)	
Less than four times per month	25463 (93·98%)
At least one time per week	1499 (5·53%)
Every day	132 (0·49%)
Occupational exposure to rubber or radiation, *n* (%)	36 (0·13%)
BMI, mean (SD)	24·73 (3·95)

***Table 1*: Characteristics of the study population.**
SD: standard deviation; BMI: body mass index.

We then examined the associations based on whether lesions are non-neoplastic or neoplastic (Table 2). Age was associated with a higher risk of non-neoplastic and neoplastic lesions. A stronger association was observed for neoplastic than non-neoplastic lesions (RR 1·01 for neo-neoplastic lesions vs. 1·05 for neoplastic lesions; *P* for heterogeneity < 0.001). A more considerable sex difference was also observed for the risk of neoplastic lesions than that of non-neoplastic lesions (RR 0·94 for neo-neoplastic lesions vs. 0·52 for neoplastic lesions; *P* for heterogeneity < 0.001). With RRs of 0·86 (95% CI 0·76−0·97) and 0·70 (95% CI 0·47−1·03), respectively, people with primary and post-secondary education had a lower risk of neoplastic lesions than people without formal schooling. Compared with people with low SES, people with medium SES had a lower chance of neoplastic lesions (RR 0·83, 95% CI 0·74 −0·93).

Figure 3 presents changes in RR with covariate adjustment. In models adjusted for age and study centre only, women had half of the neoplastic lesion risk than men (RR 0·52). The addition of sociodemographic variables to the model did not alter the RR. Full covariate adjustment reduced the RR to 0·62, which suggested 33·6% of the excess risk in men was explained by these risk factors. Adjusting for all variables yielded a slight reduction in the excess risk of neoplastic lesions among individuals with lower educational levels and lower SES.

Discussion

Overall, we found statistically significant associations of age and sex with all stages of Correa's cascade. Moreover, the effects of age and sex increase monotonically with each cascade step. People with higher educational levels and SES had a lower risk of neoplastic lesions than those without formal schooling and being in a low SES. GC risk factors included in our analysis could explain 33·6% excess risk of neoplastic lesions in men. Only a tiny fraction of socioeconomic disparities in the neoplastic lesion risk can be attributed to risk factors included in our study.

Previous longitudinal studies have shown that the elderly[19,20] and men[20,21] had a higher probability of progression to a more advanced diagnosis among people with gastric lesions. Our cross-sectional analysis observed the increasing effect of age and sex with the cascade stage advancing, demonstrating the crucial roles of age and sex in the initiation and progression of precancerous gastric lesions indirectly. It suggested the rationale for developing risk-stratified GC surveillance strategies among populations with precursor lesions. Compared with the younger women, more frequent surveillance may be more appropriate for the older men. However, some studies[19,22] did not find that sex was a significant predictor of lesion development. Thus, future studies are warranted to resolve the conflict.

Education was the most extensively used SES measure in GC disparities research.[4] Consistent with previous studies,[4,23] our analysis indicated that educational level had the most robust association with neoplastic lesions among included SES indicators (Table 2). And similar to the previous meta-analysis,[4] we did not

	RR (95% CI)		P_{het}
	Non-neoplastic lesions (*n*=13294)	Neoplastic lesions (*n*=2081)	
Age, per year	1·009 (1·007, 1·010)	1·049 (1·044, 1·053)	<0·001
Female *vs.* Male	0·94 (0·92, 0·96)	0·52 (0·48, 0·57)	<0·001
Education			
Not receiving formal schooling	1 (ref)	1 (ref)	
Primary education	1·00 (0·97, 1·04)	0·86 (0·76, 0·97)	0·08
Secondary education	1·03 (0·99, 1·07)	1·00 (0·88, 1·13)	0·47
Post-secondary education	0·99 (0·89, 1·09)	0·70 (0·47, 1·03)	0·09
Marital status			
Married	1 (ref)	1 (ref)	
Single	0·93 (0·80, 1·08)	0·65 (0·38, 1·11)	0·47
Divorced	1·14 (0·98, 1·32)	1·07 (0·63, 1·80)	0·60
Widowed	0·97 (0·92, 1·02)	1·02 (0·87, 1·21)	0·45
Family income level			
Poor	1 (ref)	1 (ref)	
Low	0·88 (0·82, 0·95)	0·89 (0·71, 1·12)	0·66
Medium	0·96 (0·90, 1·04)	0·91 (0·74, 1·13)	0·99
High	0·99 (0·93, 1·07)	1·00 (0·81, 1·24)	0·60
Socioeconomic status			
Low	1 (ref)	1 (ref)	
Medium	0·98 (0·95, 1·02)	0·83 (0·74, 0·93)	0·005
High	1·03 (0·99, 1·07)	0·90 (0·79, 1·02)	0·04

***Table 2*: Relative risks of sociodemographic factors and 95% CIs for non-neoplastic and neoplastic lesions.**
RR: relative risk; CI: confidence interval; ref: reference.

observe significant associations between income levels and GC or neoplastic lesions. Since the overall SES indicator generated by LCA is not easy to administer, future studies about GC aetiology should include the educational level at least as the capture of SES. Though social disparities in GC have been well-established,[4,24] few longitudinal studies reported that socioeconomic variables were associated with the progression of gastric lesions,[25,26] which may be due to the limited statistical power. Our analysis also failed to find any signs indicating significant roles of SES in the progression of gastric lesions.

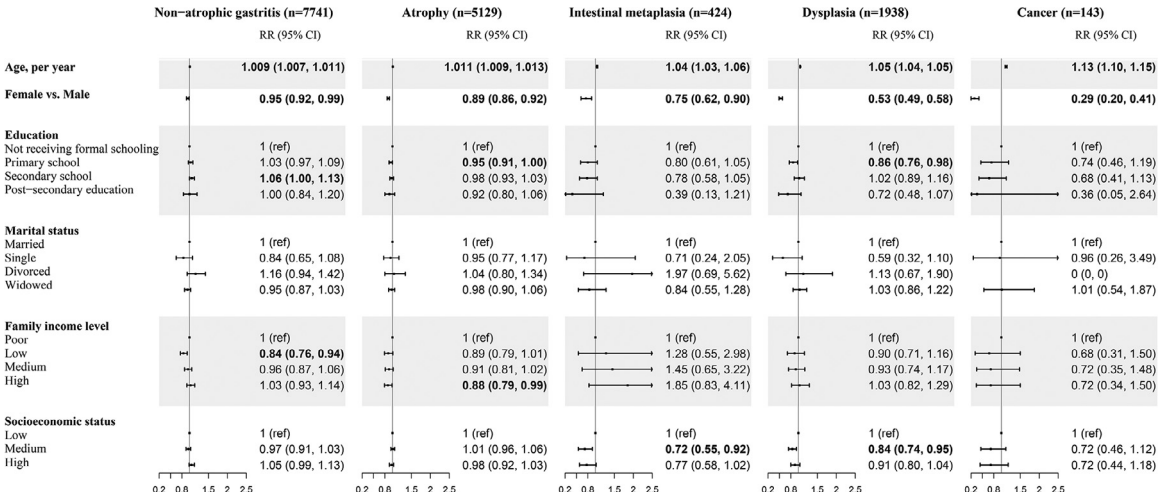

Figure 2. Forest plot of relative risk (RR) with 95% confidence interval (CI) of each gastric carcinogenesis stage in relation to sociodemographic factors. The RRs are indicated by boxes. Horizontal lines indicate the range of 95% CI. Horizontal lines with arrows reflect that the range of 95% confidence intervals exceeds the axis's limits. The dashed vertical line indicates the line of no effect. RR: relative risk; CI: confidence interval; ref: reference.

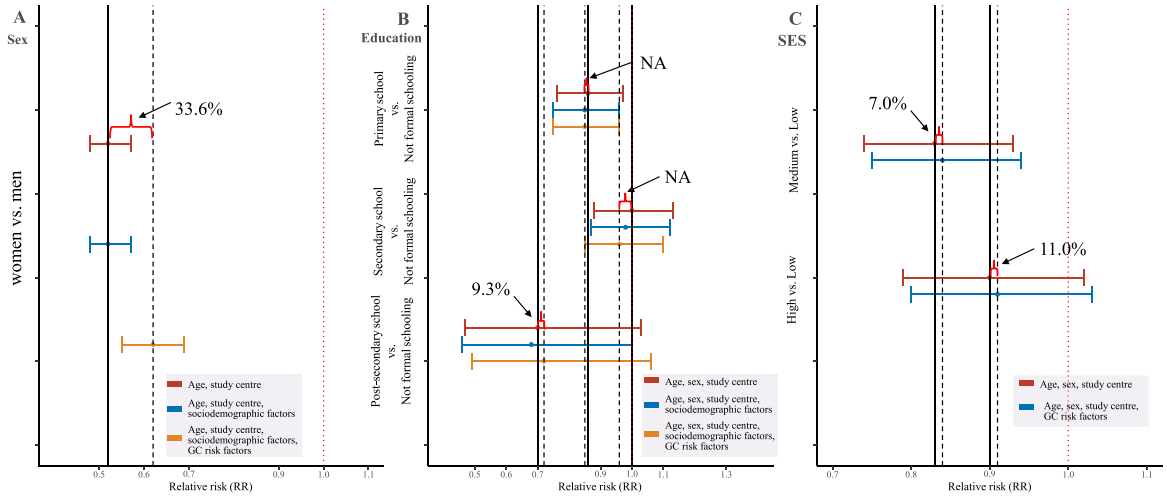

Figure 3. Relative risks of sex (A), education (B), socioeconomic status (C) for having neoplastic lesions, according to adjustment for different sets of covariates. Sociodemographic factors adjusted in the model: educational level, married status, family income level. GC risk factors adjusted in the model: smoking, alcohol drinking, fruit consumption, vegetable consumption, consumption of foods preserved by salting, occupational exposure to rubber or radiation, body mass index. NA: not available; SES: socioeconomic status; GC: gastric cancer.

Mechanisms underlying the male predominance in gastrointestinal cancer has attracted significant academic attention.[27] GC risk factors in our analysis could explain about one-third of the excess risk of neoplastic lesions in men, indicating that a population-wide reduction in GC risk factors may benefit men than women and narrow the sex gap. But a large proportion of sex disparities cannot still be justified, which may result from other mechanisms, such as the sex hormones difference.[28] A meta-analysis found that more prolonged oestrogen exposure can reduce gastric cancer risk among women.[29] Data from the UK Biobank indicated that male pattern baldness, a proxy of sex hormone levels, was associated with gastric cancer risk. But a recent cohort study failed to find the association of oestradiol concentrations with GC in men.[30] Unfortunately, we did not collect information on sex hormone levels or the potential indicators in our study. Other cohort studies are needed to examine the role of sex hormones in the sex gap. Furthermore, the effects of GC risk factors might differ among men and women. Previous studies[31] have shown that the association between smoking and GC was stronger in men than in women. Future studies are warranted to more fully elucidate the sources of the higher risk of neoplastic lesions in men.

Interestingly, our results suggested that a significant fraction of the excess risk in people with low SES or educational level cannot be attributed to the included risk factors, suggesting residual confounding or other potential mechanisms. Helicobacter pylori (Hp) infection, the most critical risk factor for GC, is more common in disadvantaged populations,[32] which might explain the SES inequalities. A pooled case-control study also observed no significant differences in SES effect across strata of smoking and drinking. But the association of SES with GC was stronger among individuals without Hp infection than those infected with Hp.[33] Future research is required to assess the impact of eliminating Hp infection on the social disparities in GC.

Our study provides significant insight to the mechanisms underlying sociodemographic disparities in GC from different angles. The analysis has several strengths, including the large sample size and the community-based survey in the general population. To our best knowledge, this is the first study to comprehensively examine sociodemographic disparities in each stage of Correa's cascade and quantify the contributions of main GC risk factors to social inequalities in neoplastic lesions. To avoid data-driven approach to select variable, we focused on well-established GC risk factors according to authoritative guidelines. We used modified Poisson regression to calculate RR to avoid the bias estimate of ERE%.

This study has several limitations. First, the study population was from high-risk areas in China, where genetic background and environmental exposure might differ from those in non-high-risk areas. The generalization of the results, therefore, must be made cautiously. Second, we included a limited set of sociodemographic factors. Not measuring other relevant variables such as childhood SES and occupation might bring residual confounding. In addition, diet, occupational exposure, and income level were all self-reported, which might result in measurement error and residual confounding. But considering our participants did not learn about their status of gastric mucosa before the questionnaire

survey, the recall bias was minimized. Finally, we used the baseline data instead of longitudinal data. Thus, we cannot ensure the temporal relationship between health behaviours and incident gastric lesions. Participants with neoplastic lesions might change their health behaviours due to symptoms even though they did not know the gastric mucosa status. From our data, we found that cases tended to experience a loss of appetite and the epigastric pain (Table S5). Nevertheless, the proportion of individuals with symptoms was low (smaller than 5%), which might not influence our results significantly.

In conclusion, age and sex were the most critical sociodemographic factors for all types of gastric lesions, whose effects increase with the progression of the cascade stage. About one-third of the excess risk of neoplastic lesions in men could be attributed to main GC risk factors, including diet, alcohol drinking, smoking, obesity, and occupational exposure. These findings advanced our understanding of carcinogenesis and suggested the potential to. People with higher educational levels and SES had a lower risk of neoplastic lesions than those without formal schooling and in a low SES. The GC risk factors included in our study could contribute to a small proportion of socioeconomic disparities in the neoplastic lesion risk, which might not be the ideal target for reducing socioeconomic gaps in GC risk. Future studies are needed to consider more factors and reveal further the mechanisms underlying the sociodemographic differences in GC.

Contributors

WC and DS conceived and designed the study. All authors participated in data collection and analyzation. DS, LL, CX, HL and MC have contributed to the interpretation of the data, DS and LL wrote the first draft of the manuscript. The dataset was accessed and verified by DS and WC. All authors contributed to the writing of the manuscript and have approved the final manuscript, and had final responsibility for the decision to submit for publication.

Data sharing statement

The data and code used to generate the reported estimates will be made available upon request to the corresponding author.

Declaration of interests

The authors declare no competing interests.

Acknowledgements

This work was supported by Sanming Project of Medicine in Shenzhen for SZSM201911015, National Natural Science Foundation of China for 81974492, and Special Project of Beijing-Tianjin-Hebei Basic Research Cooperation for J200017.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.lanwpc.2022.100437.

References

1 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–249.
2 Lin Y, Zheng Y, Wang HL, Wu J. Global patterns and trends in gastric cancer incidence rates (1988-2012) and predictions to 2030. *Gastroenterology*. 2021;161:116–127.e8.
3 La Vecchia C, Negri E, Franceschi S. Education and cancer risk. *Cancer*. 1992;70:2935–2941.
4 Uthman OA, Jadidi E, Moradi T. Socioeconomic position and incidence of gastric cancer: a systematic review and meta-analysis. *J Epidemiol Community Health*. 2013;67:854–860.
5 Gupta S, Tao L, Murphy JD, et al. Race/ethnicity-, socioeconomic status-, and anatomic subsite-specific risks for gastric cancer. *Gastroenterology*. 2019;156:59–62.e4.
6 Correa P, Haenszel W, Cuello C, Tannenbaum S, Archer M. A model for gastric cancer epidemiology. *Lancet*. 1975;2:58–60.
7 Stringhini S, Dugravot A, Shipley M, et al. Health behaviours, socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. *PLoS Med*. 2011;8:e1000419.
8 Ovrum A. Socioeconomic status and lifestyle choices: evidence from latent class analysis. *Health Econ*. 2011;20:971–984.
9 Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet*. 2020;396:635–648.
10 Chen W, Zeng H, Chen R, et al. Evaluating efficacy of screening for upper gastrointestinal cancer in China: a study protocol for a randomized controlled trial. *Chin J Cancer Res*. 2017;29:294–302.
11 Cai Q, Zhu C, Yuan Y, et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: a nationwide multicentre study. *Gut*. 2019;68:1576–1587.
12 Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. *BMJ*. 2021;373:n604.
13 World Cancer Research Fund International. Diet, nutrition, physical activity and cancer: a global perspective (the Third Expert Report): world cancer research fund international; 2018.
14 Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *J Natl Cancer Inst*. 2011;103:1827–1839.
15 VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health*. 2016;37:17–32.
16 Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159:702–706.
17 Begg CB, Gray R. Calculation of polychotomous logistic regression parameters using individualized regressions. *Biometrika*. 1984;71:11–18.
18 Szklo M, Nieto J. *Epidemiology: Beyond the basics*. Burlington: Jones & Bartlett Learning; 2019.
19 Aumpan N, Vilaichone RK, Pornthisarn B, et al. Predictors for regression and progression of intestinal metaplasia (IM): a large population-based study from low prevalence area of gastric cancer (IM-predictor trial). *PloS one*. 2021;16:e0255601.
20 de Vries AC, van Grieken NC, Looman CW, et al. Gastric cancer risk in patients with premalignant gastric lesions: a nationwide cohort study in the Netherlands. *Gastroenterology*. 2008;134:945–952.
21 Pittayanon R, Rerknimitr R, Klaikaew N, et al. The risk of gastric cancer in patients with gastric intestinal metaplasia in 5-year follow-up. *Aliment Pharmacol Ther*. 2017;46:40–45.
22 Leung WK, Lin SR, Ching JY, et al. Factors predicting progression of gastric intestinal metaplasia: results of a randomised trial on Helicobacter pylori eradication. *Gut*. 2004;53:1244–1249.

23 Sarkar S, Dauer MJ, In H. Socioeconomic disparities in gastric cancer and identification of a single SES variable for predicting risk. *J Gastrointest Cancer*. 2021.
24 Nagel G, Linseisen J, Boshuizen HC, et al. Socioeconomic position and the risk of gastric and oesophageal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST). *Int J Epidemiol*. 2007;36:66–76.
25 den Hollander WJ, Holster IL, den Hoed CM, et al. Surveillance of premalignant gastric lesions: a multicentre prospective cohort study from low incidence regions. *Gut*. 2018;68:585–593.
26 Xing J, Min L, Zhu S, et al. Factors associated with gastric adenocarcinoma and dysplasia in patients with chronic gastritis: a population-based study. *Chin J Cancer Res*. 2017;29:341–350.
27 Niedermaier T, Heisser T, Gies A, et al. To what extent is male excess risk of advanced colorectal neoplasms explained by known risk factors? Results from a large German screening population. *Int J Cancer*. 2021;149:1877–1886.
28 Chandanos E, Lagergren J. Oestrogen and the enigmatic male predominance of gastric cancer. *Eur J Cancer*. 2008;44:2397–2403.
29 Camargo MC, Goto Y, Zabaleta J, Morgan DR, Correa P, Rabkin CS. Sex hormones, hormonal interventions, and gastric cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2012;21:20–38.
30 McMenamin ÚC, Liu P, Kunzmann AT, et al. Circulating sex hormones are associated with gastric and colorectal cancers but not esophageal adenocarcinoma in the UK Biobank. *Am J Gastroenterol*. 2021;116:522–529.
31 Nomura AM, Wilkens LR, Henderson BE, Epplein M, Kolonel LN. The association of cigarette smoking with gastric cancer: the multiethnic cohort study. *Cancer Causes Control*. 2012;23:51–58.
32 Moayyedi P, Axon AT, Feltbower R, et al. Relation of adult lifestyle and socioeconomic factors to the prevalence of Helicobacter pylori infection. *Int J Epidemiol*. 2002;31:624–631.
33 Rota M, Alicandro G, Pelucchi C, et al. Education and gastric cancer risk-An individual participant data meta-analysis in the StoP project consortium. *Int J Epidemiol*. 2020;146:671–681.