Association between Socioeconomic Factors and the Risk of Gastric Cancer Incidence: results from an ecological study

Article in Iranian Journal of Public Health - August 2022

3 authors, including:

Jafar Rezaie
Urmia University of Medical Sciences
102 PUBLICATIONS 2,449 CITATIONS

Majid Babaei
Urmia University of Medical Sciences
6 PUBLICATIONS 76 CITATIONS

Some of the authors of this publication are also working on these related projects:

Regenerative Medicine View project
Exosome biology View project
Association between Socioeconomic Factors and the Risk of Gastric Cancer Incidence: Results from an Ecological Study

*Majid Babaei, Habibollah Pirnejad, Jafar Rezaie, Gholamreza Roshandel, Rana Hoseini

1. Social Determinants of Health Research Center, Clinical Research Institute, Urmia University of Medical Sciences, Urmia, Iran
2. Patient Safety Research Center, Clinical Research Institute, Urmia University of Medical Sciences, Urmia, Iran
3. Erasmus School of Health Policy & Management (ESHPM), Erasmus University Rotterdam, Rotterdam, Netherlands
4. Solid Tumor Research Center, Cellular and Molecular Medicine Research Institute, Urmia University of Medical Sciences, Urmia, Iran
5. Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran

*Corresponding Author: Email: Babaei.m@umsu.ac.ir

(Received 20 Jun 2022; accepted 18 Aug 2022)

Abstract

Background: Gastric cancer (GC), one of the most common cancers worldwide, remains the third leading cause of cancer-related mortality. The etiology of GC may arise from genetic and environmental factors. This study aimed to determine the association between GC incidence and socioeconomic status in Iran.

Methods: An ecological study was designed to investigate the relationship between socioeconomic factors and the risk of GC incidence. The data of socioeconomic variables such as income changes, unemployment rate, urbanization ratio, inflation rate, and air pollution changes in 31 provinces were collected from the Statistical Center of Iran, and the data of GC of 31 provinces were provided from the Iranian National Population-based Cancer Registry (INPCR). Data from 2014 to 2017 was analyzed using panel data analysis, the fixed effects model by EViews software.

Results: Panel data model was suitable for the present study. Results showed that there was a positive and significant relationship between GC incidence and socioeconomic factors including income changes ($P \leq 0.001$), unemployment rate ($P \leq 0.01$), inflation rate ($P \leq 0.05$), and air pollution changes ($P \leq 0.001$). The urbanization ratio showed a negative relationship and was not statistically associated with GC incidence ($P > 0.05$).

Conclusion: Our findings suggest a positive and significant association between socioeconomic status and GC incidence, proposing a GC risk factor. The key public health policies and welfare policies’ priority should therefore be to schedule for the GC management.

Keywords: Gastric cancer; Socioeconomic factors; Panel data; Income; Urbanization ratio

Introduction

Gastric cancer (GC), which forms from cells in the lining of the stomach, remains a leading cause of death globally (1). For the past years, although GC's incidence rate has decreased, it remains among the most common cancers in low- and middle-income countries (2). GC is the fourth
most common malignancy and the third leading cause of cancer-related death globally (3). The incidence and mortality rates of GC are the highest in East Asia, for example, in Japan, this cancer accounts for nearly 15% of cancer-related mortality annually over the past 4 decades (4). Furthermore, data show that these rates hold at a high level in Eastern Europe, South and Central America, and also regions of the Middle East (5).

In the Middle East encompassing three main cultures including Turkic, Semitic, and Indo-European, the incidence rate varies from very low in Egypt, with an age-standardized incidence rate (ASR) of 3.4 people, to very high in Iran, with an ASR of 26.1/100000 (4).

According to statistics, in Iran, GC is the most common cancer (third) and its incidence is increasing; accordingly, it is the first and fourth one among men and women respectively (6, 7). It happens approximately seven times more frequently in Iran than in Iraq (4). GC is a multifactorial disease and several factors may impact its progress, both genetic and environmental (8, 9). GC is associated with several risk factors, including the type of diet (low vegetable and fruit feeding), obesity, smoking habits, chronic digestive diseases, and physical inactivity (10, 11). In recent years, there has been an increasing interest in determining the other risk factors of GC incidence and its mortality rates, such as socioeconomic factors (12, 13).

A socioeconomic position involving public health strategies, accessibility of preservation, and neighborhood socioeconomic position have been reported to affect cancer (14-16). Socioeconomic inequalities are strongly associated with the incidence and mortality rates of diseases such as cancers (11, 17, 18). For example, a higher socioeconomic status was associated with a declined risk of gastric malignancy (19). Along with progress in GC treatment, however, there is an increasing concern over the GC incidence and mortality (3). Uncovering socioeconomic risk factors of GC will be step forward in controlling its morbidity and mortality. The aim of this study was to study the association between socioeconomic status and the risk of GC incidence in Iran.

### Methods

#### Data collection

We performed an ecological study to investigate the relationship between socioeconomic factors and the risk of GC incidence. The data was collected from an Iranian database related to the Statistical Center of Iran from 2014 to 2017. The database encompassing the socioeconomic variables such as income changes, unemployment rate, urbanization ratio, inflation rate, and air pollution changes of 31 provinces of Iran. Furthermore, the data of GC was obtained from the Iranian National Population-based Cancer Registry (INPCR) of the Ministry of Health and Medical Education, Iran.

#### Ethics approval

The present study was approved by a Research ethics committee of Urmia University of Medical Sciences, Urmia, Iran (IR.UMSU.REC.1400.079). There was no human participant in this work. The data was collected from an Iranian database related to the Statistical Center of Iran from 2014 to 2017 and the data of GC was obtained from the Iranian National Population-based Cancer Registry (INPCR) of the Ministry of Health and Medical Education, Iran.

#### Methodology quality assessment

The present study was used panel data that comprises cross sections in 31 provinces and a complete-time series analysis from 2014 to 2017. Data was processed by EViews software (ver. 10). For the present study, after different examinations of various models, we used panel data, the fixed-effect model for collected data.

The formula of the panel data is as follows:

\[ Y_{it} = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + \ldots + \beta_k X_{kit} + \epsilon_{it} \]

Y: depended variable; \( X_1 - X_k \): Independent variable; \( \beta_1 - \beta_k \): Coefficients; \( \alpha \): intercept; \( \epsilon \): error terms; \( i \): the number of sections; \( t \): the number of courses studied for each section.

To examine the independence of the independent variables from each other, the collinearity of the
independent variables was tested using a correlation matrix as well as multicollinearity. Then, for confirmation of the panel data method, and assessment of the intercept for each of the provinces, the following F-Limer test was used. Null Hypothesis states that $\alpha_i$ is constant for all sections:

$H_0: \alpha_0 = \alpha_1 = \ldots = \alpha_{31} = \alpha$

$H_1: \alpha_i \neq \alpha_i$

We also used a test called the Hausman test to compare fixed and random effects models in terms of the explanatory power of the dependent variable. Null Hypothesis is as follows:

$H_0$: fixed effects

$H_1$: random effects

Finally, we tested the classical hypotheses including zero mean errors, homogeneity of error variance, non-correlation of errors, normality of error terms, and independence of errors.

### Results

**Methodological quality**

Using the Panel Data method, the relationship between five independent socioeconomic variables and the incidence of GC, the dependent variable, was analyzed. The estimated equation is represented here:

$$G_{Cancer_{it}} = \alpha_i + \beta_1\text{INCOME}_{it} + \beta_2\text{UNE}_{it} + \beta_3\text{UR_RATIO}_{it} + \beta_4\text{INFLATION}_{it} + \beta_5\text{AIR_P}_{it} + \epsilon_{it}$$

$G_{Cancer}$: gastric cancer incidence; INCOME: income changes; UNE: unemployment rate; UR_RATIO: urbanization ratio; INFLATION: inflation rate; AIR_P: air pollution changes (changes in consumption of gasoline, gas oil, fuel oil, and kerosene); $\beta_1$-$\beta_5$: coefficients; $\alpha_i$: intercept; $\epsilon_{it}$: error terms; $i$: 1, 2, ..., 31 is the number of sections (provinces); $t$: 1, ..., 4: the number of studied courses for each section.

The descriptive statistics of variables are summarized in Table 1.

### Table 1: Descriptive statistics of variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>CANCER</th>
<th>INCOME</th>
<th>UNE</th>
<th>UR_RATIO</th>
<th>INFLATION</th>
<th>AIR_P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.000132</td>
<td>0.157108</td>
<td>0.123677</td>
<td>0.696129</td>
<td>0.106652</td>
<td>0.006342</td>
</tr>
<tr>
<td>Median</td>
<td>0.000120</td>
<td>0.151798</td>
<td>0.114500</td>
<td>0.684000</td>
<td>0.099870</td>
<td>0.019429</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.000310</td>
<td>0.595287</td>
<td>0.245000</td>
<td>0.952000</td>
<td>0.195077</td>
<td>0.352321</td>
</tr>
<tr>
<td>Minimum</td>
<td>4.00E-05</td>
<td>-0.070377</td>
<td>0.056000</td>
<td>0.485000</td>
<td>0.046205</td>
<td>-0.356406</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>5.94E-05</td>
<td>0.088463</td>
<td>0.038841</td>
<td>0.120994</td>
<td>0.031536</td>
<td>0.121208</td>
</tr>
</tbody>
</table>

INCOME: average income changes, UNE: unemployment rate, UR_RATIO: urbanization ratio, INFLATION: inflation rate, and AIR_P: air pollution changes related to fossil energy use (gasoline, gas oil, fuel oil, and kerosene), Std. Dev: Standard deviation.

The key issue in panel data models is the stationarity of data. As shown in Table 2, using the correlation matrix, there was no collinearity between the independent variables, thus the regression hypothesis was established. We next used the VIF test to evaluate the multicollinearity in independent variables. We observed no multicollinearity in our results (Table 3).
Table 2: Results of the correlation matrix (collinearity test)

<table>
<thead>
<tr>
<th></th>
<th>INCOME</th>
<th>UNE</th>
<th>UR_RATIO</th>
<th>INFLATION</th>
<th>AIR_P</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCOME</td>
<td>1.000000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UNE</td>
<td>-0.135759</td>
<td>1.000000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UR_RATIO</td>
<td>-0.018755</td>
<td>-0.107101</td>
<td>1.000000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>INFLATION</td>
<td>0.377047</td>
<td>-0.082514</td>
<td>-0.106092</td>
<td>1.000000</td>
<td>-</td>
</tr>
<tr>
<td>AIR_P</td>
<td>0.009814</td>
<td>-0.072111</td>
<td>-0.050855</td>
<td>-0.053962</td>
<td>1.000000</td>
</tr>
</tbody>
</table>

INCOME: average income changes, UNE: unemployment rate, UR_RATIO: urbanization ratio, INFLATION: inflation rate, and AIR_P: air pollution changes related to fossil energy use (gasoline, gas oil, fuel oil, and kerosene). The number outside the parentheses indicates the degree of correlation between the independent variables and the number inside the parentheses indicates the probability value.

Table 3: Results of multicollinearity test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient Variance</th>
<th>Uncentered VIF</th>
<th>Centered VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCOME</td>
<td>4.64E-09</td>
<td>4.940762</td>
<td>1.154490</td>
</tr>
<tr>
<td>UNE</td>
<td>2.04E-08</td>
<td>11.66907</td>
<td>1.031098</td>
</tr>
<tr>
<td>UR_RATIO</td>
<td>2.17E-09</td>
<td>36.38136</td>
<td>1.034853</td>
</tr>
<tr>
<td>INFLATION</td>
<td>3.55E-08</td>
<td>14.74279</td>
<td>1.166472</td>
</tr>
<tr>
<td>AIR_P</td>
<td>2.04E-09</td>
<td>1.016930</td>
<td>1.014208</td>
</tr>
</tbody>
</table>

G_cancer: Gastric cancer incidence, INCOME: average income changes, UNE: unemployment rate, UR_RATIO: urbanization ratio, INFLATION: inflation rate, and AIR_P: air pollution changes related to fossil energy use (gasoline, gas oil, fuel oil, and kerosene).

To verify the Panel data model, the F-Limer was used accordingly. The output of Eviews software confirmed that the Panel data model is suitable for our study (F-Statistic: 21.067756; P=0.0000, Table 4). The explanatory variables and errors are independent (Table 5); therefore, the independence of explanatory variables and errors was rejected. And the White cross-section method was used.

Table 4: Results of F-Limer test

<table>
<thead>
<tr>
<th>Test</th>
<th>F-Statistic</th>
<th>Prob</th>
<th>d.f</th>
<th>Null Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>-limer F</td>
<td>21.067756</td>
<td>0.0000</td>
<td>(30,87)</td>
<td>Pool Data</td>
<td>Panel Data</td>
</tr>
</tbody>
</table>

Table 5: Results of the independence test for the explanatory variable and errors

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>Prob</th>
<th>d.f</th>
<th>Null Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breuscher-Pagan</td>
<td>652.2650</td>
<td>0.0000</td>
<td>(465)</td>
<td>No cross-section dependency</td>
<td>There is dependency</td>
</tr>
</tbody>
</table>

Of note, when $\varepsilon_i$ and $X_i$ are not correlated, the random effects model is suitable, and when $\varepsilon_i$ and $X_i$ are correlated, the fixed effects model would be more suitable (20). Consequently, there was
no need to perform the Hausman test; a panel data model with fixed effects was selected. Additionally, the classical hypothesis was analyzed to estimate the final model. In this regard, such a hypothesis is established as (i) the mean value of the errors is zero; (ii) variances are homogenous (Table 6); (iii) there is no autocorrelation in the residuals (Durbin-Watson = 2.3, Table 8); and (iv) the residuals are normally distributed (Fig. 1).

**Table 6: Results of variance heterogeneity test**

<table>
<thead>
<tr>
<th>Test</th>
<th>F-Statistic</th>
<th>Prob</th>
<th>Null Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>1.1397971</td>
<td>0.1407</td>
<td>Homogeneity of variance</td>
<td>There is no heteroscedasticity</td>
</tr>
</tbody>
</table>

**Fig. 1: The normality of residuals. Data showed that the residuals were distributed normally**

**Association of socioeconomic factors with the risk of the GC incidence**

The output of the model, presented in Table 7 and 8, shows this model was highly significant (F-Statistic: 82.32235; \( P=0.0000 \)). The socioeconomic factors (except for the urbanization ratio) showed a significant correlation (\( P<0.05 \)) with the incidence of GC in Iran. Here, the adjusted coefficient of determination was 0.95, indicating 95% of the changes in the dependent variable are explained by the independent variables in this model which in turn showed the model was specific (Table 8). As mentioned in the classical hypotheses, Durbin-Watson test statistic value is 2.3, therefore, we concluded that there is no autocorrelation in the residuals.
Table 7: Results of hypothesis test and coefficients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCOME</td>
<td>3.70E-05</td>
<td>1.09E-05</td>
<td>3.399063</td>
<td>0.0010</td>
</tr>
<tr>
<td>UNE</td>
<td>9.34E-05</td>
<td>3.54E-05</td>
<td>2.634428</td>
<td>0.0100</td>
</tr>
<tr>
<td>UR_RATIO</td>
<td>-0.000206</td>
<td>0.000365</td>
<td>-0.562725</td>
<td>0.5751</td>
</tr>
<tr>
<td>INFLATION</td>
<td>6.29E-05</td>
<td>2.91E-05</td>
<td>2.156787</td>
<td>0.0338</td>
</tr>
<tr>
<td>AIR_P</td>
<td>2.57E-05</td>
<td>6.13E-06</td>
<td>4.189878</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

INCOME: average income changes, UNE: unemployment rate, UR_RATIO: urbanization ratio, INFLATION: inflation rate, and AIR_P: air pollution changes related to fossil energy use (gasoline, gas oil, fuel oil, and kerosene).

Table 8: Results of regression

<table>
<thead>
<tr>
<th>F-statistic</th>
<th>Prob(F-statistic)</th>
<th>Durbin-Watson stat</th>
<th>R-squared</th>
<th>Adjusted R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>82.32235</td>
<td>0.000000</td>
<td>2.301116</td>
<td>0.970690</td>
<td>0.958899</td>
</tr>
</tbody>
</table>

The distribution of GC from 2014 to 2017 in 31 provinces of Iran is illustrated in Fig. 2. The incidence of GC in the northwestern provinces of Iran is high while it is low in southern. Figure S1 presents the distribution of socioeconomic factors across different provinces from 2014 to 2017.

Fig. 2: The population-based GC incidence in Iran’s provinces (per 100,000 populations from 2014 to 2017). Data shows that the incidence of GC in the northwestern area of Iran is higher than in other area.

Available at: [http://ijph.tums.ac.ir](http://ijph.tums.ac.ir)
According to Table 7, there was a positive and significant relationship between income changes and the GC incidence ($P \leq 0.001$). A similar result was obtained when the relationship between the variable unemployment rate and GC incidence had been analyzed ($P \leq 0.01$). Furthermore, we observed a negative relationship between urbanization ratio rate and the GC incidence although this relationship was non-significant ($P \geq 0.05$; Table 7). A significant and positive relationship was found between the GC incidence and inflation rate ($P \leq 0.05$) as well as the air pollution ratio ($P \leq 0.001$). Furthermore, in terms of the importance and effectiveness of independent variables in this study, according to the results, the unemployment rate, inflation rate, income changes, and air pollution changes have the greatest impact on the incidence of GC in Iran, respectively.

**Discussion**

The correlation between socioeconomic factors and the risk of GC incidence is worth mentioning because socioeconomic factors may affect GC incidence. Factors (such as socioeconomic factors) other than genetic can affect cancer incidence (21, 22). The existing evidence that the incidence of cancer is higher in developing countries supports such a claim (23-26). As cancer incidence rates in Iran were reported to be higher than that of other regions in Asia (27, 28), here, we aimed to examine the impact of socioeconomical factors on the incidence of one of the deadliest and prevalent cancers in Iran, i.e. GC, using data panel model analysis. Similar to the results of a study, we found a significant association between socioeconomic status and GC incidence (21). Our results from income analysis was in line with those of recently reported (29). However, despite being evident a metaanalysis did not reach a statistically significant relation between the level of income and GC incidence (30).

Studies also evaluated the effect of unemployment on the incidence of GC. A historical prospective cohort study, for example, indicated that GC was associated with unemployment 2 years after diagnosis (31). Our results share similarities (32) findings where they found that there was a significant association between unemployed rate and an increased risk of GC. Furthermore, the inflation rate revealed that this factor had a significant and positive relationship with the GC incidence. To our knowledge, this finding is reported for the first time and additional studies would be useful for further confirmation. Urbanization ratio was another socioeconomic risk factor of GC introduced in publications.

Concur with these findings, we found a negative but non-significant relationship between the risk of GC and the urbanization rate. Using univariate and multivariable analysis, a population-based study conducted in the USA showed that with increasing urbanization ratio the risk of GC incidence enhanced (33), whereas Chang et al., by using multilevel logistic regression models, reported that urbanization exhibited a negative and significant association with GC incidence in Korean individuals (34). Contrary to our finding, Momenyan et al., by using quantile regression demonstrated that with an increasing urbanization ratio, the incidence rate of GC was reduced in Iran (35). This study was carried out for one year involving 345 cities. Presumably, this discrepancy was raised from the design of the studies.

We analyzed the possible relationship between GC incidence and air pollution changes as one of the urbanization results. Similar to previous studies (36, 37), we found the risk of GC was associated with air pollution change. In our opinion, it may be a relation between air pollution changes and the urbanization ratio. Therefore, it is possible that the urbanization rate solely is not a potent risk factor for GC incidence; however, air pollution may promote the GC incidence with increasing urbanization rate concurrently.

Overall, the panel data with a fixed effects model was highly significant. And using different current examinations and tests the suitability and specificity of the analysis method were confirmed. We found that the majority of socioeconomical factors were associated with the risk of
GC incidence. These results offer significant evidence for the association of socioeconomic factors with GC incidence in Iran.

Conclusion

Our research underlined the importance of socioeconomic status in GC incidence. Socioeconomic factors positively impact GC incidence, suggesting a significant risk factor for this cancer. An implication of this study is the possibility that the selected socioeconomic factors may participate in GC incidence in Iran. Further research regarding the role of these factors would be worthwhile. Another implication is that public health policies, welfare policies, and education should be taken into account for improving GC risk.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

We would like to thank the Vice Chancellor for Research of Urmia University of Medical Sciences for his support and motivation.

Funding

This work was supported by the Research Committee of Urmia University of Medical Sciences [grant number: IR.UMSU.REC.1400.079].

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interest

The authors declare that there is no conflict of interests.

References


