RESEARCH ARTICLE

Geographical variation and factors associated with gastric cancer in Manitoba

Oluwagbenga Fakanye[1☉], Harminder Singh[2‡], Danielle Desautels[3‡], Mahmoud Torabi[1☉]*

1 Department of Community Health Sciences, University of Manitoba, Manitoba, Canada, **2** Department of Internal Medicine, University of Manitoba, Manitoba, Canada, **3** Research Institute in Oncology and Hematology, CancerCare Manitoba, University of Manitoba

☉ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* Mahmoud.torabi@umanitoba.ca

Abstract

Objectives

We investigated the spatial disparities and factors associated with gastric cancer (GC) Incidence in Manitoba.

Methods

We combined information from Manitoba Cancer registry and Census data to obtain an age-sex adjusted relative risk (IRR) of GC incidence. We geocoded the IRR to the 96 regional health authority districts (RHADs) using the postal code conversion file (PCCF). Bayesian spatial and spatio-temporal Poisson regression models were used for the analysis.

Results

Adjusting for the effect of socio-economic score index (SESI), Indigenous, and immigrant population, 25 districts with high overall GC risk were identified. One unit increase in SESI was associated with reduced risk of cardia GC (CGC) by 14% (IRR = 0.859; 95% CI: 0.780–0.947) and the risk of non-cardia GC (NCGC) by approximately 10% (IRR = 0.898; 95% CI: 0.812–0.995); 1% increase in regional Indigenous population proportion reduced the risk of CGC by 1.4% (IRR = 0.986; 95% CI: 0.978–0.994). In the analysis stratified by sex, one unit increase in SESI reduced the risk of CGC among women by 26.2% (IRR = 0.738; 95% CI: 0.618–0.879), and a 1% increase in Indigenous population proportion reduced the risk of CGC among women by 1.9% (IRR = 0.981; 95% CI: 0.966–0.996).

Conclusion

Our results support a significant association between SESI and NCGC. We report regional variation of GC IRR and a varying temporal pattern across the RHADs. These results could be used to prioritize interventions for regions with high and progressive risk of GC.

Introduction

One of the cancers that has consistently contributed to the global burden of cancer is Gastric Cancer (GC). It is the fifth most diagnosed cancer worldwide, which accounts for 6.1% of all cancer cases for both sexes. It stands as the fourth most diagnosed cancer in men accounting for 7.8% of all cancers affecting men globally and the seventh most diagnosed cancer in women accounting for 4.3% of all cancers affecting women globally [1].

Almost half (43.5%) of all the GC cases are diagnosed at an advanced stage (stage IV), where cancer has metastasized to other parts of the body [2]. Therefore, only one third of all diagnosed GC cases live beyond five years, which is one of the lowest cancer survival rates [3]. In some parts of the world, researchers have studied the geographical and temporal variation of GC with the sole purpose of understanding its propagation [4–7]. However, there is very limited information about the geographical variation in Canada. The purpose of this study was to identify hot-spot districts and associated area-level risk factors of GC, across the 96 Regional Health Authority Districts (RHADs) in Manitoba using a twenty-five year (1992–2016) Canadian Cancer Registry (CCR) dataset. Specifically, this study investigated the geographical variation of GC incidence in Manitoba, explored factors influencing the geographic difference of GC incidence in Manitoba, and also investigated the geographic variation of GC incidence over time in Manitoba.

Materials and method

Study area

Manitoba, the study region, has an approximate population of 1,278,365 based on the 2016 population census. 631,400 (49.39%) identified as male, while 646,965 (50.61%) identified as female. Winnipeg is the capital of the province, with more than half (55.96%) of the entire Manitoba population. It is a diversified multi-cultural ethnic province with a higher proportion of Indigenous peoples than many other Canadian provinces [8]. The province has five regional health authorities (RHAs) created from the initial eleven RHAs in 2012. These RHAs are tasked with overseeing health services, both acute and community-based care. The RHAs are further sub-divided into districts (RHADs), [9]. The RHADs (Fig 1), were adopted in this study as the area-level aggregate point. At the time of this study, the province was divided into 96 RHADs, which were used as the analysis unit in this study as used in other similar geographical studies [10]. ArcGIS, version 10.2.2 [11], was used to create choropleth maps displayed in subsequent sections. All maps presented in this study are created from the shapefiles which form part of the dataset used in the analysis. It does not require a copyright permission.

Data

Using the universal disease classification code, as defined by the international classification of disease for oncology third Edition (ICD-O3) [12], we extracted all cases of GC (C160 –C169) in Manitoba from 1992 to 2016 from the Canadian Cancer Registry (CCR) (Table 1). Prior to 1992, there was no unified dynamic database for cancers in Canada [13].

We also obtained demographic data (age and sex) from the CCR data. Upon the satisfaction of the ethical condition requirement by the Research Data Centre (RDC) and a course completion on research ethics by the Tri-Council Policy statement Ethical Conduct (TCPS2: CORE), we were granted access to the CCR dataset from the Research Data Centre (RDC) repository located at the University of Manitoba campuses (Bannatyne and Fort-Gary). All data extracted from the CCR used in this study were fully anonymized before access to avoid bridging of confidentiality. We created an age-sex standardized incidence risk (ASSIR). We obtained
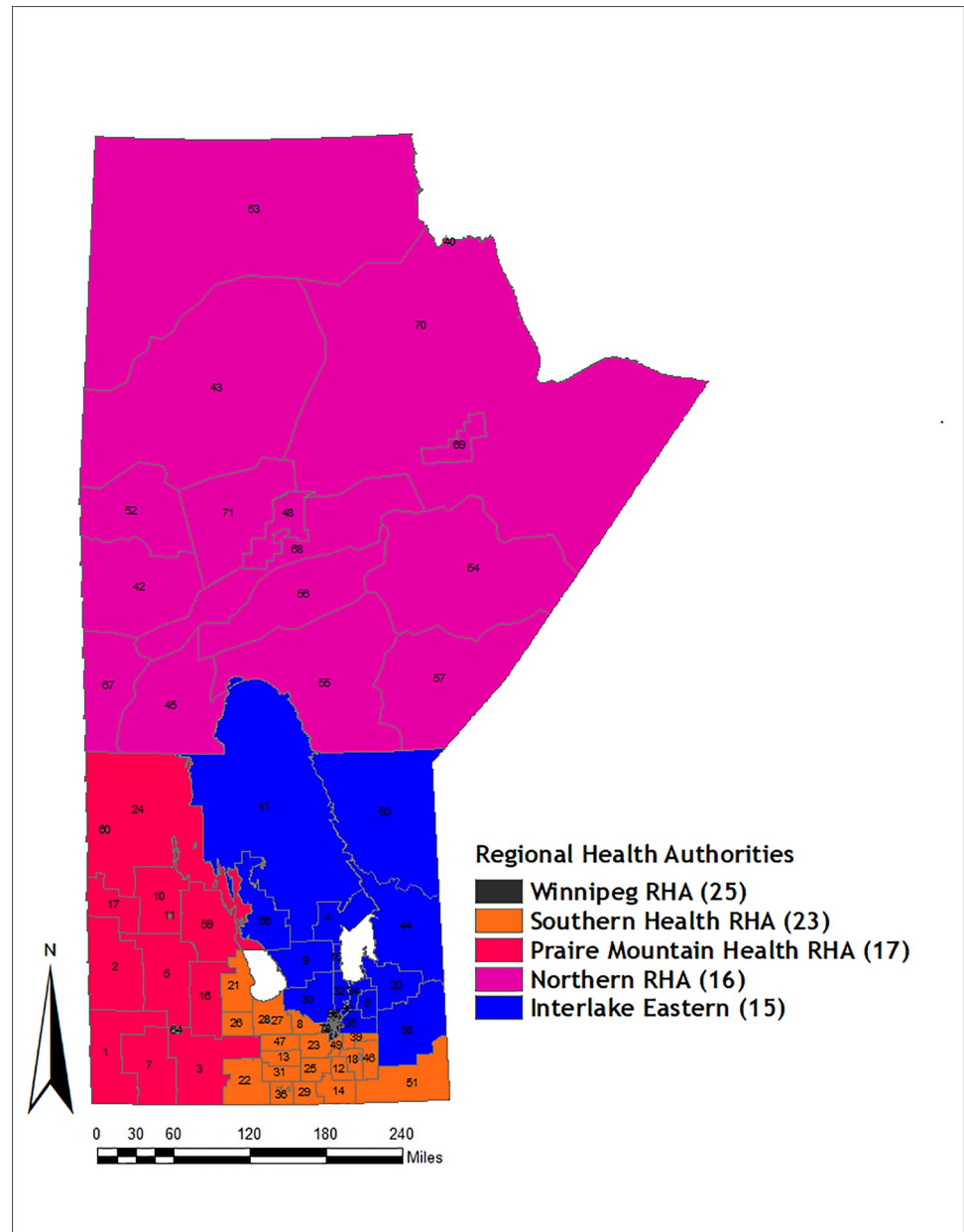
Fig 1. Map of five Regional Health Authorities (RHA) and the corresponding Regional Health Authority Districts in Manitoba. Number of districts for each RHA is shown in parenthesis in the legend. *Note: The numbers on the map and subsequent ones in this study represents area identification tag.

https://doi.org/10.1371/journal.pone.0253650.g001

information for average socioeconomic status in RHADs, proportion of individuals of Indigenous and immigrant population in each RHAD from the 2016 Census datasets. The 2016 Census data was used based on the assumption that the population of Manitoba within the studied period has been relatively steady (1.1 million– 1.2 million).

Socioeconomic status score index. Several studies have linked Socioeconomic status (SES) with the etiology of GC [14–16], which partly contributed to the need to adjust for it in this study. The socioeconomic status score index (SESI) used in this study was computed via a linear combination of income, employment and education using factor analysis (FA); a

Table 1. Description of all GC topographies using ICD-O3.

ICDO-O3 code	Topography sub-site
C160	Cardia
C161	Fundus
C162	Body
C163	Antrium
C164	Pyloric
C165	Lesser curvature of the stomach
C166	Greater Curvature of stomach
C168	Overlapping lesion of the stomach
C169	Stomach unspecified

https://doi.org/10.1371/journal.pone.0253650.t001

method that has been widely used in SESI computation [17–19]. We adopted the FA algorithm implemented by [20], where SESI value ranged from 0.643397 to 9.4163. The lowest score indicates poor socio-economic status while the highest score indicates a good socio-economic status (high standard of living).

Immigrant population. Considering the diversity in Manitoba population, another factor of interest is the distribution of immigrant across the province. This factor is considered important in this study as our interest is on geographical variation of GC which could be linked to corresponding geographical variation in the distribution of immigrants across the province. There have been studies that looked into the association between GC and immigration in different countries. However, the result has been inconsistent [21–23]. The measure of immigrant population variable used in this study was the districts' percentage population of individuals residing in Manitoba during the study period who identified themselves as an immigrant after 1965 (according to the definition of 2016 Canadian Census).

Indigenous population. Another essential factor of interest that has extensively been researched in the epidemiological study of GC in Canada is the Indigenous population [6, 24, 25]. Despite the extensive study of this factor, little information is available about its impact on geographical variation of GC in Manitoba. The measure of Indigenous population variable, as used here, represents the districts' percentage of Manitoba population who identified as First Nation, Metis or Inuit.

Statistical analysis

Information extracted from the CCR and the 2016 Census data were merged using the postal code conversion file (PCCF) which provides a correspondence between the Canadian six-alphanumeric postal code and Statistics Canada's standard geographic area. We standardized the GC incidence by age and sex since our primary focus did not include these two variables. This was also done to eliminate the possible impact of age and sex on the parameter estimates which may lead to spurious estimates. GC cases were stratified into cardia and non-cardia GC based on topographical sub-site, and into 5-year interval (1992–1996, 1997–2001, 2002–2006, 2007–2011, 2012–2016) to examine the risk of GC over time; note that we split our data into 5-year category to avoid sparsity in our dataset. A Moran's I statistic was used as a confirmatory test to assess the spatial dependency of GC, as a justification for using spatial regression model. Bayesian spatial and spatio-temporal Poisson regression models are mathematically defined later in Eqs (1.2) and (1.4) respectively were used to address the research objectives [26–30].

Bayesian regression model. Let $y_{ijk}; i = 1, \ldots, 96; j = 1, 2; k = 1, \ldots, 12$ represents the GC count in RHAD i, sex group j and age group $k = less\ than\ 35, 35-39, 40-44, \ldots, 85+$; y_{ik} is the GC count in sex j and age group k; y_i; $I = 1, \ldots, 96$ is the GC count in RHAD i; n_{ijk} is the population of people in RHAD i belonging to sex group j and age group k; n_{jk} is the population of people in sex j and age group k; x_i is the covariates in RHAD i, (SESI, the proportion of Indigenous, the proportion of immigrants), and assuming the total number of cases of GC in each RHAD follows a Poisson distribution, i.e. $Y_i \sim Poisson(\lambda_i E_i)$, where λ_i is relative risk and E_i is the expected number of GC count in RHAD i defined in Eqs (1.0) and (1.1),

$$Y_i = \sum_{j=1}^{2} \sum_{k=1}^{12} y_{ijk} \tag{1.0}$$

$$E_i = \sum_{j=1}^{2} \sum_{k=1}^{12} \frac{y_{jk}}{n_{jk}} \times n_{ijk} \tag{1.1}$$

The spatial regression model used in this study is mathematically defined as

$$log(\lambda_i) = \alpha + \beta x_i + u_i + \eta_i \tag{1.2}$$

where α and β are mean ratio (intercept) and regression coefficients, and η_i and u_i are the spatially correlated and uncorrelated random effect respectively [26–28]. The u_i, $u_i \sim N(0, \sigma_u^2)$, is the unstructured heterogeneity, "noise," which denotes the variation unaccounted for in this study. The η_i is the structured heterogeneity described by a Gaussian conditional autoregressive (CAR) distribution which is defined as

$$\eta_i | \eta_{j,\ i \neq j}, \tau_\eta \sim N\left(\frac{1}{\mathcal{N}(i)} \sum_{j \sim i}^{n} w_{ij} \eta_j, \ \frac{\sigma_\eta^2}{\mathcal{N}(i)}\right) \tag{1.3}$$

The $\mathcal{N}(i), i = 1, \ldots, 96$, is the set of neighbors(s) to a specific RHAD which can be defined in different ways such as proximity or boundary sharing between regions [31]. The neighborhood as used in this study is defined as two regions sharing boundaries, and w_{ij} represent the relationship between any two regions (i,j), which is defined as

$$w_{ij} = \begin{cases} 1, & i\ and\ j\ are\ adjacents \\ 0, & otherwise \end{cases}$$

and σ_η^2 is the spatial dispersion of region i. In other to account for the temporal effect in the spatial model, Eq (1.2) was extended as shown in Eq (1.4) below:

$$log(\lambda_{it}) = \alpha + \beta x_i + u_i + \eta_i + \gamma_t + \varphi_t \tag{1.4}$$

where the additional two components (γ_t, φ_t) represent the unstructured and structured temporal effects. The structured temporal effect, φ_t, is modeled by imposing a random walk of order one, RW (1), a distribution which is defined as a step function given as

$$\varphi_t | \varphi_{-t} \sim \begin{cases} Normal\ (\varphi_{t+1}, \sigma_\varphi^2), & for\ t = 1 \\ Normal\ \left(\dfrac{\varphi_{t-1} + \varphi_{t+1}}{2}, \dfrac{\sigma_\varphi^2}{2}\right), & for\ t = 2, \ldots, T-1 \\ Normal\ (\varphi_{t-1}, \sigma_\varphi^2), & for\ t = T \end{cases}$$

and an exchangeable prior $\gamma_t \sim (0, \sigma_\gamma^2)$ is imposed on the unstructured temporal effect.
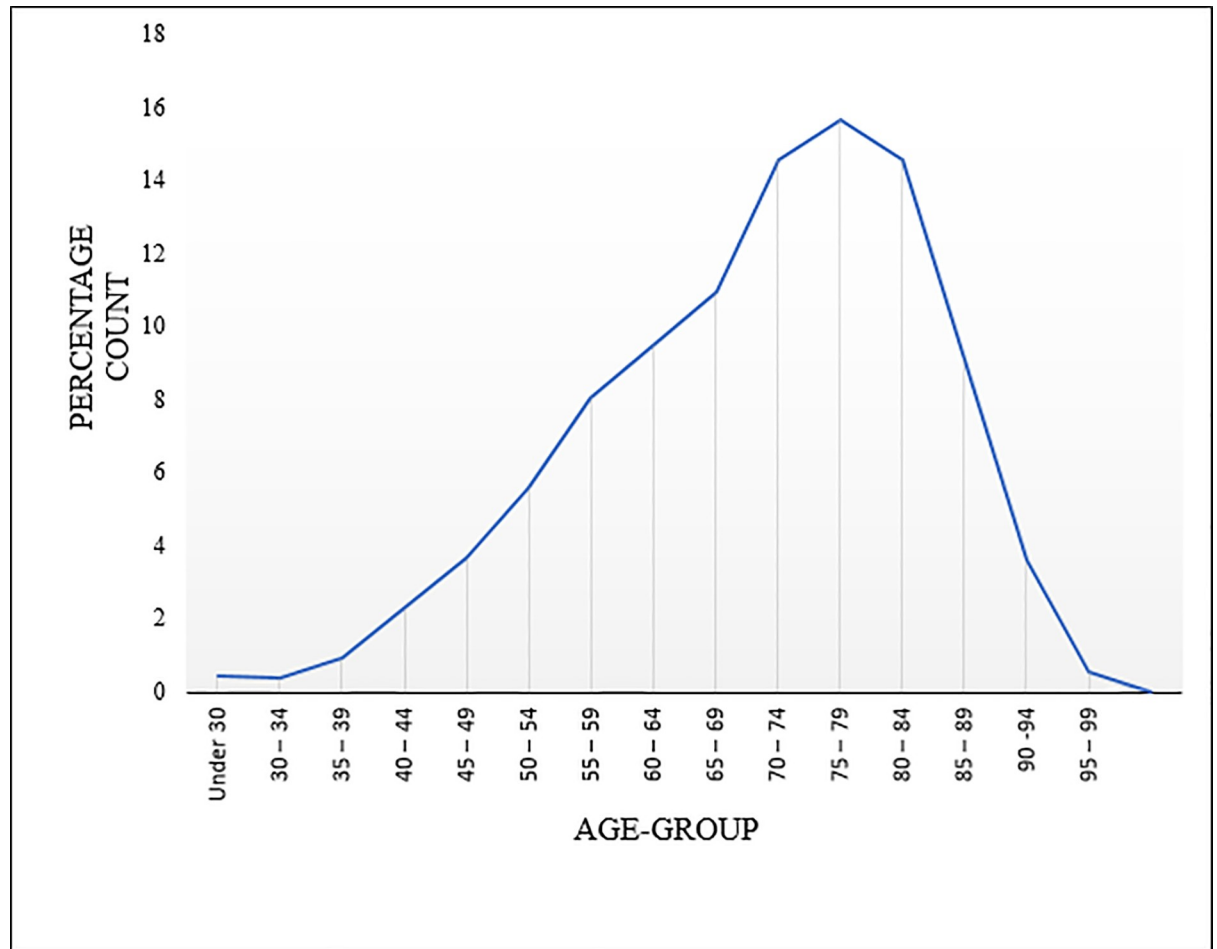
Fig 2. GC distribution by age group.

The number of parameters and hyperparameters to be estimated are $\theta = (\alpha, \beta)$ and $\psi = (\sigma_u^2, \sigma_\gamma^2, \sigma_\varphi^2, \sigma_\eta^2)$ respectively. The model fitting was implemented using Integrated Nested Laplace Approximation (INLA) in R programming language. We imposed different minimally informative prior on the log of the structured effect precision to examine the effect of prior on the parameter estimates. In summary, we imposed an intrinsic conditional autoregressive (ICAR) to account for the spatial dependency in both spatial and spatio-temporal model, using a minimally informative uniform prior distribution for the intercept $\alpha \sim U(-\infty, +\infty)$, a normal prior distribution for the slopes $\beta \sim N(0, 1000)$, and a log Gamma prior for both spatially structured and unstructured precision log $\tau \sim \log Gamma\ (0.1, 0.01)$. We present the incidence risk ratios (IRR) and the corresponding 95% credible interval for the model parameters.

Results and discussion

Results

A total of 3,172 cases of GC were diagnosed between 1992 and 2016 in Manitoba. Approximately 31% of the GC cases were cardia gastric cancer (CGC), 49% were non-cardia gastric cancer (NCGC) and 20% were unspecified. The highest number of GC (approximately 16%) occurred in the group 75–79 as shown in Fig 2.
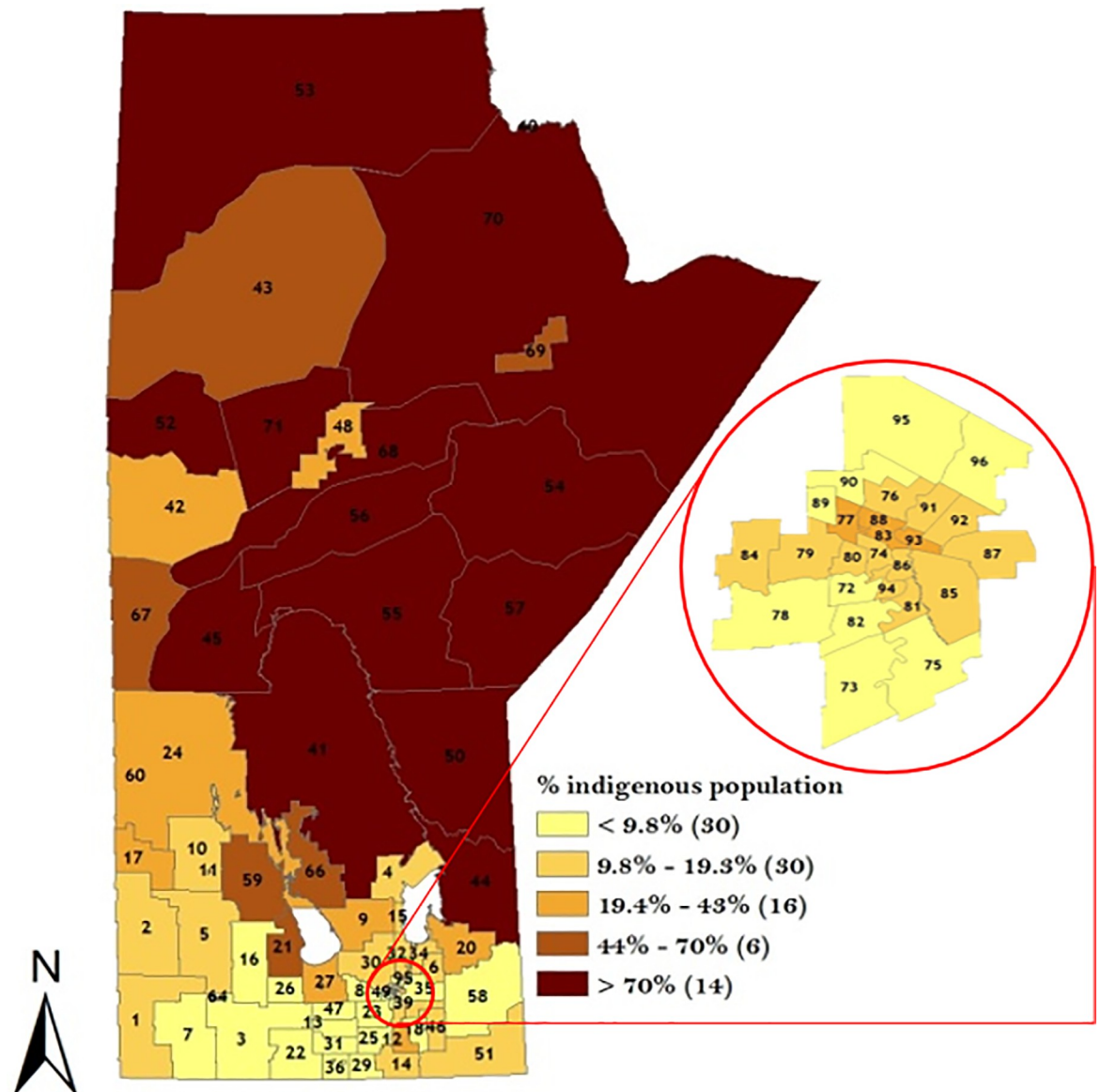
Fig 3. Geographical distribution of Indigenous population across the 96 RHADs based on 2016 Canadian Census data. Numbers in the map are the district labels from 1 to 96. Values in parenthesis represents the counts in each distribution category.

A descriptive analysis of the potential risk factors is presented in Figs 3–5. The proportion of Indigenous people was highest in the north-western part of the Manitoba (Fig 3) while the proportion of immigrants was highest in the Winnipeg Regional Health Authority (WRHA) (Fig 4). Also, the SESI revealed that majority of the districts in the northern Manitoba had the lowest SESI, which translates to having a low socio-economic status, while WRHA and the majority of the southern part of Manitoba had a high SESI (Fig 5).

We first fitted a multivariable Poisson regression model to examine the spatial dependency of overall GC IRR. This was done by examining the unexplained variation in GC IRR as shown in Fig 6. The result revealed the presence of geographical dependency of GC which is evident based on the variation in color gradient in the map. Region(s) with lighter color
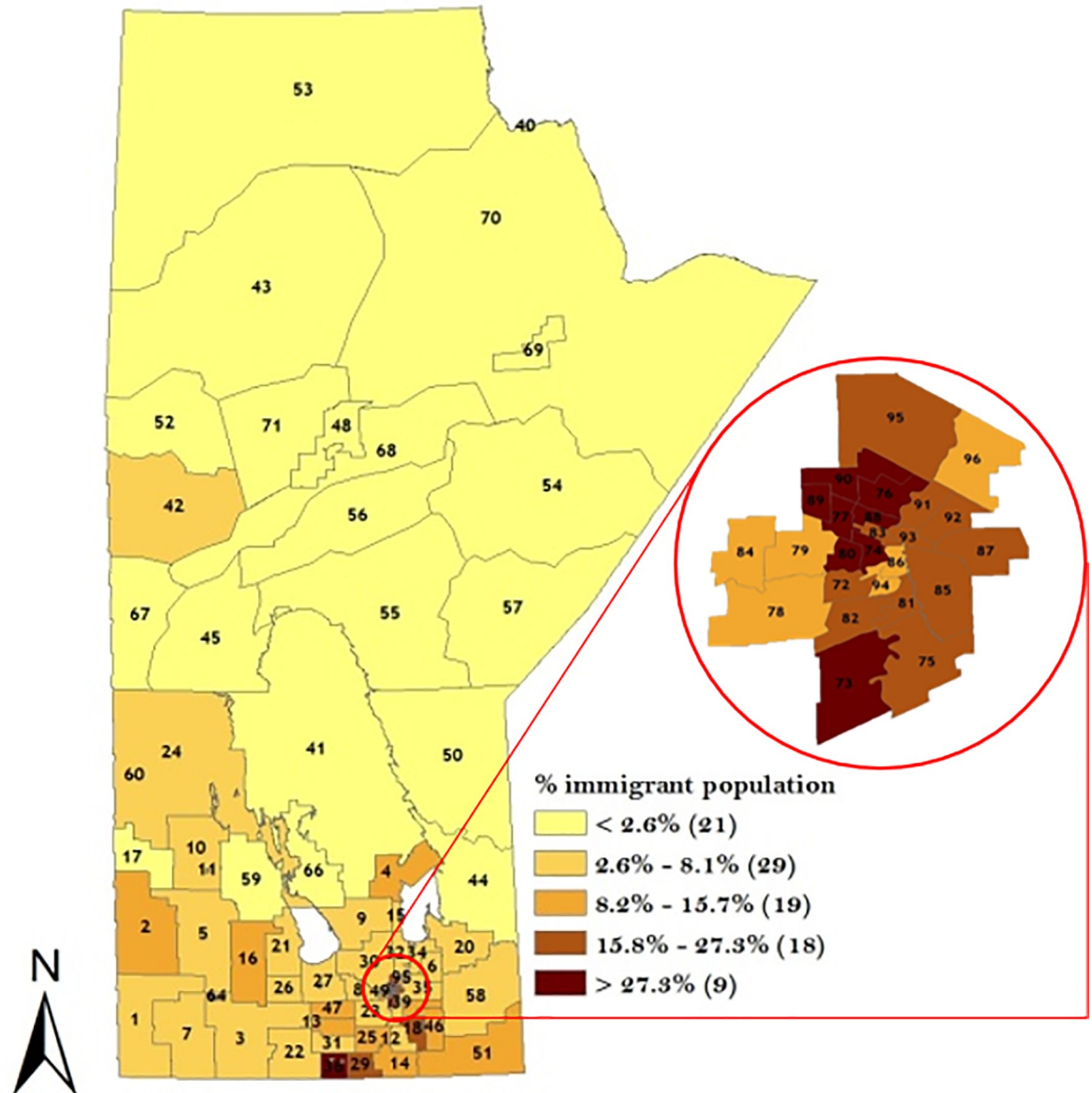
Fig 4. Geographical distribution of immigrant population across the 96 RHADs based on 2016 Canadian Census data. Numbers in the map are the district labels from 1 to 96. Values in parenthesis represents the counts in each distribution category.

indicates low GC IRR, while darker color indicate regions with high GC IRR as used in this study for all maps.

This result was followed by a confirmatory test using Moran's I statistics defined as

$$I = \frac{n}{S_0} \left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \right) \tag{1.5}$$

where y_i represents the count of GC cases in region I; $\bar{y}$ represents the average count of GC cases; w_{ij} represents the distance weight matrix obtained assuming a neighborhood structure where regions share a direct boundary with one another; n is the number of regions (n = 96); $S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$ is the aggregate weight. The Moran's I statistic confirmed the existence of
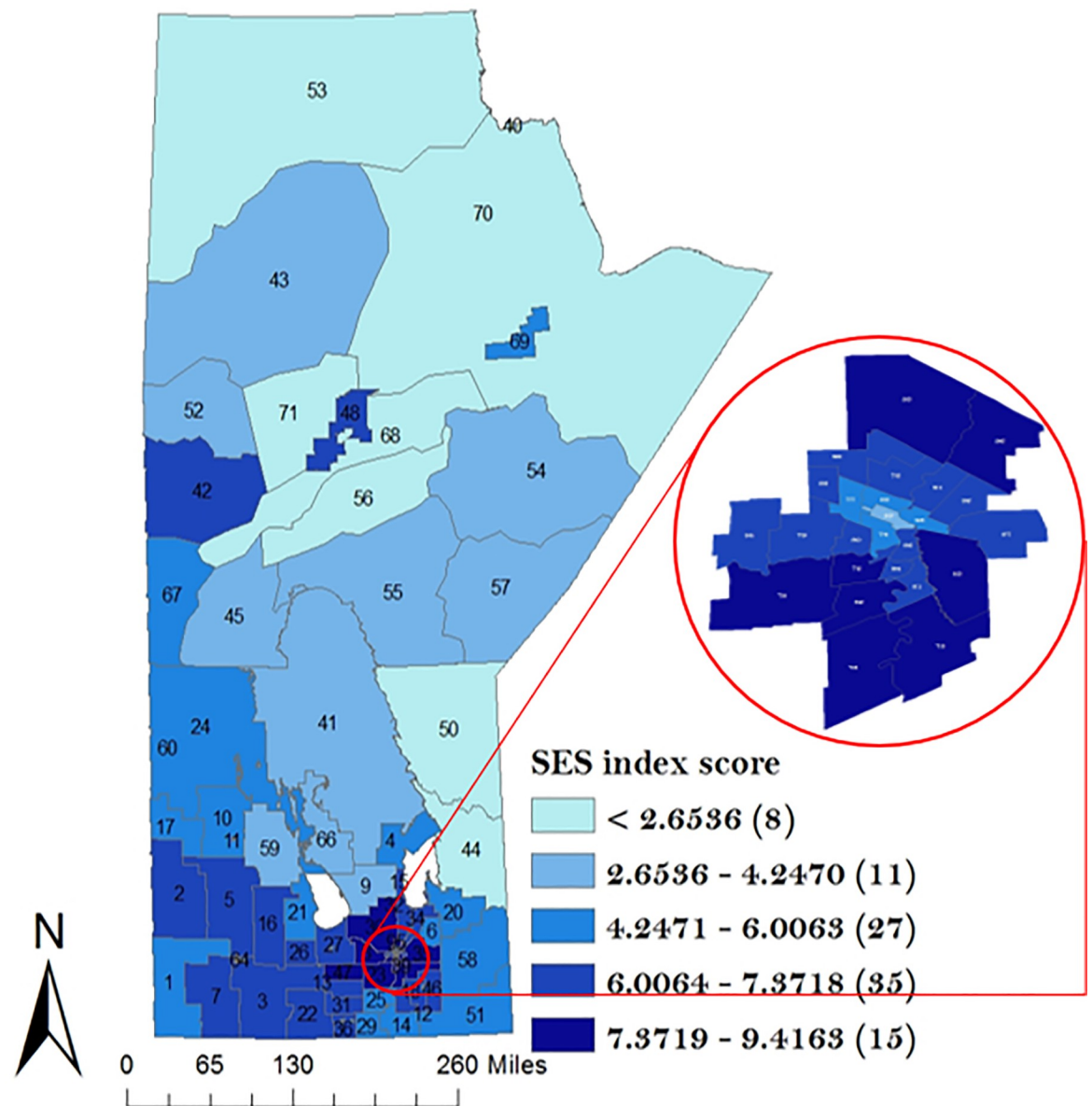
Fig 5. Geographical distribution of socio-economic status score index across the 96 RHADs based on 2016 Canadian Census data.
Numbers in the map are the district labels from 1 to 96. Values in parenthesis represents the counts in each distribution category.

spatial dependency (clusters) (0.1563; p-value = 0.007). Tables 2 and 3 present the effects of the risk factors on our dataset partitioned by sex and topographical sub-sites respectively.

From Table 2, it is evident that the covariates did not significantly predict the risk of GC for the overall GC dataset and the sex-stratified GC dataset. Table 3 presents the results for the topographical sub-site partitioning into CGC and NCGC. Unlike the overall result, SESI was a significant factor associated with CGC. A unit increase in SESI decreased the risk of CGC by 14%, and the risk of NCGC by 10%. 1% increase in regional Indigenous population proportion reduced the risk of CGC by 1.4% which was marginally significant, while the risk of NCGC among the Indigenous was not significant.
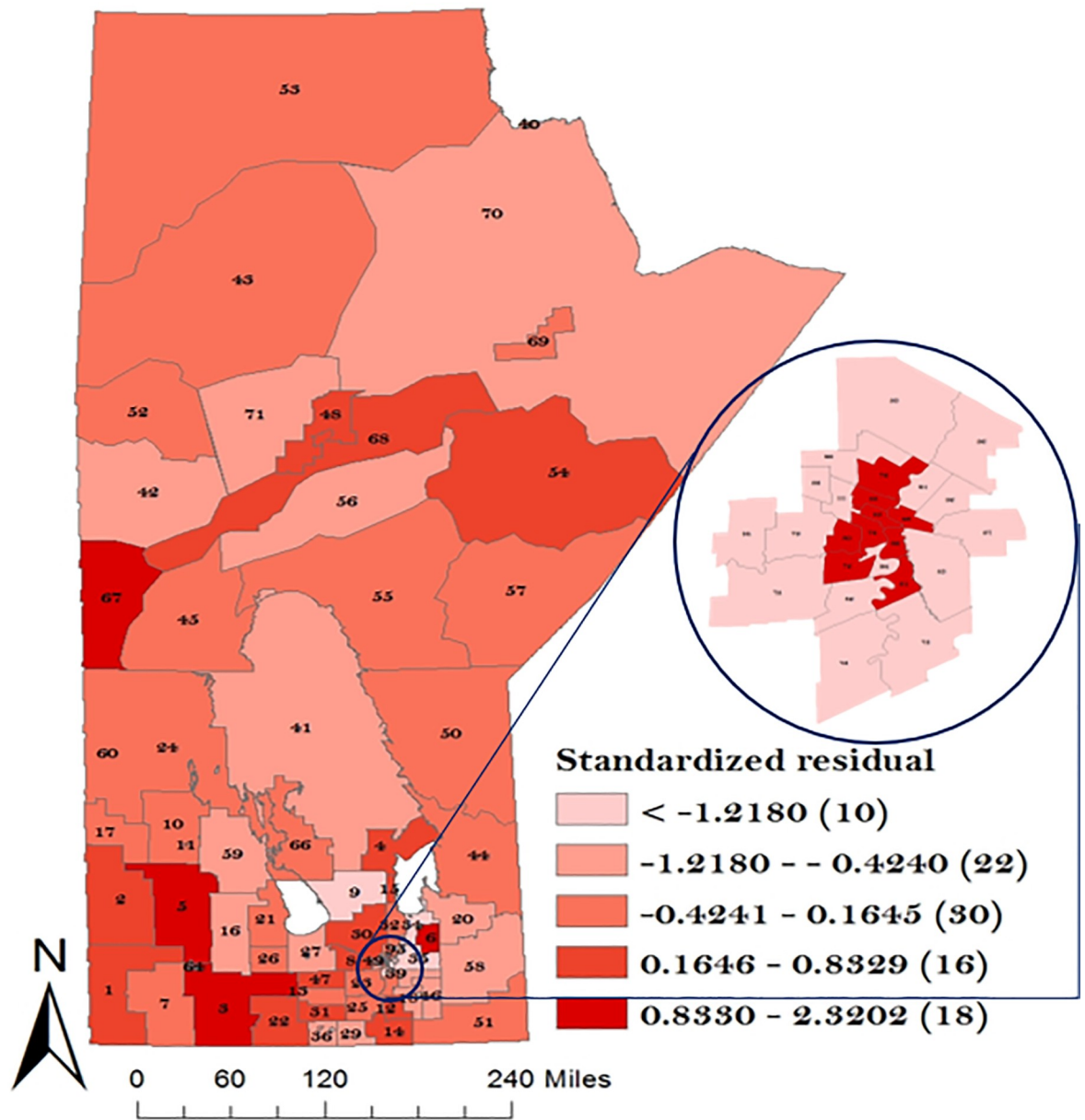
Fig 6. Map of unexplained variation in overall GC incidence risk ratio. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legend.

Table 2. Incidence Risk Ratio (IRR) and 95% credible interval for overall, male and female GC dataset using spatial Poisson regression model.

Parameter	Overall population		Male population		Female population	
	IRR	95% Credible Interval (Lower, Upper)	IRR	95% Credible Interval (Lower, Upper)	IRR	95% Credible Interval (Lower, Upper)
SESI	0.921	(0.836, 1.013)	0.916	(0.833, 1.008)	0.906	(0.814, 1.010)
Immigrant	1.003	(0.992, 1.014)	1.004	(0.993, 1.014)	1.003	(0.992, 1.014)
Indigenous	1.001	(0.994, 1.008)	1.000	(0.992, 1.007)	1.004	(0.961, 1.012)

Table 3. Incidence Risk Ratio (IRR) and 95% credible interval for cardia and non-cardia GC dataset using spatial Poisson regression model.

Parameter	Cardia GC		Non-cardia GC	
	IRR	95% Credible Interval (Lower, Upper)	IRR	95% Credible Interval (Lower, Upper)
SESI	0.859	(0.780, 0.947)	0.898	(0.812, 0.995)
Immigrant	0.994	(0.983, 1.003)	1.002	(0.990, 1.014)
Indigenous	0.986	(0.978, 0.994)	1.000	(0.990, 1.007)

Table 4. Incidence Risk Ratio (IRR) and 95% credible interval for cardia GC dataset stratified by sex using spatial Poisson regression model.

Parameter	Male		Female	
	IRR	95% Credible Interval (Lower, Upper)	IRR	95% Credible Interval (Lower, Upper)
SESI	0.930	(0.827, 1.012)	0.738	(0.618, 0.879)
Immigrant	1.003	(0.986, 1.006)	0.994	(0.979, 1.009)
Indigenous	1.010	(0.982, 1.049)	0.981	(0.966, 0.996)

Table 5. Incidence Risk Ratio (IRR) and 95% credible interval for Non-cardia GC dataset stratified by sex using spatial Poisson regression model.

Parameter	Male		Female	
	IRR	95% Credible Interval (Lower, Upper)	IRR	95% Credible Interval (Lower, Upper)
SESI	0.910	(0.811, 1.022)	0.922	(0.819, 1.040)
Immigrant	1.010	(0.997, 1.022)	1.001	(0.987, 1.014)
Indigenous	1.001	(0.992, 1.010)	1.004	(0.995, 1.013)

Further stratification of CGC by sex, as shown in Table 4, revealed no significant impact of SESI, immigrant and Indigenous variables on the risk of CGC for the men population. A different result for the female population was observed, where a unit increase in district SES score decreased the risk of CGC among women by a significant 26.2%, and the risk of CGC among Indigenous women population was reduced by 1.9% with 1% increase in regional Indigenous population proportion. The result of the non-cardia GC sex-stratified spatial model displayed in Table 5 showed no significant impact of the covariates on the risk of NCGC.

In order to understand the district-specific risk of GC stratified by sex and topographical sub-site, the smoothed spatial random effect from the model was plotted in Figs 7–10. The result of the spatial variation of GC stratified by sex for the male population is displayed in Fig 7 which identified 22 districts with GC incidence risk ratio greater than the rest of the population. Note that 7 RHADs out of the 22 RHADs (i.e., 32%) had a high GC incidence risk ratio relative to the rest of the population which were located in WRHA. Five districts (Winnipeg Churchill in the northern RHA, Brandon East end and Dauphin in Prairie Mountain RHA, and downtown East & Point Douglas North in Winnipeg RHA) had the highest risk ratios (between 1.70 and 3.65).

For the women GC sub-population, a total of 16 RHADs were identified with a high GC incidence risk ratio relative to the rest of the population. Note that 10 out of the 16 RHADs (i.e., 63%) were located in WRHA. One RHAD (district 06 in Interlake RHA) was identified with the highest GC incidence risk ratio as shown in Fig 8. The result of CGC spatial variation in Fig 9 identified 11 RHADs with a high CGC incidence risk ratio relative to the rest of the population. Note that 6 out of the 11 RHADs (i.e., 55%) were located in WRHA. The result of the NCGC in Fig 10 identified 27 RHADs with a higher risk of NCGC relative to the rest of the
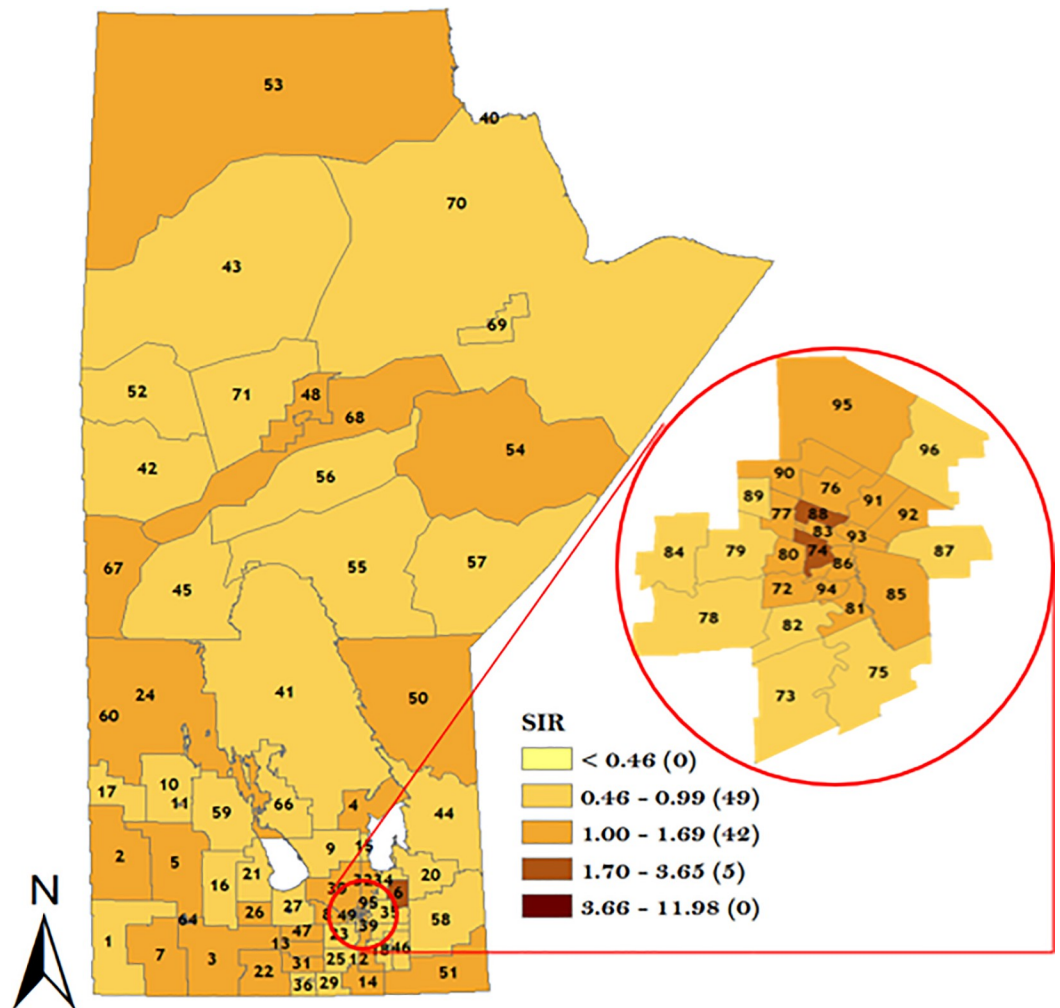
Fig 7. Map of standardized IRR of overall GC in Manitoba men using spatial Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

population. Note that 11 out of the 27 RHADs (i.e., 41%) with a high NCGC incidence risk ratio was located in WRHA. Similar to CGC, 4 RHADs (Souris River in Prairie Mountain RHA, Winnipeg Churchill in northern RHA, Rural East in Southern Health RHA, and Point Douglas South in Winnipeg RHA) were identified with the highest NCGC incidence risk ratios.

Similarly, we also considered district-specific risk over time for each data strata, by mapping the spatio-temporal Poisson regression model (Figs 11–14). The results shown in Figs 11–14 revealed the presence of variation in risk of GC, CGC, and NCGC over time for each district. This is evident based on the variation in color gradient across the RHADs over time. More explanations about time trends of GC, CGC, and NCGC will be provided in the Discussion section.

Discussion

In general, using the ecological regression model on the stratified dataset of GC, a significant association between SESI and CGC, and a marginally significant association between
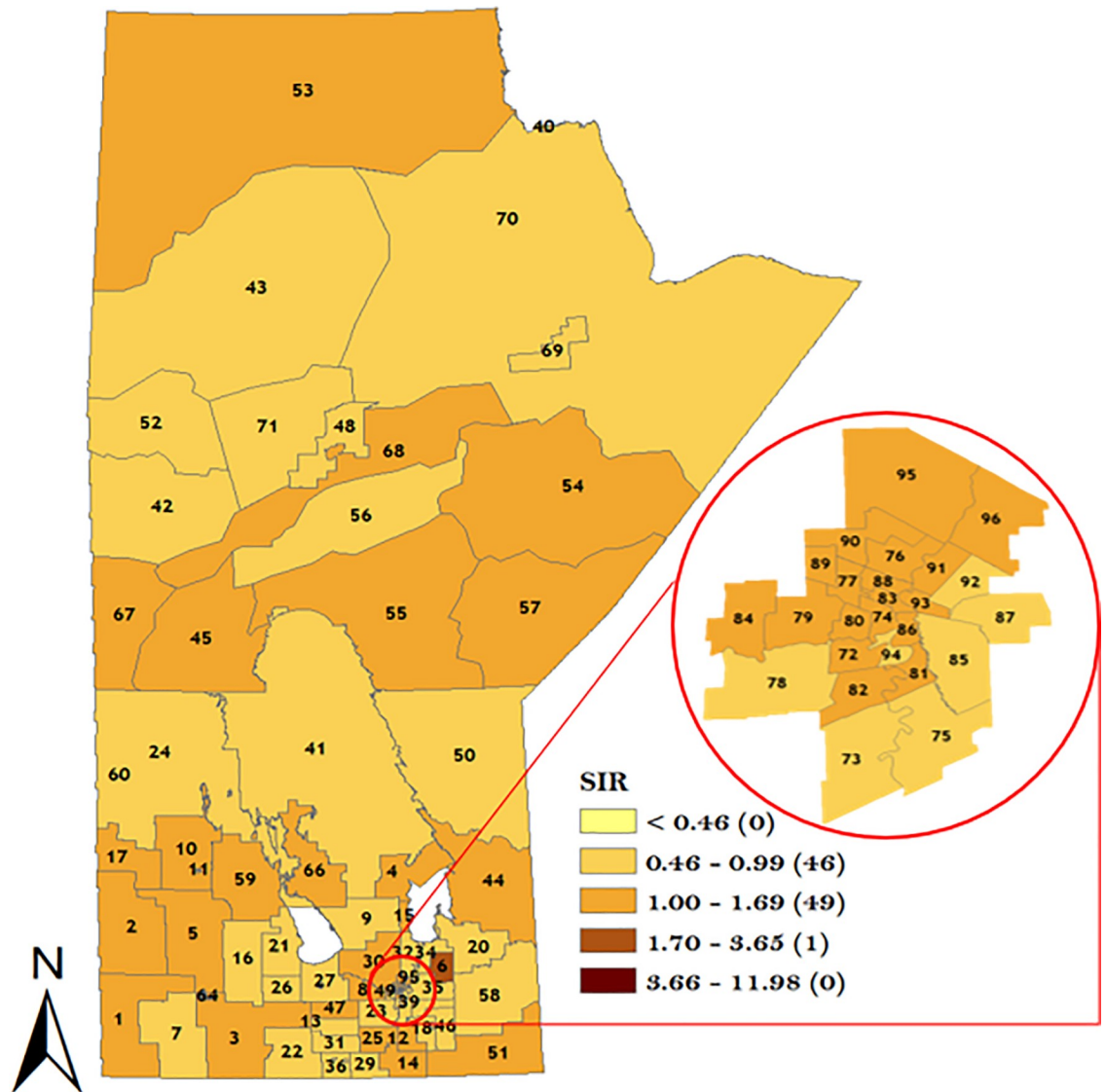
Fig 8. Map of standardized IRR of overall GC in Manitoba women using spatial Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

https://doi.org/10.1371/journal.pone.0253650.g008

Indigenous population and CGC was observed, where both covariates decreased the risk of CGC. This result is partly supported by literature as some studies also suggest a decrease in the risk of GC among both sexes in Indigenous populations in other parts of the world [32]. Despite this, one might have predicted that Indigenous population proportion would be associated with a higher GC risk due to the fact that many of the Indigenous people in the study region have lower socioeconomic status and live in more remote regions, which could impede their access to good healthcare. We also observed a marginally significant association between SESI and NCGC.

A closer look at the district-specific risk identified five districts exhibiting high risk across all data sub-groups. These districts include Little Saskatchewan and Brandon East End in Prairie Mountain RHA and Point Douglas North, Downtown East, and St. Boniface West in WRHA. These districts were also associated with high percentage of no post-secondary school
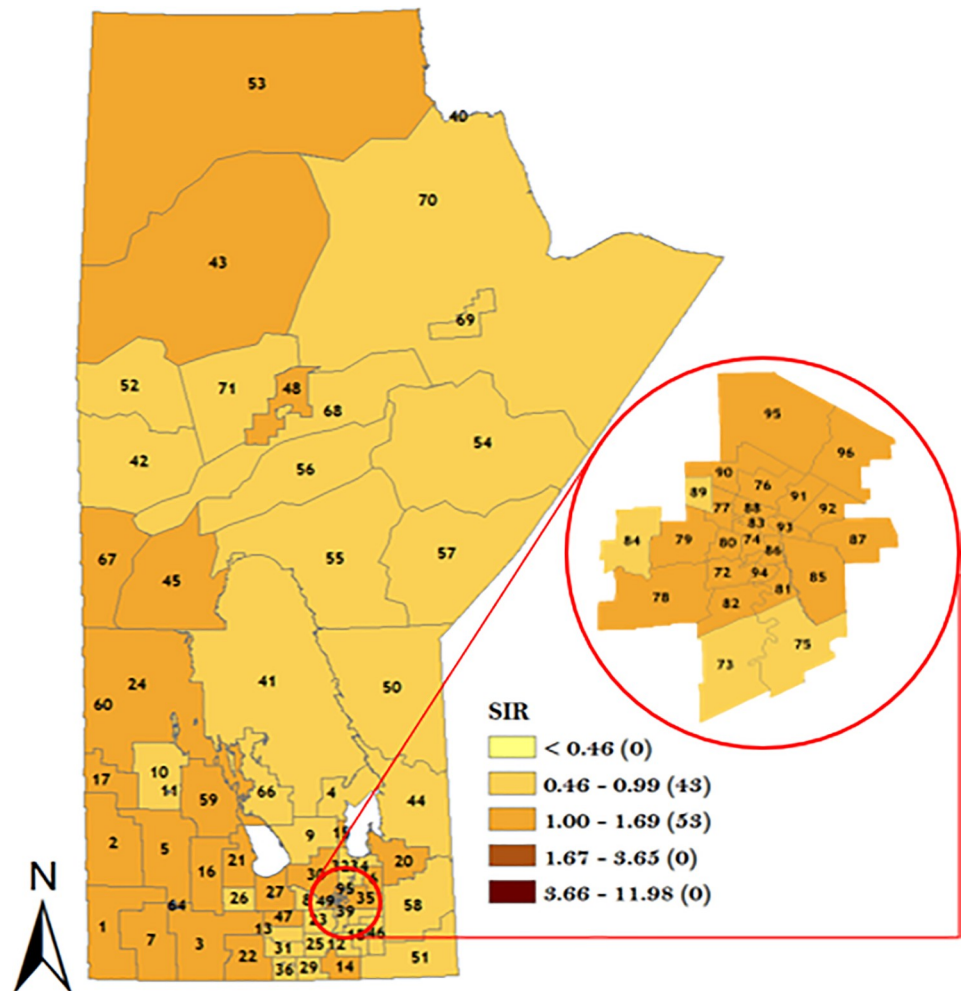
Fig 9. Map of standardized IRR of overall CGC in Manitoba using Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

https://doi.org/10.1371/journal.pone.0253650.g009

education (35.7% - 48.8%), unemployment (6.4% - 11.1%), a low income ($21,841 - $43, 648), and moderate proportion of immigrants (27.3%).

The spatio-temporal result presented in Figs 11–14 demonstrated apparent changes in the risk of GC across the RHADs over time. The study of the variation in risk of GC for each RHAD overtime gave us the opportunity to investigate regions with decreasing risk, increasing risk, and steady risk over time. Our findings support the fact that needs in respect to GC eradication across the province are not uniform. Therefore, policies regarding resource allocation for the detection and treatment of GC should be based on level of GC risk.

Until now, the epidemiological study of GC both in space and time in Manitoba was not well documented. Hence, this study was the first of its kind to examine the geographical variation of GC in Manitoba. This study offers a detailed glimpse into spatial epidemiology of GC in Manitoba. It will enable public health agencies (especially those with a keen interest in cancer surveillance) such as CancerCare Manitoba and Manitoba Health to have a clear picture of the spatial variation of GC in Manitoba. This study will then help policy makers determine resource allocation and plan for possible prevention/intervention. In addition to using the
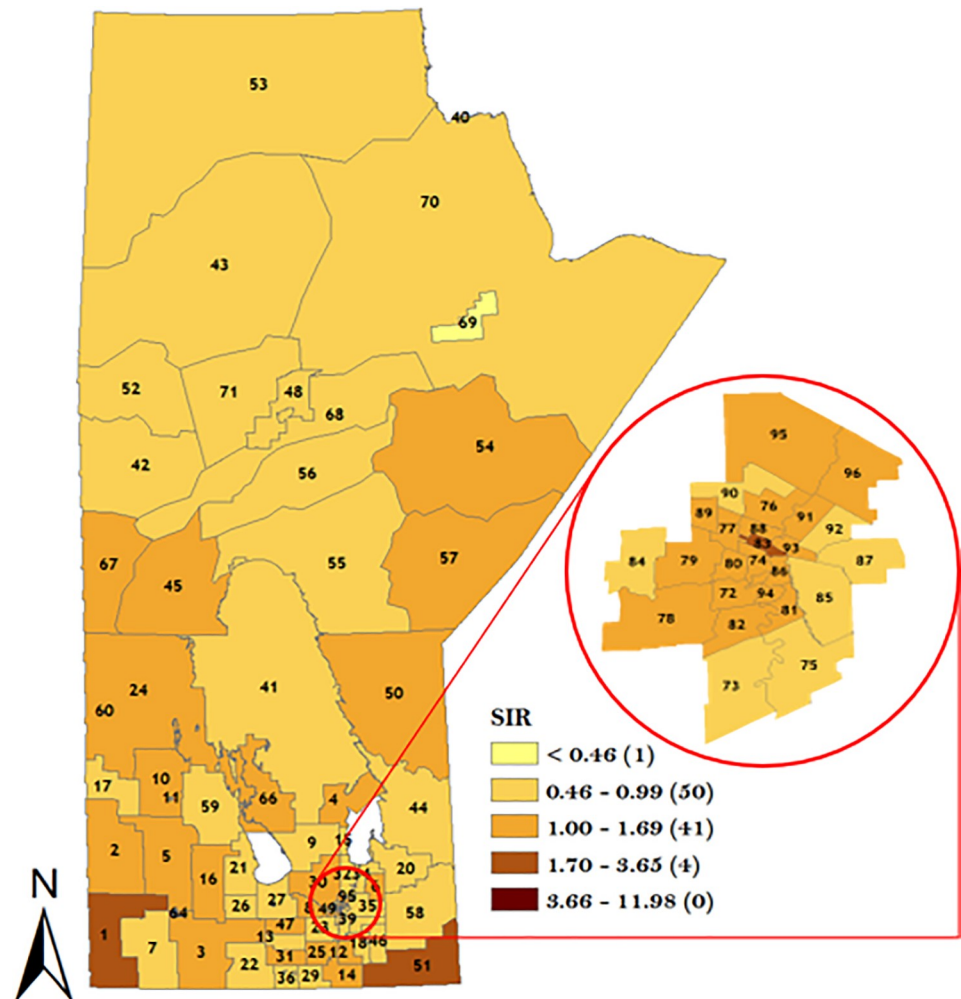
Fig 10. Map of standardized IRR of overall NCGC in Manitoba using Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

findings in this study for a logical resource allocation program, it will also serve as a GC diseases surveillance and monitoring tool for interested organizations.

Strengths and limitations

One of the main strengths of this study is that it is based on population data and therefore not subject to selection bias. The spatial and spatio-temporal methods used in this study leveraged on districts neighborhood to obtain a reliable risk estimate. The inclusion of time random effect also helped us to acknowledge and account for time effect, thereby avoiding over-estimation. The time effect also helped us investigate the trend of the GC over time, which is a desired tool in disease trend surveillance. The main advantage of using spatio-temporal regression model is to capture all complex variations (risk factors and spatio-temporal variation) unlike the other models such as e.g. geographically weighted regression. One could also use cluster detection methods such as SaTScan [33] which is a simulation-base method to identify clusters with high risk without properly assuming necessary assumptions on the nature of outcome. Also, the cluster detection methods are not able to simultaneously identify possible risk factors
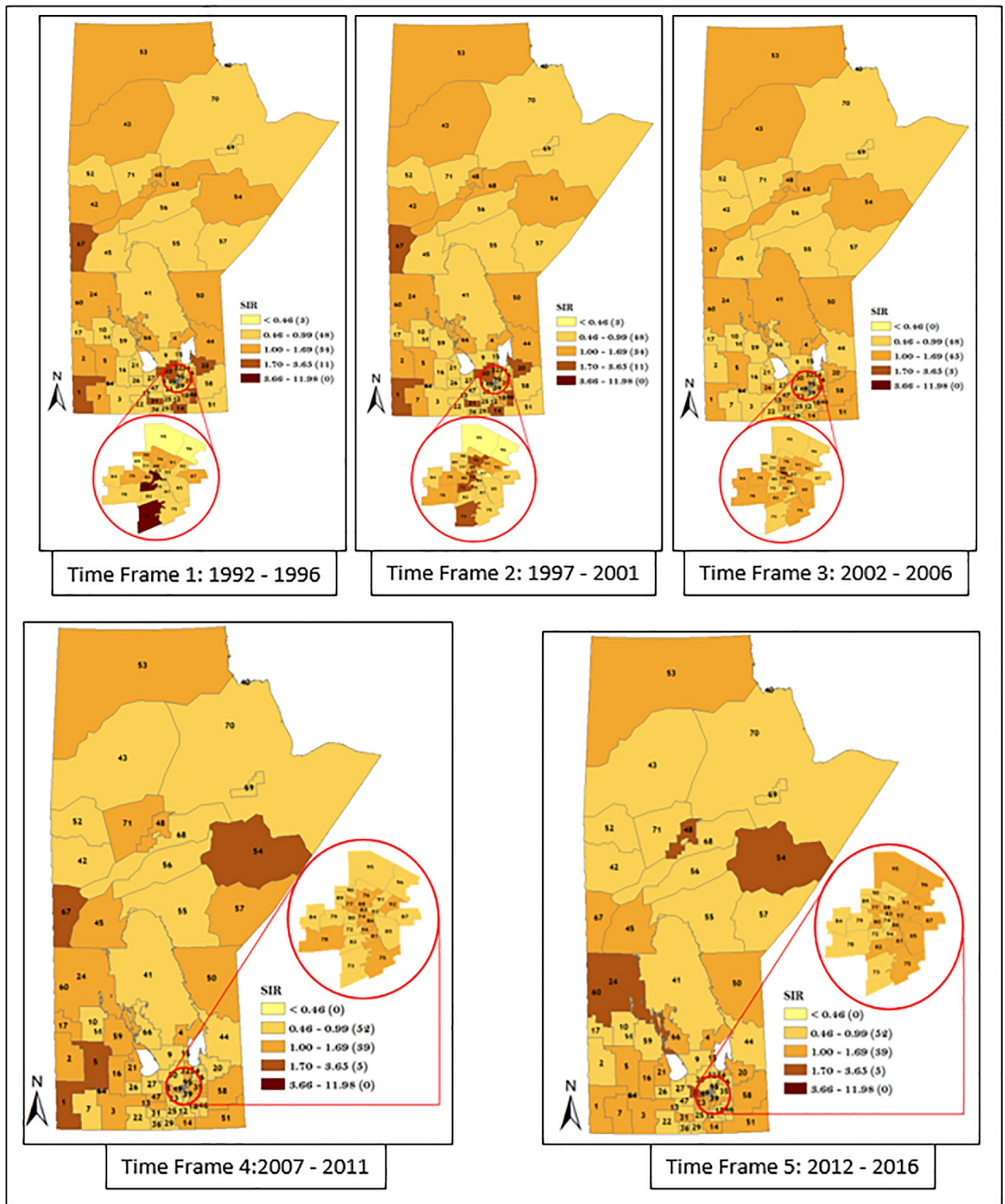
Fig 11. Maps of district-specific risk of overall men GC for the 96 RHADs in Manitoba for five time periods using spatio-temporal Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.
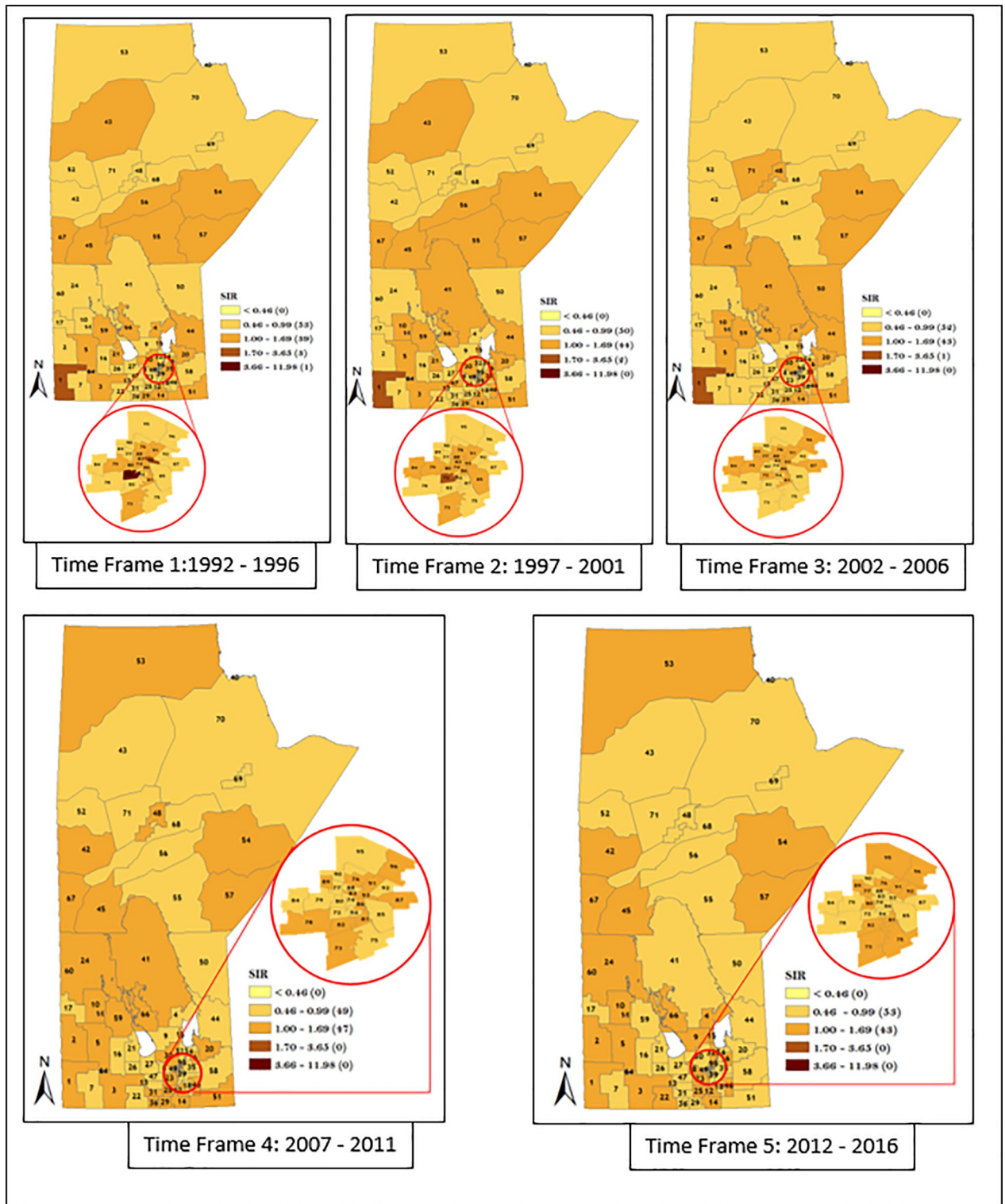
Fig 12. Maps of district-specific risk of overall women GC for 96 RHADs in Manitoba for five time periods, using spatio-temporal Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

https://doi.org/10.1371/journal.pone.0253650.g012
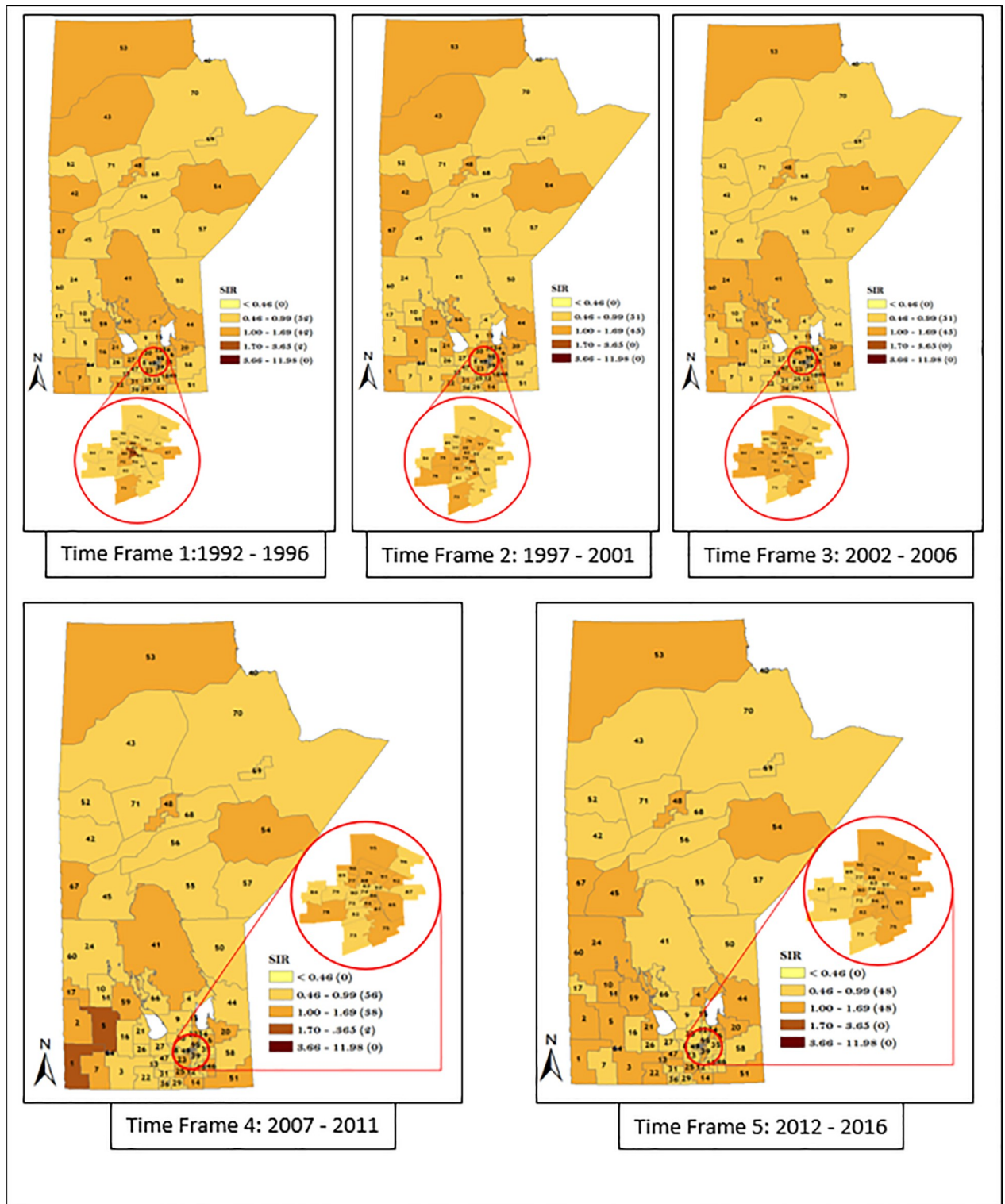
Fig 13. Maps of district-specific risk of CGC for 96 RHADs in Manitoba for five time periods using spatio-temporal Poisson regression model.
Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.

Fig 14. Maps of district-specific risk of NCGC for 96 RHADs in Manitoba for five time periods using spatio-temporal Poisson regression model. Numbers in the map are the district labels from 1 to 96. Number of districts in each risk category is shown in parenthesis in the legends.
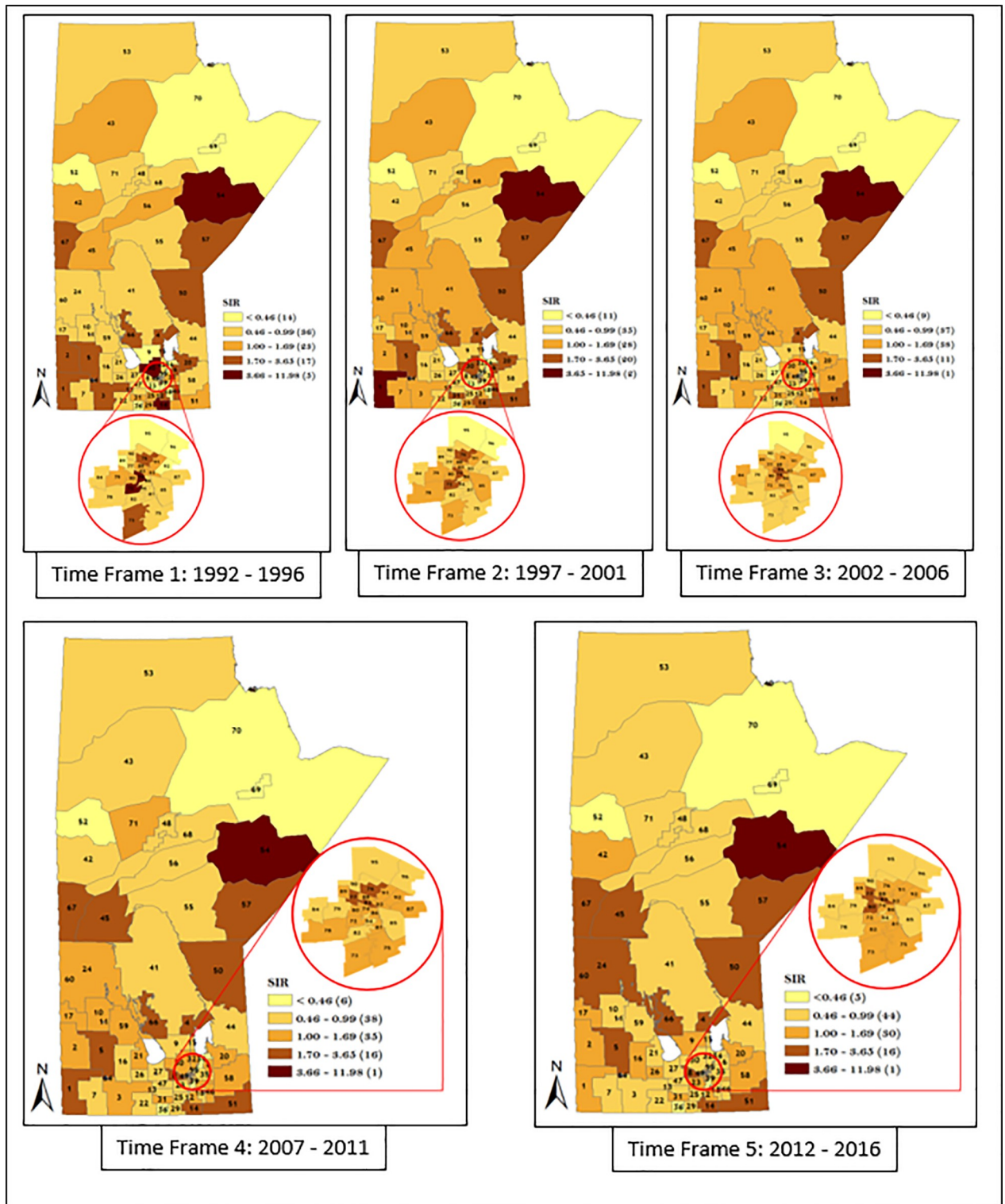
of the disease (e.g., GC) unlike our sophisticated spatio-temporal regression model. However, this study is limited by the fact that we were unable to adjust for several other factors such as smoking, obesity, lifestyle, Helicobacter pylori, and food, which are important risk factors associated with the etiology of GC. Some of factors used in this work such as Indigenous and Immigrant population may be used as proxy for covariates such as lifestyle and food.

Conclusions

In conclusion, our study has demonstrated that the risk of GC varied across the province of Manitoba.

Not only was the incidence risk of GC found to differ across small regions within the province, several regions exhibited different progressive risk (decreasing, increasing or steady) over time. While the overall incidence risk of GC at the provincial level was steady over time, this was not true for all RHADs. More studies to investigate the driving force behind the progressive increase in risk of GC for the identified regions are highly recommended. This will further enable policy makers to address the issue more efficiently as each district may require a different approach to eradicating GC risk.

Finally, we demonstrated the use of spatial and spatio-temporal regression model as a robust model for investigating variation in space and time. The model leverages on both regional and temporal dependency to arrive at a more reliable estimate. This modeling approach is highly recommended in health study where the dataset is collected over space and time.

Acknowledgments

Author Contributions

Conceptualization: Oluwagbenga Fakanye, Harminder Singh, Danielle Desautels, Mahmoud Torabi.

Data curation: Oluwagbenga Fakanye.

Formal analysis: Oluwagbenga Fakanye.

Funding acquisition: Mahmoud Torabi.

Investigation: Mahmoud Torabi.

Methodology: Oluwagbenga Fakanye, Mahmoud Torabi.

Project administration: Mahmoud Torabi.

Resources: Mahmoud Torabi.

Software: Oluwagbenga Fakanye.

Supervision: Harminder Singh, Danielle Desautels, Mahmoud Torabi.

Validation: Oluwagbenga Fakanye, Harminder Singh, Danielle Desautels, Mahmoud Torabi.

Visualization: Oluwagbenga Fakanye, Mahmoud Torabi.

Writing – original draft: Oluwagbenga Fakanye.

Writing – review & editing: Oluwagbenga Fakanye, Harminder Singh, Danielle Desautels, Mahmoud Torabi.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLO-BOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [published online ahead of print 2018]. CA Cancer J Clin. https://doi.org/10.3322/caac.21492.

2. Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2018. Toronto, ON: Canadian Cancer Society; 2018. Available at: cancer.ca/Canadian-Cancer-Statistics-2018-EN (accessed December 23, 2018).

3. American Cancer Society: Cancer Facts and Figures 2019. Atlanta, Ga: American Cancer Society, 2019. Available online at https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2019/cancer-facts-and-figures-2019.pdf. Accessed May 05, 2019.

4. Mahar A. L., Coburn N. G., Kagedan D. J., Viola R., Johnson A. P. (2016). Regional variation in the management of metastatic gastric cancer in Ontario. *Current Oncology*, 23(4), 250. https://doi.org/10.3747/co.23.3123 PMID: 27536175

5. Sharafi Z., Asmarian N., Hoorang S., Mousavi A. (2018). Bayesian spatio-temporal analysis of stomach cancer incidence in Iran, 2003–2010. *Stochastic environmental research and risk assessment*, 32(10), 2943–2950.

6. Colquhoun A., Hannah H., Corriveau A., Hanley B., Yuan Y., Goodman K. J., CANHelp Working Group. (2019). Gastric Cancer in Northern Canadian Populations: A Focus on Cardia and Non-Cardia Subsites. *Cancers*, 11(4), 534. https://doi.org/10.3390/cancers11040534 PMID: 30991639

7. van Loon AJ, Goldbohm RA, van den Brandt PA (1998). Socioeconomic status and stomach cancer incidence in men: results from The Netherlands Cohort Study. *J Epidemiol Community Health* 52,166–71. https://doi.org/10.1136/jech.52.3.166 PMID: 9616421

8. John M. Bumsted, Peter McLintock and Others: Manitoba. Encyclopedia Britannica, Inc, 2019. Available at https://www.britannica.com/place/Manitoba Accessed on May 13, 2019.

9. Fransoo R., Martens P., The Need to Know Team, Burland, E, Prior, H., Burchill, C., Chateau, D., Walld, R., Sex Differences in Health Status, Health Care Use, and Quality of Care: A Population-Based Analysis for Manitoba's Regional Health Authorities. Winnipeg, Manitoba Centre for Health Policy, November, 2005.

10. Torabi M., & Rosychuk R.J. (2012). "Hierarchical Bayesian Spatiotemporal Analysis of Childhood Cancer Trends." *Geographical Analysis*, (44(2), 109–120.

11. ESRI, 2014. ArcGIS for Desktop, version 10.2.2. Environmental Systems Research Institute, Redlands, CA, USA. (http://www.esri.com/).

12. Fritz P, Percy C, Jack A, Shanmugaratnuers K, Solin L, Parkin D: International classification of diseases for oncology. 3 edition. Geneva: World Health Organization; 2000.

13. Statistics Canada. 2019. Canadian Cancer registry. https://www23-statcan-gc-ca.uml.idm.oclc.org/imdb/p2SV.pl?Function=getSurvey&Id=1211637 (accessed on December 18).

14. Nagel G., Linseisen J., Boshuizen H. C., Pera G., Del Giudice G., Westert G. P., et al. (2007). Socioeconomic position and the risk of gastric and oesophageal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST). *International journal of epidemiology*, 36(1), 66–76. https://doi.org/10.1093/ije/dyl275 PMID: 17227779

15. Wu C. C., Hsu T. W., Chang C. M., Yu C. H., Wang Y. F., & Lee C. C. (2014). The effect of individual and neighborhood socioeconomic status on gastric cancer survival. *PloS one*, 9(2), e89655. https://doi.org/10.1371/journal.pone.0089655 PMID: 24586941

16. Dong E., Duan L., & Wu B. U. (2017). Racial and ethnic minorities at increased risk for gastric cancer in a regional US population study. *Clinical Gastroenterology and Hepatology*, 15(4), 511–517. https://doi.org/10.1016/j.cgh.2016.11.033 PMID: 27939654

17. Holt J, Lo C (2008). The geography of mortality in the Atlanta metropolitan area. *Computers*, *Environment and Urban Systems*, 32(2), 149–164.

18. Krishnan V (2010). Constructing an area-based socioeconomic status index: A principal component analysis approach. *Early Childhood Council Annual Conference 2010, 2010*, May 7, Christchurch, New Zealand.

19. Vincent Kyle, Sutherland Jason M. "A review of methods for deriving an index for socioeconomic status in British Columbia." Vancouver, Canada: University of British Columbia, Centre for Health Services

and Policy Research. Available at: http://healthcarefunding.ca/files/2013/04/Review-of-Methods-for-SES-Index-for-BC.pdf (2013).

20. Abdul Rahman N., Abd Naeeim N. S. (2018). Construction of a Socio-Economic Status (SES) Index in Peninsular Malaysia Using the Factor Analysis Approach. Pertanika Journal of Social Sciences & Humanities, 26(3).

21. Kamineni A., Williams M. A., Schwartz S. M., Cook L. S., & Weiss N. S. (1999). The incidence of gastric carcinoma in Asian migrants to the United States and their descendants. Cancer causes & control, 10 (1), 77–83. https://doi.org/10.1023/a:1008849014992 PMID: 10334646

22. Lee J., Demissie K., Lu S. E., & Rhoads G. G. (2007). Cancer incidence among Korean-American immigrants in the United States and native Koreans in South Korea. Cancer Control, 14(1), 78–85. https://doi.org/10.1177/107327480701400111 PMID: 17242674

23. Mousavi S. M., Sundquist J., & Hemminki K. (2011). Does immigration play a role in the risk of gastric cancer by site and by histological type? A study of first-generation immigrants in Sweden. Gastric cancer, 14(3), 285–289. https://doi.org/10.1007/s10120-011-0033-5 PMID: 21431296

24. Moore S. P., Antoni S., Colquhoun A., Healy B., Ellison-Loschmann L., Potter J. D., et al. (2015). Cancer incidence in indigenous people in Australia, New Zealand, Canada, and the USA: a comparative population-based study. The Lancet Oncology, 16(15), 1483–1492. https://doi.org/10.1016/S1470-2045(15)00232-6 PMID: 26476758

25. Arnold M., Moore S. P., Hassler S., Ellison-Loschmann L., Forman D., & Bray F. (2014). The burden of stomach cancer in indigenous populations: a systematic review and global assessment. Gut, 63(1), 64–71. https://doi.org/10.1136/gutjnl-2013-305033 PMID: 24153248

26. Besag J., York J., Mollié A. (1991). Bayesian image restoration, with two applications in spatial statistics. Annals of the institute of statistical mathematics, 43(1), 1–20.

27. Knorr-Held L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. Statistics in medicine, 19(17-18), 2555–2567. https://doi.org/10.1002/1097-0258(20000915/30)19:17/18<2555::aid-sim587>3.0.co;2-# PMID: 10960871

28. Torabi M, Green C., Nugent Z, et. al., (2014). Geographical variation and factors associated with colorectal cancer mortality in a universal health care system. Can J Gastroenterol Hepatol. 28(4), 191–197. https://doi.org/10.1155/2014/707420 PMID: 24729992

29. Torabi M., Bernstein C. N., Yu B. N., Wickramasinghe L., Blanchard J. F., Singh H. (2020). Geographical variation and factors associated with inflammatory bowel disease in a central Canadian province. Inflammatory Bowel Diseases. 26(4), 581–590. https://doi.org/10.1093/ibd/izz168 PMID: 31504519

30. Torabi M (2014). Bowel disorders and its spatial trend in Manitoba, Canada. BMC Public Health. 14:285. https://doi.org/10.1186/1471-2458-14-285 PMID: 24673850

31. Waller L.A., & Gotway C.A. (2004). Applied spatial statistics for public health data (Vol. 368). John Wiley & Sons.

32. Sebastián M. S., Hurtig A. K. (2004). Cancer among indigenous people in the Amazon Basin of Ecuador, 1985–2000. Revista Panamericana de Salud Pública, 16, 328–333. https://doi.org/10.1590/s1020-49892004001100006 PMID: 15729982

33. Desjardins M. R., Hohl A., & Delmelle E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. Applied Geography, 118, 102202. https://doi.org/10.1016/j.apgeog.2020.102202 PMID: 32287518