Cancer risk factors analysis in New Zealand

Todd (Zhen) Zhang, Shubo Feng, Dan Wei, Zichen Zhou

2023-10-16

Background

In the year 2020, New Zealand recorded a total of 27,072 new cancer registrations, reflecting an overall age-standardized rate of 338.1 registrations per 100,000 individuals.(Ora, 2023) This considerable volume of new cancer cases imposes a substantial socio-economic burden on the nation. However, it is worth noting that cancer is often far more manageable and treatable when detected at an early stage. This underscores the critical importance of cancer screening.

In New Zealand, a comprehensive national screening program is in place, covering breast, cervical, and bowel cancers, and it is accessible to all residents. For instance, the colorectal cancer (CRC) screening program is initiated at the age of 60 for the majority of the population. Nevertheless, an exception is made for Māori and Pasifika communities, who commence screening at 50 years of age.(Zealand, n.d.) This approach is founded on the fact that ethnicity has been established as a notable risk factor for CRC, with both Māori and Pasifika populations exhibiting a higher incidence of the disease.

The implementation of risk factors-adjusted cancer screening is essential for cost-effective cancer screening. Prior research has illuminated numerous social and environmental risk factors that are linked to cancer incidence. For instance, temperature has emerged as a significant environmental factor associated with the occurrence of cancer.(Voskarides, 2023) Also, studies have indicated that individuals with higher levels of education tend to have lower cancer rates.(Larsen et al., 2020)

This report collected and analyzed regional cancer incidence and risk factors, aiming at identifying potential risk factors associated with cancer incidence in New Zealand. By doing so, we aspire to provide valuable insights in refining and optimizing cancer screening strategies.

2. Methods

2.1 Data

2.2 Analysis pipline

2.2.1 Cancer Overview

2.2.2 Correlation analysis

3. Results

3.1 Cancer Overview

3.1.1 Age distribution

Table 1: Number of New Registered Cancers

population	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
adult	26,572	26,577	27,075	27,642	28,755	28,458	28,807	29,351	29,205	30,199
child	32	54	36	54	42	42	32	40	36	38

The number of adult cancer registrations is over 800 to 1000 times greater than the number of child cancer registrations.

This table presents cancer registration data from 2011 to 2020, categorized into adults and children. While there has been an upward trend in adult cancer registrations, child cancer registrations have remained relatively stable.

Given the low proportion of child cancer registrations in the overall dataset, the small percentage of child data does not significantly impact our analysis. And our analysis primarily focuses on the general cancer situation in New Zealand. Furthermore, the cancer data we've collected from various regions doesn't differentiate between adults and children. As a result, we can't analyze adult data separately in subsequent analyses.

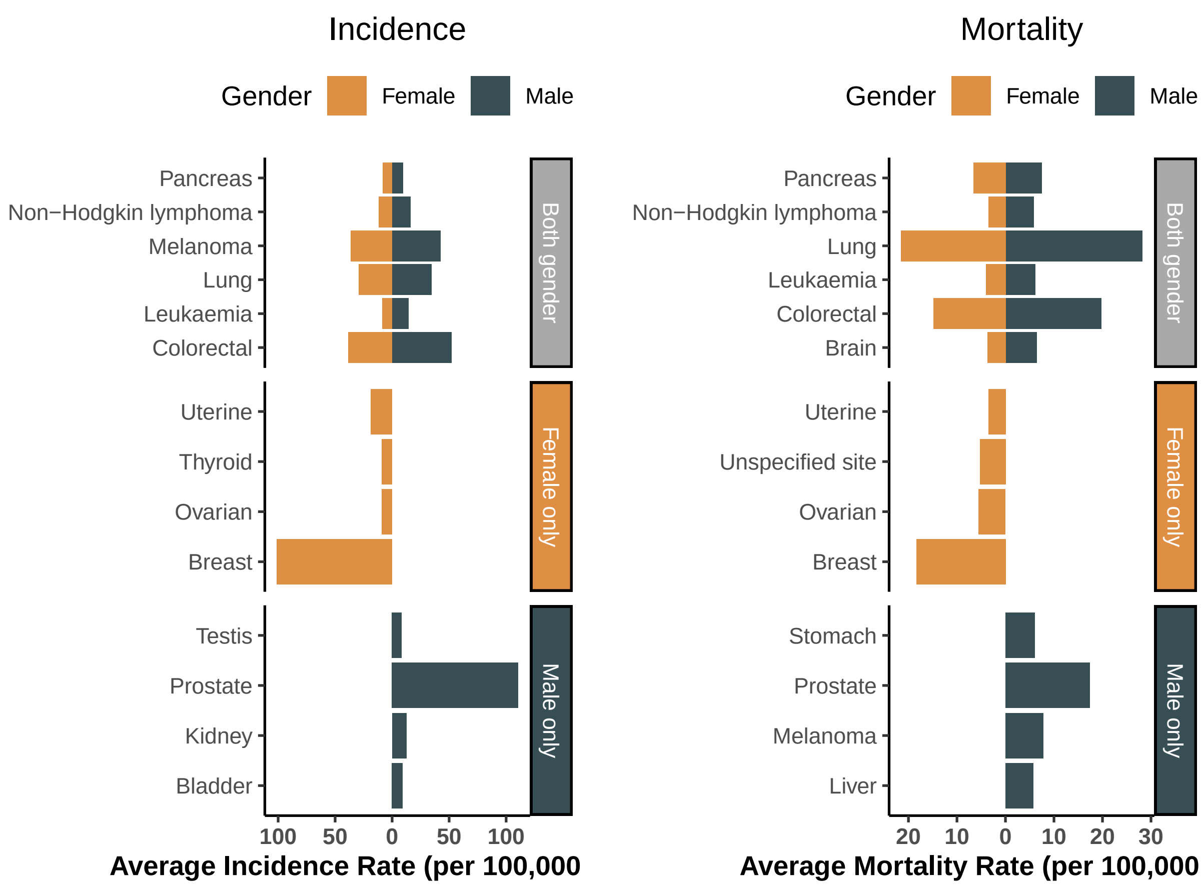

Figure 1: Average Incidence Rate and Average Mortality Rate for each cancer type by sex

3.1.2 Gender variance Females had the highest average incidence of "Breast cancer" from 2011 to 2020, while males had the highest for "Prostate cancer". "Colorectal cancer" also showed high average incidence rates for both genders.

For average mortality rates over the same period, both genders exhibited the highest rates from “Lung cancer”. For females, “Breast cancer” was the second-leading cause of mortality, while for males, it was “Colorectal cancer”.

This figure visually represents these average incidence and mortality rates for each cancer type by gender based on data from the website https://tewhatuora.shinyapps.io/cancer-web-tool/. However, some data points were missing: female data for “Kidney cancer” and “Bladder cancer” and male data for “Thyroid cancer”.

Cancers exclusive to one gender, like “Uterine cancer”, “Ovarian cancer”, and “Breast cancer” in females and “Prostate cancer” and “Testicular cancer” in males, were analyzed separately. For other cancer types with data for both genders, separate analyses were conducted. Generally, average incidence and mortality rates were similar between genders, though males had slightly higher incidence rates.

Among cancer types with data for both sexes, the average incidence and mortality rates were generally similar between males and females, with slightly higher average incidence rates among males. In subsequent analyses, cancers with data for both genders were analyzed using “all sex” data, while cancers with data available for only one gender were analyzed using the corresponding gender-specific data.

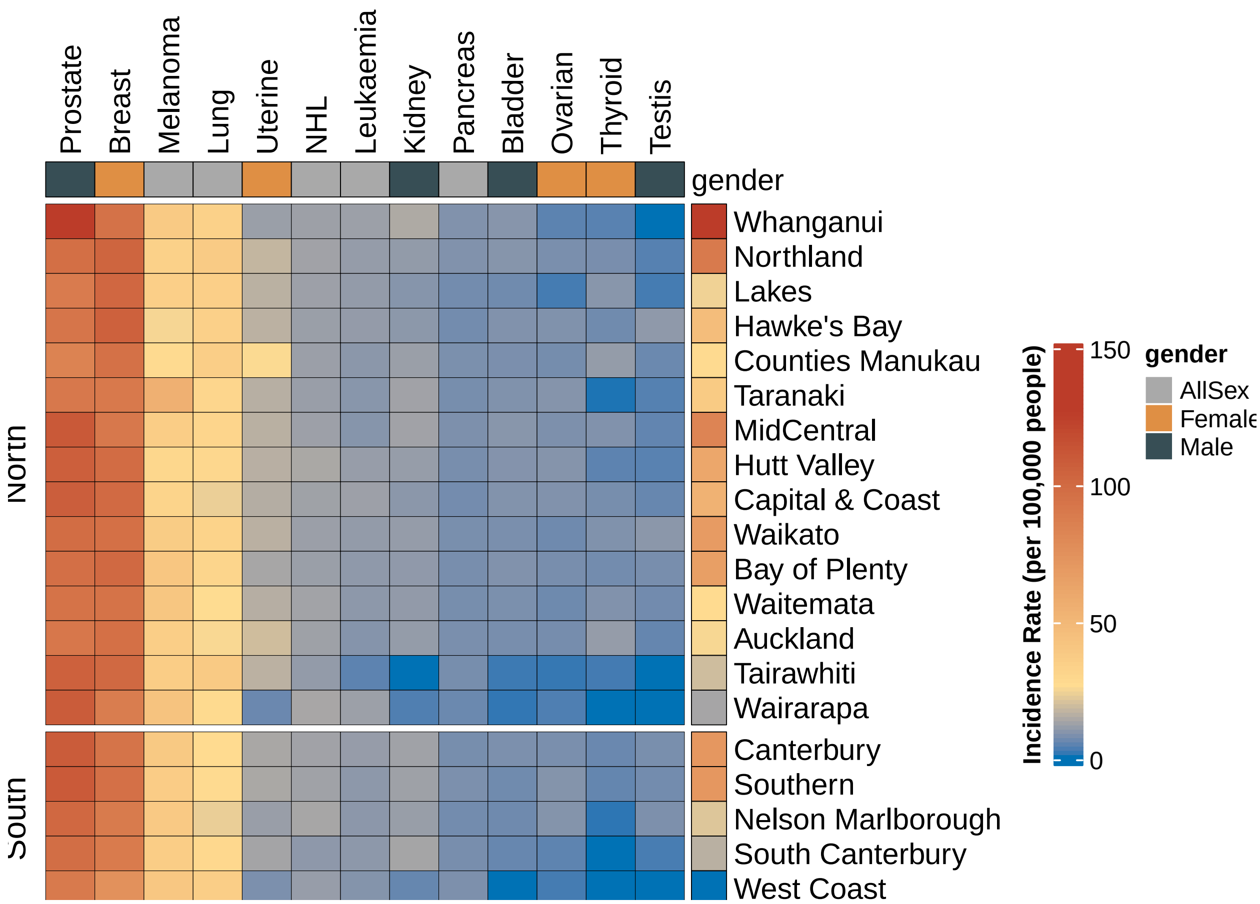

Figure 2: Heatmap illustrating the regional distribution of average cancer incidence during 2011 to 2020 in New Zealand. Gradient color indicated the average cancer incidence rate for the corresponding region.

3.1.3 Regional distribution From 2011 to 2020, Whanganui had the highest cancer incidence rate among all New Zealand regions, with the West Coast recording the lowest.

The heatmap uses shades of orange to represent higher cancer incidence rates and shades of blue for lower rates. Whanganui and Northland had notably high rates, whereas the West Coast had the lowest.

The heatmap also shows that “Prostate cancer” and “Breast cancer” had the highest incidence rates in every region, while “Thyroid cancer” and “Testicular cancer” had the lowest rate.

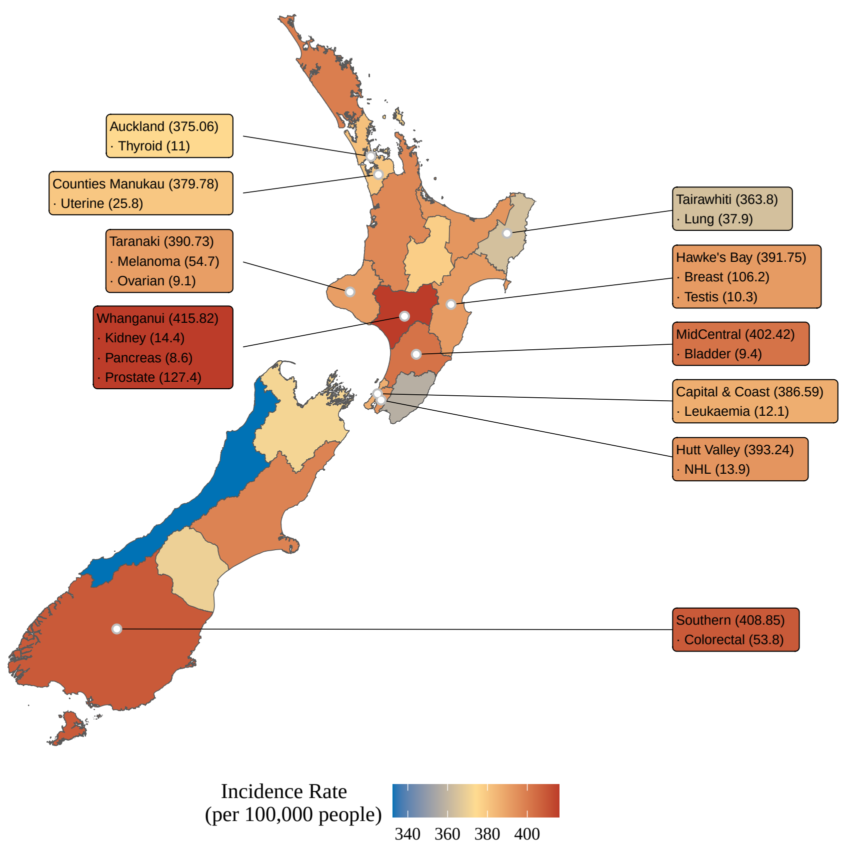

Figure 3: Regional distribution of overall cancer incidence in New Zealand. Gradient color indicated the average cancer incidence rate (All cancer types) for the corresponding region during 2011 to 2020. Lables highlight the most common region for corresponding cancer type

This heatmap provides a clearer depiction of regional cancer rates. Whanganui topped the list with an incidence rate of 415.82/100,000, primarily driven by “Kidney cancer”, “Pancreas cancer”, and “Prostate cancer”. The Southern region followed with 408.55/100,000, where “Colorectal cancer” was most prevalent. The West Coast had the lowest incidence rate.

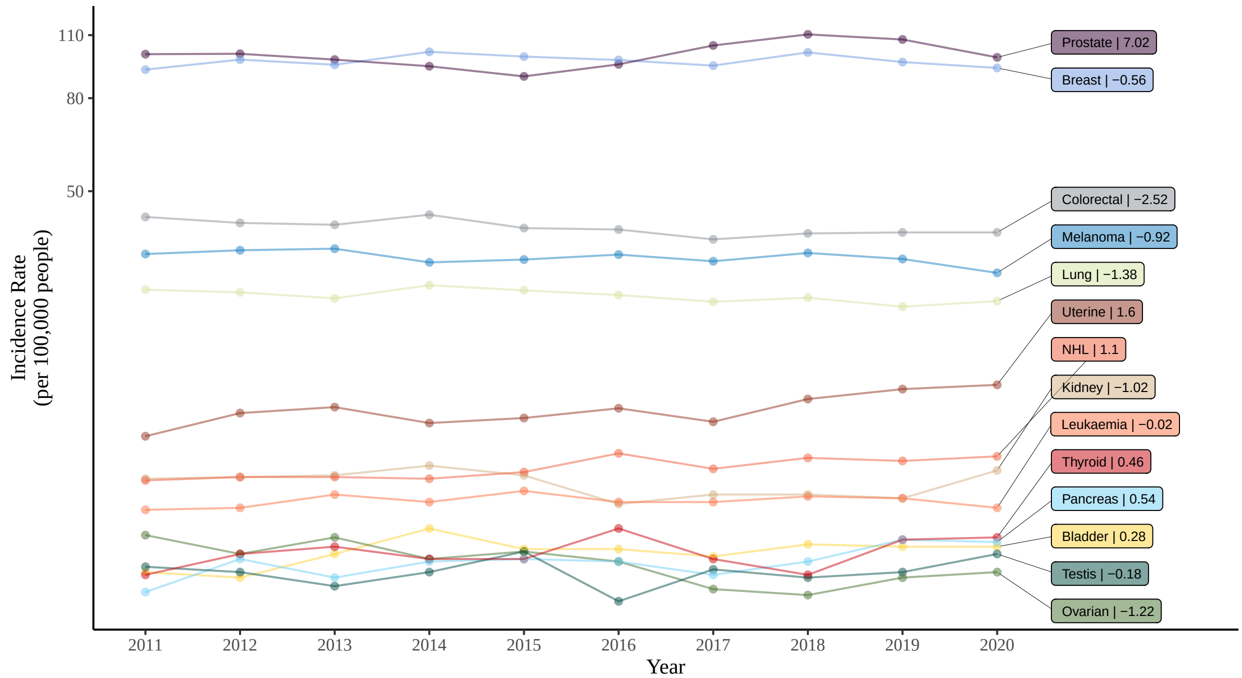

Figure 4: Temporal trends for cancer incidence of each cancer type during 2011 to 2020. The number in each label indicate the difference between average incidence of 2016-2020 and 2011-2015.

3.1.4 Tempory trends Between 2011 and 2022 year, the incidence rates of “Prostate cancer”, “Uterine cancer”, “Non-Hodgkin lymphoma cancer”, “Thyroid cancer”, “Pancreas cancer”, and “Bladder cancers” all increased. In contrast, “Colorectal cancer” showed a significant decline.

The number in this figure’s label represent the difference between median cancer incidence rates from 2016 to 2020 and those from 2011 to 2015, indicating growth or decline trends over these five-year periods.

Lastly, the incidence rates for “Thyroid cancer” and “Testicular cancers” fluctuated substantially during this decade, while other cancers had more stable trends.

3.2 Correlation analysis

4. Discussion

Discussion

Our analysis delves into the multifaceted landscape of cancer risk factors, including various cancer types in New Zealand. The results we present not only corroborate prior research findings but also introduce new dimensions to our understanding of cancer etiology, underscoring the importance of considering environmental and social factors in cancer prevention strategies.

Colorectal Cancer Risk Factors: Our findings affirm well-established associations between PM2.5 (Ku et al., 2021) or cold exposure and increased colorectal cancer risks (Lehrer & Rosenzweig, 2014). This connection emphasizes the pivotal role of environmental factors, specifically air quality, in shaping colorectal cancer risk. Furthermore, the inclusion earthquakes as a risk factor for colorectal cancer, merits further investigation, especially in earthquake-prone regions like New Zealand. The association between earthquakes and colorectal cancer may be attributed to environmental changes that occur following seismic activity.

Lung Cancer Risk Factors: Our analysis reveals risk factors for lung cancer, including overweight (Vedire, Kalvapudi, & Yendamuri, 2023), lack of physical activity (Cannioto et al., 2018), hypertension (Lindgren, 2003), and smoking (Walser et al., 2008), which align with well-documented risk factors for this cancer type. Lifestyle choices and health status are crucial contributors to lung cancer risk. Moreover, the inclusion of low income and education levels underscores the socio-economic determinants of lung cancer, as these factors are closely associated with smoking habits (Larsen et al., 2020). The surprising link between dental health care visits and lung cancer emphasizes the significance of oral health in cancer prevention. Several studies have demonstrated that poor oral health is linked to an increased risk of lung cancer (Yoon et al., 2019), suggesting the importance of regular dental check-ups for maintaining oral health and potentially reducing lung cancer risk.

Melanoma Risk Factors: For melanoma, our analysis highlights risk factors such as PM2.5 concentration and average annual earthquake depth, indicative of the role of environmental factors in this skin cancer. Additionally, teeth loss and mean height are intriguing findings that merit further exploration. Understanding the mechanisms behind these associations is essential to develop effective strategies for melanoma prevention in New Zealand.

Breast Cancer and Drinking: The inclusion of heavy episodic drinking as a risk factor for breast cancer, as indicated by previous studies such as (McDonald, Goyal, & Terry, 2013), underscores the complexity of breast cancer etiology. This finding emphasizes the importance of addressing lifestyle factors, including alcohol consumption, in breast cancer prevention efforts.

Prostate Cancer and Work Hours: While our analysis suggests that lower work hours may be a risk factor for prostate cancer, it's important to acknowledge a potential bias. Lower work hours could be associated with older age, a known vulnerability factor for prostate cancer. However, due to the absence of age data in regional cancer incidence, further age-specific analysis is needed to confirm or rule out this bias.

Our analysis has limitations, including the inability to conduct subgroup analyses based on ethnicity or age group due to the lack of relevant data, potentially hindering our understanding of demographic-specific cancer incidence variations. Additionally, our correlation analyses were executed at the regional level, whereas comparing risk factors between cancer and non-cancer individuals at the individual level would have yielded more accurate results, although such data were not publicly available. Furthermore, some data sources lacked comprehensive variable documentation, introducing ambiguity in our interpretation of certain risk factors, such as "heavy episodic drinking" from the New Zealand Health Survey, for which a precise definition was absent.

In summary, our findings expand our knowledge of cancer risk factors for various types of cancer in New Zealand. These results underscore the multifaceted nature of cancer etiology and the importance of considering environmental and social factors in cancer prevention strategies. Our findings can serve as a foundation for the development of targeted prevention and screening strategies tailored to the unique risk factors associated with different cancer types, ultimately contributing to a reduction in the cancer burden in New Zealand. Further research on less-explored risk factors, such as earthquake-related cancer risks, will be crucial in advancing our understanding and improving prevention efforts.

References

Cannioto, R., Etter, J. L., LaMonte, M. J., Ray, A. D., Joseph, J. M., Qassim, E. A., ... Moysich, K. B. (2018). Lifetime physical inactivity is associated with lung cancer risk and mortality. *Cancer Treatment and Research Communications*, *14*, 37–45. https://doi.org/10.1016/j.ctarc.2018.01.001

Ku, M.-S., Liu, C.-Y., Hsu, C.-Y., Chiu, H.-M., Chen, H.-H., & Chan, C.-C. (2021). Association of ambient fine particulate matter (PM _{2.5}) with elevated fecal hemoglobin concentration and colorectal carcinogenesis: A population-based retrospective cohort study. *Cancer Control*, *28*, 107327482110412. https://doi.org/10.1177/10732748211041232

Larsen, I. K., Myklebust, T. Å., Babigumira, R., Vinberg, E., Møller, B., & Ursin, G. (2020). Education, income and risk of cancer: Results from a norwegian registry-based study. *Acta Oncologica*, *59*, 1300–1307. https://doi.org/10.1080/0284186X.2020.1817548

Lehrer, S., & Rosenzweig, K. E. (2014). Cold climate is a risk factor for thyroid cancer. *Clinical Thyroidology*, *26*, 273–276. https://doi.org/10.1089/ct.2014;26.273-276

Lindgren, A. (2003). Blood pressure, smoking, and the incidence of lung cancer in hypertensive men in north karelia, finland. *American Journal of Epidemiology*, *158*, 442–447. https://doi.org/10.1093/aje/kwg179

McDonald, J. A., Goyal, A., & Terry, M. B. (2013). Alcohol intake and breast cancer risk: Weighing the overall evidence. *Current Breast Cancer Reports*, *5*, 208–221. https://doi.org/10.1007/s12609-013-0114-z

Ora, T. W. (2023). *Cancer web tool.* Retrieved from https://www.tewhatuora.govt.nz/our-health-system/data-and-statistics/nz-health-statistics/health-statistics-and-data-sets/cancer-data-and-statistics/cancer-web-tool

Vedire, Y., Kalvapudi, S., & Yendamuri, S. (2023). Obesity and lung cancer—a narrative review. *Journal of Thoracic Disease*, *15*, 2806–2823. https://doi.org/10.21037/jtd-22-1835

Voskarides, K. (2023). The double face of cold in cancer. *Translational Oncology*, *28*, 101606. https://doi.org/10.1016/j.tranon.2022.101606

Walser, T., Cui, X., Yanagawa, J., Lee, J. M., Heinrich, E., Lee, G., … Dubinett, S. M. (2008). Smoking and lung cancer: The role of inflammation. *Proceedings of the American Thoracic Society*, *5*, 811–815. https://doi.org/10.1513/pats.200809-100TH

Yoon, H.-S., Wen, W., Long, J., Zheng, W., Blot, W. J., & Cai, Q. (2019). Association of oral health with lung cancer risk in a low-income population of african americans and european americans in the southeastern united states. *Lung Cancer*, *127*, 90–95. https://doi.org/10.1016/j.lungcan.2018.11.028

Zealand, B. C. N. (n.d.). *New zealand's national bowel screening.* Retrieved from https://bowelcancernz.org.nz/about-bowel-cancer/early-detection-and-prevention/screening/

Answer to questions

Q4: What difficulties you had to overcome to wrangle the data sources into the target data model? Mapping Coordinates to DHB Regions: Our research focused on the relationship between cancer occurrences and socio-environmental factors, utilizing "region and year" as a common identifier. While cancer data was available in "DHB regions," other datasets primarily used "council regions." Initially, we aimed to map "DHB regions" to "Council regions," but this proved impractical. For humanities datasets, we filtered out those lacking DHB region information. For environment datasets, with available coordinates information, we overcame the challenge by using geographical coordinates and "R library sf" to map them to DHB geometrics.

Representing Multiple Data Points for Each Common Identifier: In the case of environment-related datasets, such as earthquakes, we simplified representation by selecting key statistical measures - maximum, average, minimum, and frequency of events - for each DHB region within a given year. The same approach was applied to datasets concerning air quality, groundwater quality, and temperature.

Creating Variables for Enhanced Representation of Humanities Factors: Original humanities datasets expressed counts within various categories, e.g., the number of individuals with a particular qualification level (e.g., number of people with qualification > level 3). To ensure data comparability, we converted these counts to percentages. We further analyzed the percentage of individuals with qualifications at or above specific levels, generating new variables based on cumulative sums of percentage (e.g., percentage of people with qualification > level 3). This approach was also applied to income, work hours, and birth numbers datasets.

Q6: what you managed to achieve and what you failed to do Achievements: We successfully collected and cleaned comprehensive cancer incidence and mortality data, as well as risk factors data. Our primary goal was to establish a connection between these diverse data sources using the "region + year" as a common identifier. This allowed us to conduct a thorough analysis of cancer incidence patterns across different regions and investigate correlations with various risk factors. The analysis revealed valuable insights and identified significant risk factors associated with the most common cancer types, including Lung, Prostate, Breast, Melanoma, and Colorectal cancers. These findings have practical implications for optimizing cancer prevention strategies.

Limitations: One notable limitation was our inability to perform subgroup analyses based on ethnics or age group, despite the potential variations in cancer incidence within different demographic groups. Additionally, our correlation analyses were conducted at the regional level, treating each region as an independent sample. Ideally, comparing risk factors among cancer and non-cancer individual could provide more accurate and robust risk factor results. Unfortunately, such individual-level data are not publicly available.