Author’s response to reviews

Title: Ethnic Inequalities in Cancer Incidence and Mortality: Census-Linked Cohort Studies with 87 Million Years of Person-Time Follow-Up

Authors:

Andrea Teng (andrea.teng@otago.ac.nz)
June Atkinson (june.atkinson@otago.ac.nz)
George Disney (george.disney@otago.ac.nz)
Nick Wilson (nick.wilson@otago.ac.nz)
Diana Sarfati (diana.sarfati@otago.ac.nz)
Melissa McLeod (melissa.mcleod@otago.ac.nz)
Tony Blakely (tony.blakely@otago.ac.nz)

Version: 1 Date: 10 Jul 2016

Author’s response to reviews:

University of Otago, Wellington

11th July 2016

Dear Editor

Re: BCAN-D-16-00569

Title: Ethnic Inequalities in Cancer Incidence and Mortality: Census-Linked Cohort Studies with 87 Million Years of Person-Time Follow-Up
Thank you for the reviewer’s comments and the opportunity to revise this article. Please find attached our response and the revised version of the article with our improvements. We hope that these changes will allow the revised version to be published in BMC Cancer.

Kind regards

Dr Teng (nee McDonald) on behalf of the co-authors

Reviewer Feedback and Our Responses

Reviewer #1

This is an interesting article linking census data and cancer registries in New Zealand to assess inequalities in cancer-related mortality in the Maori and Pacific population compared to the European/Other population.

The cancer subtypes and categories were well thought out to reflect larger socioeconomic/epidemiological issues (ex: tobacco usage, obesity, infection, and cancer prevention via screening). It may be interesting to comment on other factors that may contribute to cancer incidence or mortality (ex: alcohol consumption) if available.

Our response:

• Thank you for the favourable feedback. A sentence was added in the Introduction about the association of alcohol with breast cancer on line 150 “Alcohol is a risk factor for liver, breast and colorectal cancers. [15, 16]”

• Alcohol patterns were added into figure 1.

• In the Discussion, a paragraph was added “Total alcohol consumption does not differ much between Māori and European/Other and differences by ethnicity have not changed much over time (Fig 1). European/Other are more likely to have drunk alcohol in the last year but Māori and Pacific (males only) have higher rates of hazardous drinking patterns.[23] However, for Māori and Pacific women there may be some suggestion that
alcohol consumption (in the last year) has increased over time to match European/Other and narrowed the gap (Fig 1). This may have contributed somewhat to the pattern of inequalities observed for breast and colorectal cancer in women, but it does not contribute to explaining the similar patterns in men for colorectal cancer.”

I have a general question regarding the practice of colorectal cancer screening in New Zealand via colonoscopy or fecal occult blood testing/sigmoidoscopy (Line 148). In the United States, colon cancer screening is a mainstay of preventative care. I am curious to know why this is not the practice in New Zealand and whether access to screening/access to care may have contributed to the general decrease in mortality and incidence noted among the European population compared to the increased incidence and mortality among Maori and Pacific populations observed in Table 1.

Our response:

• Healthcare provision in New Zealand is predominantly funded through general taxation and provided through the public system. Colonoscopies are offered if indicated e.g. for diagnostic workup relating to symptoms. There has never been a population-based screening program for colorectal cancer in New Zealand. However, the government has recently announced that such a program will be introduced, region by region, using immunochemical faecal occult blood test.

• It is possible that European/Other have had more access to colonoscopy and iFOBT screening in the past e.g. through private insurance, but this access is low and will have had little impact, particularly during the 1980s and 90s when SRRs were lowest.

Also, for figure 2, is there any issue with census data/capture of the pacific population in 1981-1984 compared to the subsequent years resulting in the marked increase in SRDs over time? If so, would this interfere with any of the conclusions extrapolated from this data regarding the Pacific female population?

Our response:

• The exact same methods were used in 1981-84 as in other years. In 1981-84 the Pacific population was much smaller than it is now and results are consistent with other years when CIs are considered. See Table 2 and Figure S6 in the supplementary figures for an indication of the uncertainty.
Lines 175, 191, 193, 199 should clarify reference to Figure 3a, rather than Figure 3.

Our response:

• See last comment below – these now revert back to “Figure 3”.

Line 219 regarding incidence data for endometrial, liver, and cervical cancer - it may be helpful to clarify "unlike mortality trends" in this area with a repeat reference/explanation from line 140.

Our response:

• New text added: “with data for endometrial, liver and cervical cancer, not available for mortality trends due to small numbers” (line 237)

Line 253: I would consider rephrasing the sentence "lower colorectal cancer incidence in Maori males and females significantly declined" - the double negative in this sentence may be initially confusing in reflecting an overall increase in colorectal cancer incidence in this population

Our response:

• New sentence added as a replacement: “Furthermore, colorectal cancer incidence in Māori males and females increased towards the European/Other incidence thus significantly narrowing the ethnic difference that was previously favouring Māori.”

My last general comment would be in rearranging and renumbering the figures to correspond with the order of reference in the text. For example, as Table 1 and Figure 3a are discussed in the text in context of each other and Figure 4 and 3b are discussed together, it may make sense separate figure 3a from 3b and rename figure 3a as figure 3 and figure 4 as figure 4a and figure 3b as 4b.
Our response:

• To simplify things, we have removed the a’s and b’s from the figure names. Then we switched the order of the final two tables to align with the text as recommended.

  o Fig1=1
  o Fig2=2
  o Fig3a=3
  o Fig3b=5
  o Fig4=4

Reviewer #2

I think the data is impressive and the interesting results will make an important contribution to the knowledge on ethnic inequalities in mortality and the role of cancer in the excess mortality observed in minority ethnic group in New Zealand.

I believe the manuscript is not fully ready for publication though and requires more work, a bit of editing and refined writing in the description of linkage, bias, methods and consequently results to ensure that the reader follow smoothly what is going on.

Our response:

• Thank you for the favourable comment on the impressive data. Please see responses to the specific comments below.

More specific comments:

METHODS:

- line 110: replace "date" by "data"

  o Our response: Done.
- lines 115-117: In one case, you call 2001-06 cohort and next sentence 2001 cohort. Be consistent in how you name your cohort.

  Our response: Changed to “2001 census” (the incidence cohort is actually 2001-04). Cohort names are different for incidence and mortality data – where follow-up is for a different length of time.

- lines 115-116: Why using 3 years of mortality data for earlier cohort and 5 years for the 2 most recent cohorts? Don't you have mortality data all over the period 1981-2011?

  Our response: Earlier cohort linkages only linked census individuals to deaths for the shorter follow-up frame of three years due to particular concerns at that time about diminishing record linkage rates over time since the census. The recent cohorts from recent updates have included a full five year follow-up (with little diminishing linkage with increasing years since census date). We present all the data that are available.

- lines 118-119: Was probability record linkage not done on person level data? No name - first name? Does that mean your linkage only provide data on aggregated level which further enable calculation of directly standardised rate? Can you explain the linkage in more details so we can further understand that it was an individual or aggregated type of matching.

  Our response: It was individual matching. This is clarified on line 119 “The probabilistic record linkage was done with QualityStage software using an individual’s address....” It is true that we did not use name as a matching variable because names were not available for the datasets.

- line 125: You say you corrected for linkage bias. How is the correction done? Can you give an example of the linkage bias in your study and whether it has been assessed through some pre-analysis? Did you do some work to assess the linkage bias or not prior adjusting for it? If so can you explain the process and the results?

  Our response: An example of the bias was added “The percentage of deaths linked to a census record ranged from 71% (1981 mortality linkage)[12] to 83% (2006 registration linkage).[13]”

  An explanation of the process has also been added “For example, if 20 out of 25 eligible cancers for Māori males aged 50-54 years old of high deprivation living in the north of New Zealand, were linked back to their census record, each of the 20 linked records was weighted up by 25/20 = 1.25.”
- lines 139-140: Rephrase "Mortality rate differences were presented for the same cancers where numbers were adequate". By "same cancer", do you mean for each individual cancer (stratified analysis)? What are adequate numbers? Did you have a threshold, minimum N?

o Our response: Now rephrased: “Mortality rate differences were presented for all these cancers except for endometrial, cervical and liver cancer mortality where there was a smaller number of deaths.”

- lines 142-143: It is indeed sad not to be able to account for other mortality data or migration. But what about this "other than for 2006-2011" into parenthesis? Does that mean you did adjust for death for that particular cohort? Or you just limited to cancer death for all cohorts regardless of the data available in the most recent cohort?

o Our response: We altered the sentence to make this clearer; “We could not censor for non-cancer mortality (other than for 2006-11) and out migration because census-cancer and census-mortality datasets were not linked in earlier cohorts. In the 2006-11 cohort, however, we did censor for non-cancer mortality.”

Was there no possibility to link your data to overall mortality? Could you explicitly explain why?

o Our response: Cancer was linked to overall mortality, but in the earlier cohorts; census-cancer and census-mortality linkage were undertaken separately.

- line 147: What is "Changing reproductive patterns"? Could you specify or give an example?

o Our response: “such as access to contraception and later child bearing” inserted.

- lines 148-149: In other countries, we found minority ethnic groups having lower participation into screening but also lower risk of Cancer. I believe Colorectal screening might not impact on incidence as cancer will get detected at some point but lower participation might lead to a diagnosis at a later stage and henceforth screening participation might impact on mortality.

o Our response: Screening with iFOBT, flexible sigmoidoscopy and colonoscopy have all been found to reduce CRC incidence because they identify and remove adenomas which are precancerous lesions.
- line 151: I don't understand why you have used the "WHO World Standard Population" rather than New Zealand population for standardisation.

  Our response: New text added to the sentence: “to maximise international comparability”

- lines 152-154: I’m used to the terminology (Directly) "Standardised Mortality Rates (SMR)". So if I understand, here you have Standardised rate (specify SMR here?), rate differences (which you further refer to as SRD in Results? So you should add SRD here) and rate ratios (would be some kind of RR of standardised rates? Apart from the second part of table 1, there was no rate ratio described in your results section so why including it?).

  Our response: We have clarified meaning and added “standardised” in front of the rate differences and rate ratios, as it applies to all of these.

  Rates were standardised for cancer incidence and cancer mortality so we stick to the terminology here of standardised rates rather than specifying standardised mortality rates.

  Standardised rate ratios are included because it is important to consider both absolute and relative inequalities. A sentence has been moved from the Discussion and added in to the end of the Results: “Stomach and liver cancer accounted for the greatest relative ethnic inequalities (SRRs) with three- to six-fold greater incidence in Māori and Pacific populations compared to European/Other (2006-11).”

- line 157: Do you assess trend across cohort (time)? What are the variables in the model?

  Our response: Trends are assessed across time using the middle date for each of the cohorts as an independent variable in a linear regression equation, with the standardised rates and rate ratios (logs) and standardised rate differences as dependent variables.

  Sentence added to Methods: “Statistical tests of increasing or decreasing linear trend (linear regression) were calculated on log rates, log rate ratios and rate differences, using the mid-date of each cohort period as the independent variable”


  Our response: Done.
RESULTS:

- lines 161-163: The decline here seems to be over 3 years for the oldest cohort and 5 years for the most recent one. My point there is that analysis is done over time for each cohort rather than across cohort and time so it might be confusing from the reader side.

What trend did you get for the other 4 cohorts?

o Our response: The trend was taken from the linear regression analysis across all six cohorts. We have now added “All-cause mortality declined for all ethnic groups across all six cohorts from 1981-84 and 2006-11.”

- lines 161-165: You said you did not have the other non-cancer mortality data earlier? Is it a different data source you refer to here like national data not linkable to your cancer data? Could you add either reference or sources?

o Our response: See comment above. This has been specified more clearly in the Methods.

- Figure 2: Are you calculating the difference between SMR in Maori and SMR in European/Other to get SRD? Could you add more information to the methods and/or figure to make that clear?

I suppose you are calculating SMR over 3 years of mortality data (or 5 years). Could you again give this notion into figure/method as well as detailing denominators? Is it 3 years data over N number of people of a specific ethnic group standardised by X, Z, Y?

o Our response: Addition to the Methods: “Standardised rates, were calculated in 1-74 year olds for each ethnic group in each cohort (number of events per person years of follow-up). Standardised rate differences (SRDs) and standardised rate ratios (SRRs) among 1-74 year olds were calculated for Māori and Pacific compared to European/Other”.

o The cohort dates are on the x-axis. We have changed the title to “Contribution of cancer and cardiovascular disease (CVD) to ethnic inequalities in all-cause mortality over time for Māori and Pacific peoples (compared to the European/Other population) 1-74 years old in New Zealand, census-linked mortality data for six cohorts between 1981 and 2011”

Again, is figure 2 based on all-cause mortality data which is another source of data (not linkable to your cancer data)?
Our response: Yes, in previous years cancer analyses were separate and not linked to mortality analyses.

- line 167: specify "in the cohort 1981-84" otherwise we get confused between the notions of time, overtime, trends, by cohort, across cohort.

Our response: Done.

This comment applies to the whole manuscript.

Our response: Done.

- line 167: replace "double this" by "doubled".

Our response: Done

- line 172: "from 1981 to 2011"

Our response: Done.

- line 173: "in the cohort 2006-11"

Our response: Done

- line 175: What do you mean by "From 1981-84 to 2006-11"? Is it overtime or for each cohort?

Our response: Now addressed.

A general comment is that it is difficult to disentangle the notion of time and cohort. Sometimes the analysis is done on each cohort with a single estimate for 3 years data and sometimes, the analysis is done overtime to look at trend over 3 years for each cohort and then there are also trends over the full period considering the 6 cohorts results as a continuous time. It is hard not to get a bit lost. It can certainly be improved by a clear analysis structure and adding detailed explanations on each part.
• Our response: All trends are presented over time across all six cohorts using the mid-dates. No trends are presented within cohorts. See comments above which have clarified this.

DISCUSSION:
- lines 244-247: Too long, split the sentence in two.
  o Our response: Done.

- line 247: I would replace "what" by "which"
  o Our response: Done.

- line 254: replace ",;" by ",,"
  o Our response: Done.

- line 277: You introduce a US reference here but could you explain how your results relate to it? There is also literature on the matter elsewhere in the world like in Europe for example.
  o Our response: Changed to “Similar to our study, in the United States (US) Non-Hispanic Black population (compared to White) there were higher rates of obesity,[21] endometrial cancer (7.5 per 100 000 vs 4.0), breast cancer (30.6 vs 21.7), and also colorectal cancer incidence (male 27.7 vs 18.5, female 18.5 vs 13.0) which we did not identify.[22] “
  o The European literature on specific cancer inequalities is largely focussed on socioeconomic position (SEP).

- lines 301-304: Could it be due to higher co-morbidity and/or later diagnosis?
  o Our response: Yes we think so. This point has now been added to the text.

- line 315: Explain the reason why. Was it that your data did not have any socio-economic variable available? Or the methods could not incorporate an additional socio-economic standardisation?
Our response: We did not standardise for SEP because in our view SEP is not a confounder of the ethnicity-mortality relationship. Instead SEP is on the causal pathway ie, it is a mediator.

Analyses elsewhere have examined the mediation effect of SEP on inequalities and we are planning further analyses to publish in this area in a separate paper. SEP is available in our data.

We will remove this confusing paragraph to save on word space.

Final comments: Thank you very much for this helpful Reviewer feedback which has helped to improve the manuscript further. Of note is that we have taken this opportunity to update the numbers in the manuscript via a slight improvement in the person-time calculation. This has been done so as to make our mortality figures consistent with forthcoming analyses and papers. This has meant there are very slight changes in the numbers for the rates, SRDs, SRRs in mortality figures BUT there is no change to the overall pattern of results (or associated changes in how the results are described in the text).