PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

| TITLE (PROVISIONAL) | Reduction in Myocardial Infarction Admissions in Liverpool after the Smoking Ban: Potential Socio-Economic Implications for Policymaking |
|---------------------|--------------------------------------------------------------------------------------------------------------------------|
| AUTHORS             | Liu, Andrew; Castillo, Maria; Capewell, Simon; Lucy, John; O'Flaherty, Martin                                                  |

VERSION 1 - REVIEW

| REVIEWER            | Stanton A. Glantz, PhD |
|---------------------|------------------------|
|                     | Professor of Medicine  |
|                     | Director, Center for Tobacco Control Research and Education |
|                     | University of California San Francisco |
|                     | No competing interests. |
| REVIEW RETURNED     | 18-Jun-2013             |

THE STUDY

This paper uses data from Liverpool to evaluate changes in hospital admissions for acute myocardial infarction and other cardiac conditions following implementation of England’s smokefree law. Consistent with earlier work, it shows a drop in AMI hospital admissions. In a new result, it examines the effects of the law on health disparities and finds consistent effects across all economic groups.

Both questions are important and the data are fine, but there are several issues with the analysis.

First, there is no need to smooth the data before doing the regression analysis. Doing so artificially reduces the variance in the data and makes the fit look better than it is. There is also the chance that the period used for the moving average smoother could affect the assessment of break points in the data.

Second, the use of jointpoint regression with 5 breakpoints seems overkill, especially given that there is no reason to expect so many sudden changes in AMI rates. The authors would be better served to use the approach that most others have of modeling the possibility of a sudden change followed by a slope change after the law. Alternatively, they could allow for two slope changes, one immediately following the law and one some time later. (Which of these approaches is most reasonable can be better assessed when considering the unsmoothed data.)

Estimating the incident rate ratios by comparing two year periods before and after the law is a poor way to do the analysis because it does not account for any underlying secular trends. It would be much better to treat the data as a time series and directly estimate the IRR using a negative binomial regression. Doing so would also
collapse the two different analyses into a single analysis, which would shorten and simplify the paper.

The comparison across income groups also needs to be integrated into the overall interrupted time series analysis, using and ordinal variable if possible.

Other points:

The authors do not need to be so defensive about the use of interrupted time series analysis (compared with randomized controlled trials). As the authors note, randomized controlled are impossible in such studies. Interrupted time series studies are widely used to assess causality in evaluating public policies.

Many abbreviations (MI, CHD, HES, NHS, etc) need to be defined the first time they are used.

The authors needs to say precisely when in 2007 the law took effect and include this date in their analysis.

Page 2, line 10: This statement understates the strength of existing evidence. Doing so is not necessary to justify this work.

Page 2, line 42: “Is” not “appears to be.”

Page 3, line 20: Drop this bullet. It implies that the authors did something radically different from others, when in fact they did.

Page 3, line 37: Drop this bullet.

Page 4, line 51: Saying health inequalities “were not widened” presumes that they would be. “Were not affected” would be a more neutral statement.

Page 9, line 23: For the reasons stated above, the lab period is an artifact of the way the analysis was constructed. In addition, the statement of a 3-6 month “lag” is not an accurate representation of what the authors found; they found changes starting immediately. It was several months before the changes flattened out. But, as noted above, this interpretation could be an artifact of the way the model was constructed.

Page 10, line 30: As noted above, the existence of this secular trend is why simply comparing the two years before the law with the two years after is not good enough.

Page 10, line 52: This paper is worth discussing in the context of the authors’ findings, since if finds similar results using an entirely different population and different methodology: Tobacco control policies are egalitarian: a vulnerabilities perspective on clean indoor air laws, cigarette prices, and tobacco use disparities. Dinno A, Glantz S. Soc Sci Med. 2009 Apr;68(8):1439-47. doi: 10.1016/j.socscimed.2009.02.003. Epub 2009 Mar 11. PMID: 19282078.

Page 11, line 25: Drop this paragraph for reasons discussed previously in this review.

Page 11, line 42: Same comment.
If the analysis is redone in a more direct way, the figures and tables will have to be redone and many of the tables will probably no longer be needed.

**RESULTS & CONCLUSIONS**

While there are serious problems with the paper as it is currently presented, all these can be fixed by redoing the analysis appropriately.

**THE STUDY**

Statistical methods: There are important details missing about calculation of rates, derivation of denominators, calculation of moving averages. There is no justification for why a joinpoint analysis has been chosen, whether the data exhibit seasonality or auto-correlations, and why an ARIMA based analysis was not undertaken.

**RESULTS & CONCLUSIONS**

Comments on alcohol, salt and dietary fat in the final section "public health implications" are not supported by data.

**REPORTING & ETHICS**

The supplied STROBE statement is far too light in detail. Most of the questions are answered "time trend study, not applicable". I think that using admin data in this way creates a type of retrospectively assessed population-based cohort study, and the cohort study questions should be applicable in most instances. Just because the data are population based doesn't mean there aren't eligibility criteria, case ascertainment, missing data or that numbers of events shouldn't be reported.

**GENERAL COMMENTS**

This paper reports an interrupted time series analysis of MI admission rates in Liverpool in relation to a legislative ban on smoking in public places. While the health effects of smoking are well known, and public policy efforts to reduce smoking related harms are well justified on principle, evaluation of individual policy interventions is important to determine their effectiveness, document the case for extending programmes to other jurisdictions, to aid in refining programme implementation, and to monitor the possibility of inadvertent consequences.

While there may have been other options for implementing the smoking ban policy that would provide a stronger test of their impact, such as introducing the legislation at different times in different places and comparing changes in outcomes between those places, a before-after comparison in appropriate for a national implementation of this type of policy action. Clearly, observational data such as these do not prove cause and effect. However, in situations where fully randomised designs are not possible, or weren't implemented, observational data may form an important component of the evaluation of new policy or programme initiatives.

I have some questions about the level of detail presented in the paper, which left me with a number of unanswered questions, and also some concerns about the time series analysis techniques employed. I feel the paper would be improved by providing more detail in several areas.
Context

I would suggest including more detail in the introduction. For international readers it would be helpful to have some description of the nature of the smoking bans, when they were introduced, and what they covered. It would also be helpful to have some context on the health services in the Liverpool area. I am not familiar with the HES data base. How many health services does it cover, and what is the population of the region?

Time series analysis

The authors have undertaken a piecewise linear regression analysis using the Joinpoint software. There is no discussion in the paper as to why this particular technique has been chosen, and whether the data meet the conditions for its use. My understanding of the Joinpoint programme is that it was originally developed for describing long-term trends in cancer incidence from annualised data, and that it assumes there is no auto-correlation within the time series. I would have thought that an ARIMA model of some type would have been the most likely first choice of analysis tool for this type of question. It would be helpful to explain why the Joinpoint approach has been chosen, and whether any tests of stationarity in the time series or autocorrelational structure were undertaken.

In the “Trend Analysis” section of the paper, the authors start by saying that plots of age-specific mortality rates were smoothed using 3 year moving averages. Firstly I presume they mean admission rates. It is not clear why 3 year moving averages have been applied to quarterly time series data. I presume this means a 12-point moving average? It is usually not recommended to adopt a moving average with an even number of points included as this can phase shift the series. Also with only 30 quarterly data points available for analysis, an important issue is treatment of the end points (i.e. the first six quarters and the last six quarters) in calculating the moving averages. No mention is made of how this was done, and I feel this detail needs to be included. I presume that the smoothed data was then used as input to the Joinpoint regression, but this is not specified in the text.

The data in most of the tables and graphs appear to be rates rather than numbers of admissions, which is appropriate, but there is no mention where the denominators for rates have been sourced. A common challenge with use of administrative data of this type is how well population counts match the catchment areas of the health services. It is not entirely clear, but I presume the analysis has been restricted to people with a residential address in one of the 30 Liverpool wards, but does the HES data relate to use of health services in the Liverpool region, or can you account for use of health services out of region?

Were you restricted in data access to only using quarterly data? It seems you have enough events to create a monthly time series. I would think this would better allow you to test for seasonality, and might better help determine the lag between intervention and movement in the time series. This is quoted as being between 3 and 6 months, but this is difficult to judge from a quarterly time series.

Another concern with the Joinpoint approach is that it is entirely data
driven. As you have a clear hypothesis as to whether there was a change in admission rates subsequent to the legislation, an ARIMA model that specifically tested this would be closer to the hypothesis of the study. An artefact of the Joinpoint approach is that it inserts turning points in the series that are unrelated to the study hypothesis. For instance, the top line in Figure 2 has a turning point after only four quarters. Is this just an artefact of the first point in the series being unusually high?

Additional queries

The abstract cites I20-I25 as the definition of CHD for the study, while the methods section cites I20-I22. Which was used?

The question of mechanism that may underpin a change in MI rates is only touched on lightly in the paper. The introduction suggests a reduced exposure to second-hand smoke. It is possible that a reduction on smoking prevalence, or intensity could also follow from the smoking ban. Are you able to refer to any data on smoking rates for England or the Liverpool area?

The paper notes an increase in CHD admissions overall, at the same time as the reduction in MI admissions. Given the substantial size of reduction in MI rates observed, are you able to comment on whether there have been any changes to coding systems, practices, or procedures during the study period? Has the reduction in MI admissions been offset by an increase in admissions for any other diagnosis within CHD? A table of admission numbers over time by individual diagnosis would be helpful.

Trend analyses have been conducted by sex and socio-economic status. I think an analysis by age group would also be instructive. As the data have been age-standardised, these data should be available for analysis.

The discussion notes a possible explanation for why a higher rate of decline of MI events was observed in this study compared with another was the inclusion of multiple events, and the possibility that the intervention had a greater effect in reducing repeat or relapse MIs. It would be helpful if you could quantify in the paper the number or proportion of MI cases that are new or repeat cases, as this seems to be an important point to clarify.

I think the y-axis origin in Figures 1 and 2 should be 0 not 50. This won’t affect the ability to discern the shape and magnitude of the series, and would better contextualise the size of the movement in the series.

Public health implications

The authors conclude their paper with comments on the generalisability of their findings to other health issues, such as alcohol, salt and saturated fat consumption. I would caution that every public health issue is different and there is a long track record of approaches that have worked in tobacco control not being translatable to alcohol or dietary interventions. I would also suggest that the track record on dietary salt and saturated fat is not strong. The recent Institute of Medicine report on dietary salt indicates there
is little evidence that reductions in dietary salt have had much population benefit. The reduction in dietary saturated fat has not seen a decline in obesity, diabetes or metabolic disease. I would suggest a more circumspect conclusion from the reported data would be to focus on the hypothesis at hand and the issue of smoking and CHD.

**VERSION 1 – AUTHOR RESPONSE**

Reviewer 1: Stanton A. Glantz,

Thank you for your advice and feedback. We offer the below notes in response to the specific points received.

1. First, there is no need to smooth the data before doing the regression analysis. Doing so artificially reduces the variance in the data and makes the fit look better than it is. There is also the chance that the period used for the moving average smoother could affect the assessment of break points in the data.
   a. 3-period moving averages were used, to help reduce the exaggerated effect that outlying points can have on trend analysis models when these points are very close to either end of the study period.

2. Second, the use of jointpoint regression with 5 breakpoints seems overkill, especially given that there is no reason to expect so many sudden changes in AMI rates. The authors would be better served to use the approach that most others have of modeling the possibility of a sudden change followed by a slope change after the law. Alternatively, they could allow for two slope changes, one immediately following the law and one some time later. (Which of these approaches is most reasonable can be better assessed when considering the unsmoothed data.)
   a. One of the benefits of the Joinpoint analysis was that it is able to use a Bayesian statistical approach to select the optimum number of breakpoints for the trend. The Bayesian Information criterion method (BIC) approach finds the model with the best fit by penalising the cost of extra parameters, favouring trends with fewer segments. We selected this method as it tends to balance more conservative results with providing an adequate fit to the observed data.

3. Estimating the incident rate ratios by comparing two year periods before and after the law is a poor way to do the analysis because it does not account for any underlying secular trends. It would be much better to treat the data as a time series and directly estimate the IRR using a negative binomial regression. Doing so would also collapse the two different analyses into a single analysis, which would shorten and simplify the paper.
   a. The crude rate ratios comparing an early period and late period we felt was useful to contextualise some of the findings for use in comparison with other studies which have a similar ‘headline’ figure, and also to include another way by which different strata (e.g. socioeconomic) may be compared with each other within our study.
   b. We accept the clear limitations that such an incident rate calculated using this method will have, certainly in regards to accounting for background trends. Thus we have made the changes so that where these rate ratios are mentioned, we make the limitations clear and/or provide appropriate justifications for their use.

4. The comparison across income groups also needs to be integrated into the overall interrupted time
series analysis, using and ordinal variable if possible.
a. The data used for the income groups is identical to the data included in the main analysis, only recategorised into different strata (i.e. by socioeconomic status as opposed to by sex).

5. The authors do not need to be so defensive about the use of interrupted time series analysis (compared with randomized controlled trials). As the authors note, randomized controlled are impossible in such studies. Interrupted time series studies are widely used to assess causality in evaluating public policies.
Many abbreviations (MI, CHD, HES, NHS, etc) need to be defined the first time they are used. The authors needs to say precisely when in 2007 the law took effect and include this date in their analysis.
Page 2, line 10: This statement understates the strength of existing evidence. Doing so is not necessary to justify this work.
Page 2, line 42: “Is” not “appears to be.”
Page 3, line 20: Drop this bullet. It implies that the authors did something radically different from others, when in fact they did.
Page 3, line 37: Drop this bullet.
Page 4, line 51: Saying health inequalities “were not widened” presumes that they would be. “Were not affected” would be a more neutral statement.
a. All the above done.

6. Page 9, line 23: For the reasons stated above, the lab period is an artifact of the way the analysis was constructed. In addition, the statement of a 3-6 month “lag” is not an accurate representation of what the authors found; they found changes starting immediately. It was several months before the changes flattened out. But, as noted above, this interpretation could be an artifact of the way the model was constructed.
a. References to a ‘lag period’ have been removed as required.

7. Page 10, line 52: This paper is worth discussing in the context of the authors’ findings, since if finds similar results using an entirely different population and different methodology: Tobacco control policies are egalitarian: a vulnerabilities perspective on clean indoor air laws, cigarette prices, and tobacco use disparities. Dinno A, Glantz S. Soc Sci Med. 2009 Apr;68(8):1439-47. doi: 10.1016/j.socscimed.2009.02.003. Epub 2009 Mar 11. PMID: 19282078.
a. Paper has been added to the discussion and cited.

8. Page 11, line 25: Drop this paragraph for reasons discussed previously in this review.
Page 11, line 42: Same comment.
a. Done

Reviewer 2: David Lawrence

Thank you for your advice and feedback. We offer the below notes to support the specific points received.

1. Statistical methods: There are important details missing about calculation of rates, derivation of denominators, calculation of moving averages. There is no justification for why a joinpoint analysis has been chosen, whether the data exhibit seasonality or auto-correlations, and why an ARIMA based analysis was not undertaken.
a. We have brought in an additional author to apply our data to an ARIMA model, and included this in our method, analysis, results and discussion. We felt that this would still run alongside our primary
Joinpoint analysis however, as tests on the suitability of the models for use on our dataset suggested that it would not be suitable as a sole analysis method, given the relatively small numbers used. We do feel, however, that the inclusion of the ARIMA model in our study certainly adds weight to the analysis.

2. Comments on alcohol, salt and dietary fat in the final section “public health implications” are not supported by data.
   a. We have adjusted the text to be more cautious in dealing with issues outside of smoking and CHD in particular and, where we have done so, made sure to include references.

3. The supplied STROBE statement is far too light in detail. Most of the questions are answered “time trend study, not applicable”. I think that using admin data in this way creates a type of restrospectively assessed population-based cohort study, and the cohort study questions should be applicable in most instances. Just because the data are population based doesn’t mean there aren’t eligibility criteria, case ascertainment, missing data or that numbers of events shouldn’t be reported.
   a. We have completed the STROBE questionnaire in more detail as requested.

4. While there may have been other options for implementing the smoking ban policy that would provide a stronger test of their impact, such as introducing the legislation at different times in different places and comparing changes in outcomes between those places, a before-after comparison is appropriate for a national implementation of this type of policy action. Clearly, observational data such as these do not prove cause and effect. However, in situations where fully randomised designs are not possible, or weren’t implemented, observational data may form an important component of the evaluation of new policy or programme initiatives.
   a. We have acknowledged this and removed the unnecessary passages from the manuscript.

5. I would suggest including more detail in the introduction. For international readers it would be helpful to have some description of the nature of the smoking bans, when they were introduced, and what they covered. It would also be helpful to have some context on the health services in the Liverpool area. I am not familiar with the HES data base. How many health services does it cover, and what is the population of the region?
   a. We have included additional description on the background of the smoking ban, HES data, PCTs etc. for international readers who may not be familiar with these terms.

6. The authors have undertaken a piecewise linear regression analysis using the Joinpoint software. There is no discussion in the paper as to why this particular technique has been chosen, and whether the data meet the conditions for its use. My understanding of the Joinpoint programme is that it was originally developed for describing long-term trends in cancer incidence from annualised data, and that it assumes there is no autocorrelation within the time series. I would have thought that an ARIMA model of some type would have been the most likely first choice of analysis tool for this type of question. It would be helpful to explain why the Joinpoint approach has been chosen, and whether any tests of stationarity in the time series or autocorrelational structure were undertaken.
   a. The Joinpoint method is aimed particularly at detecting ‘changepoints’ in trends, and quantifying those changes. Given that we could ‘expect’ certain changes to take place during the period (once close to the ban, and perhaps another some time later) we felt it was appropriate to use this method.
   b. An ARIMA analysis was added, supporting our main findings. However, the small numbers meant that ARIMA would not be enough on it’s own so it runs alongside our JoinPoint analysis.

7. In the “Trend Analysis” section of the paper, the authors start by saying that plots of age-specific mortality rates were smoothed using 3 year moving averages. Firstly I presume they mean admission rates. It is not clear why 3 year moving averages have been applied to quarterly time series data. I presume this means a 12-point moving average? It is usually not recommended to adopt a moving average with an even number of points included as this can phase shift the series. Also with only 30 quarterly data points available for analysis, an important issue is treatment of the end points (i.e. the first six quarters and the last six quarters) in calculating the moving averages. No mention is made of how this was done, and I feel this detail needs to be included. I presume that the smoothed data was then used as input to the Joinpoint regression, but this is not specified in the text.
   a. 3-period moving averages were used, to help reduce the exaggerated effect that outlying points
can have on trend analysis models when these points are very close to either end of the study period.

8. The data in most of the tables and graphs appear to be rates rather than numbers of admissions, which is appropriate, but there is no mention where the denominators for rates have been sourced. A common challenge with use of administrative data of this type is how well population counts match the catchment areas of the health services. It is not entirely clear, but I presume the analysis has been restricted to people with a residential address in one of the 30 Liverpool wards, but does the HES data relate to use of health services in the Liverpool region, or can you account for use of health services out of region?

a. Although we do not think that out-of-area healthcare use of this type was significant, we were not able to analyse this.
b. The HES data looked at admissions in the City of Liverpool for those with a Liverpool residential address.

9. Were you restricted in data access to only using quarterly data? It seems you have enough events to create a monthly time series. I would think this would better allow you to test for seasonality, and might better help determine the lag between intervention and movement in the time series. This is quoted as being between 3 and 6 months, but this is difficult to judge from a quarterly time series.

a. Although it was certainly possible to perform a monthly analysis on the CHD data, the numbers were too small for the MI or socioeconomic strata data, thus it was grouped into quarters. For the sake of easier comparability, the CHD analysis was also done by quarter. For potential future expanded studies with a larger population, we would look first to see if we could perform a monthly time series.

10. Another concern with the Joinpoint approach is that it is entirely data driven. As you have a clear hypothesis as to whether there was a change in admission rates subsequent to the legislation, an ARIMA model that specifically tested this would be closer to the hypothesis of the study. An artefact of the Joinpoint approach is that it inserts turning points in the series that are unrelated to the study hypothesis. For instance, the top line in Figure 2 has a turning point after only four quarters. Is this just an artefact of the first point in the series being unusually high?

a. One of the benefits of Joinpoint was that it restricts the artefacts that may arise from the exaggerated effect of early/late outlying points. In this case Joinpoint did not consider changepoints within 4 points (one year) of either end of the study period.
b. Nevertheless, with relatively small numbers (MIs/socioeconomic strata) these artefacts are still more likely to occur.

11. The abstract cites I20-I25 as the definition of CHD for the study, while the methods section cites I20-I22. Which was used?

a. I20 to I22 was used. Incorrect references to this have been changed.

12. The question of mechanism that may underpin a change in MI rates is only touched on lightly in the paper. The introduction suggests a reduced exposure to second-hand smoke. It is possible that a reduction on smoking prevalence, or intensity could also follow from the smoking ban. Are you able to refer to any data on smoking rates for England or the Liverpool area?

a. This is certainly a possible cause for the reduced rates seen, and we were keen at the time to see if this could be linked with the data we had access to. Unfortunately the HES data we had access to could not link other behavioural factors such as smoking status. We have adjusted the manuscript to mention this.
b. This may be possible by using other data sources in conjunction, but unfortunately for this study we are unable to go back and do this.

13. The paper notes an increase in CHD admissions overall, at the same time as the reduction in MI admissions. Given the substantial size of reduction in MI rates observed, are you able to comment on whether there have been any changes to coding systems, practices, or procedures during the study period? Has the reduction in MI admissions been offset by an increase in admissions for any other diagnosis within CHD? A table of admission numbers over time by individual diagnosis would be helpful.
a. ICD-10 codes have undergone major changes in early 2004, and after 2012, outside of our study period, making at an ideal time period to use this coding system. Small adjustments and corrections have taken place over this span, but they are not thought to have a significant effect on coding records overall.
b. The admissions data was available by overall ICD code (I20, I21 or I22) but not by subdiagnosis. 
c. Data for admissions by MI, and by all CHD overall are available in Tables 1 & 2.
14. Trend analyses have been conducted by sex and socio-economic status. I think an analysis by age group would also be instructive. As the data have been age-standardised, these data should be available for analysis.
   a. Analysis by age group for our study population would have yielded too small numbers unfortunately, especially for the younger brackets, and was not performed. Looking at a wider population (E.G. Regional) we would look to include analysis by age.
15. The discussion notes a possible explanation for why a higher rate of decline of MI events was observed in this study compared with another was the inclusion of multiple events, and the possibility that the intervention had a greater effect in reducing repeat or relapse MIs. It would be helpful if you could quantify in the paper the number or proportion of MI cases that are new or repeat cases, as this seems to be an important point to clarify.
   a. Again, unfortunately the HES data at the time did not allow us to look at it in this manner. In some cases, this was implicit e.g. a diagnosis code for 'Previous MI' – however we felt that we needed corroboration with other patient identifiers (to link episodes by patient), before we could reliably analyse the data in this way.
16. I think the y-axis origin in Figures 1 and 2 should be 0 not 50. This won’t affect the ability to discern the shape and magnitude of the series, and would better contextualise the size of the movement in the series.
   a. This has been changed in the manuscript as requested.
17. The authors conclude their paper with comments on the generalisability of their findings to other health issues, such as alcohol, salt and saturated fat consumption. I would caution that every public health issue is different and there is a long track record of approaches that have worked in tobacco control not being translatable to alcohol or dietary interventions. I would also suggest that the track record on dietary salt and saturated fat is not strong. The recent Institute of Medicine report on dietary salt indicates there is little evidence that reductions in dietary salt have had much population benefit. The reduction in dietary saturated fat has not seen a decline in obesity, diabetes or metabolic disease. I would suggest a more circumspect conclusion from the reported data would be to focus on the hypothesis at hand and the issue of smoking and CHD.
   a. We have ensured that we take describing issues or implications outside of CHD or smoking very cautiously. We have also removed some overly assumptive generalisations, ensuring that, more often, we concentrate on drawing conclusions about CHD/smoking specifically.
findings.

Just a couple of minor points on the new material added:

p12 - "The sample size could also mask the real effect of the smoking ban" seems clumsily worded. Perhaps you mean to say something like "Because of the sample size, the study may be underpowered to adequately estimate the real effect of the smoking ban".

"Joinpoint regression seems to be a more adequate and robust methodology" - I'm not sure on what basis this claim is supported - I would suspect the power of the analysis is largely determined by the sample size for both techniques.

p. 7 "man trend analysis"?