PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Identification of myocardial infarction type from electronic hospital data in England and Australia: a comparative data linkage study
AUTHORS	Nedkoff, Lee; Lopez, Derrick; Goldacre, Michael; Sanfilippo, Frank; Hobbs, Michael; Wright, Lucy

VERSION 1 – REVIEW

REVIEWER	Marek Gierlotka Silesian Centre for Heart Diseases, Zabrze, Poland No Competing Interest
REVIEW RETURNED	10-Sep-2017

GENERAL COMMENTS	In a large datasets of electronic hospital data the authors tested if it is possible to monitor trends in myocardial infarction and its subtypes (NSTEMI, STEMI) at the national population level. They point to some very important limitations of this approach. The paper is well designed and well written. I have no major remarks. The topic is interesting, especially nowadays, when a lot of papers use hospital or administrative data to study epidemiological data.

REVIEWER	Michael Falster Centre for Big Data Research in Health, UNSW Sydney, Australia
REVIEW RETURNED	20-Sep-2017

GENERAL COMMENTS	This paper sought to explore the utility of hospital admissions data to explore trends in types of myocardial infarction, identified using the 4th digit of ICD-10 codes, in population-level linked datasets in Western Australia and the UK. This was a well written paper, and a very useful exploration of the impact changes in coding standards can have on the ability to identify specific types of morbidities. The study is limited by a lack of clinical data to validate the type of MI diagnosis, but is also upfront about this limitation, and still raises valid concerns about the reliability of ICD-10 hospital data. 1. It would seem appropriate that key implications for use of hospital admissions data for identifying and monitoring MI types were conveyed in the abstract. 2. Given there have been changes in coding practice over time, I question whether an average annual percent change is an appropriate statistic. For example, Table 1 reports a significant decreasing trend in Subsequent MI in England, but Figure 1 shows a relatively flat trend with a sharp reduction in 2012 – the 'average trend' reflects a change in coding standards at a single point in time.

	Alternate analytic approaches, such as an interrupted time series analysis or joinpoint regression, would perhaps be a better means to explore trends in types of MI, independent to the changes in coding practice. However, this may be outside the scope of the current paper and the (appropriately cautious) conclusions made. 3. I don't think Figure 2 is needed – this presents the same information as Figure 1 with two of the categories combined – information which could easily be inferred. 4. Consistent colours between figures (e.g. for STEMI) should be used to aid readability. 5. Presentation of time series data could be enhanced by including markers indicating chronological changes in coding practice within each setting. For example, the sharp reduction in Subsequent MI in England in 2012 clearly represents the change in coding standards described on page 6 line 13, but this link is not articulated in the paper. 6. Ethics approvals for the use of linked data in England have not been stated.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Comment:
In a large datasets of electronic hospital data the authors tested if it is possible to monitor trends in myocardial infarction and its subtypes (NSTEMI, STEMI) at the national population level. They point to some very important limitations of this approach.
The paper is well designed and well written. I have no major remarks.
The topic is interesting, especially nowadays, when a lot of papers use hospital or administrative data to study epidemiological data.

Response:
Thank you for the comments on our paper.

Reviewer 2
Comments and Responses:
1. It would seem appropriate that key implications for use of hospital admissions data for identifying and monitoring MI types were conveyed in the abstract.

2. Given there have been changes in coding practice over time, I question whether an average annual percent change is an appropriate statistic. For example, Table 1 reports a significant decreasing trend in Subsequent MI in England, but Figure 1 shows a relatively flat trend with a sharp reduction in 2012

– the 'average trend' reflects a change in coding standards at a single point in time. Alternate analytic approaches, such as an interrupted time series analysis or joinpoint regression, would perhaps be a better means to explore trends in types of MI, independent to the changes in coding practice. However, this may be outside the scope of the current paper and the (appropriately cautious) conclusions made.

Response:
The reviewer is correct that presenting average annual % changes from Poisson regression models may not adequately reflect trends prior to 2011, particularly in the English data where changes in trends are noted from 2012. We have therefore re-run separate Poisson regression models for the period 2000-2011, for each MI subtype. We have inserted the following into the text of the Results section, which highlight the differences between trends for the overall period vs 2000-2011, for subsequent and unspecified MI (the attenuation of trends for STEMI and NSTEMI was limited compared with the other two groups when the same approach was taken):

"The age- and sex-adjusted average change in subsequent MI rates in England was -4.6%/year (95% CI -4.5%, -4.7%) for the overall study period. However when the period was restricted to 2000 to 2011, there was a marginal increase in rates (+0.9%/year, 95% CI +0.7, +1.0%). Similarly for unspecified MI, there was an attenuation of the upward trend in rates when the period was restricted (+0.6%/year, 95% CI +0.5%, +0.7%, 2000 to 2011). There was a sharp downturn in rates of subsequent MI in England in 2012, declining from 43/100,000 person-years to 1/100,000 person-years by 2013, with concomitant increases in STEMI and unspecified MI rates." (page 8)

3. I don't think Figure 2 is needed – this presents the same information as Figure 1 with two of the categories combined – information which could easily be inferred.

Response:
Agreed – this figure has been removed, and the description of the analysis retained in the text.

4. Consistent colours between figures (e.g. for STEMI) should be used to aid readability.

Response: The colours have been changed in Figure 3 so are now consistent across figures.

5. Presentation of time series data could be enhanced by including markers indicating chronological changes in coding practice within each setting. For example, the sharp reduction in Subsequent MI in England in 2012 clearly represents the change in coding standards described on page 6 line 13, but this link is not articulated in the paper.

Response: The two main coding changes that occurred during our study period were the introduction of STEMI/NSTEMI into the nomenclature of ICD-10-AM in 2004, and the change in coding standard for the I22 code in England in 2012. To ensure that the link between these changes and the study results are clear, we have added an extra small paragraph to the Discussion to highlight these:

"The change to coding standards in Australia in 2004, where STEMI and NSTEMI were first listed alongside the relevant I21 codes, appear to have had little impact on trends in these subtypes. The trends in proportions of NSTEMI and STEMI cases predated the inclusion of these descriptors in the Australian version of ICD-10. In contrast, changes to the coding standards in England in 2012 appear to have contributed to a marked decline in the proportion of cases coded as subsequent MI from 2012, with most of the shift in cases being towards higher proportions of STEMI and unspecified MI, rather than NSTEMI." (page 10)

6. Ethics approvals for the use of linked data in England have not been stated.
We have added a statement to the Ethics acknowledgement at the end of the manuscript regarding the England ethics approval as requested ("Ethical approval was obtained from the Central and South Bristol Multi-Centre Research Ethics Committee (04/Q2006/176) for the building and analysis of the English national record-linked data by the Unit of Health-Care Epidemiology and have additionally added a funding statement relevant to the English data at the end of the manuscript.

VERSION 2 – REVIEW

REVIEWER	Michael Falster Centre for Big Data Research in Health, UNSW Sydney, Australia
REVIEW RETURNED	29-Sep-2017

GENERAL COMMENTS	No further comments. The authors have addressed all my concerns, and this will make an excellent publication.