Author's response to reviews

Title: Data validity of routine electronic hospital admissions data with Scottish Stroke Care Audit (SSCA) data for identifying stroke in the Scottish population.

Authors:

Melanie Turner (m.e.turner@abdn.ac.uk)
Mark Barber (mark.barber@nhs.net)
Hazel Dodds (hazeldodds@nhs.net)
Martin Dennis (martin.dennis@ed.ac.uk)
Peter Langhorne (Peter.Langhorne@glasgow.ac.uk)
Mary-Joan Macleod (m.j.macleod@abdn.ac.uk)

Version: 3
Date: 7 October 2015

Author's response to reviews: see over
Mary Joan Macleod MBChB, PhD, FRCP
Senior Lecturer
University of Aberdeen

6th October 2015

Dear Editors

Data validity of routine electronic hospital admissions data with Scottish Stroke Care Audit (SSCA) data for identifying stroke in the Scottish population.

We thank the reviewers for their comments regarding our paper. We have addressed these as detailed below.

Reviewer 1:

1. It would be helpful to specify in the abstract that the SMR01 is a discharge database.

   *We have changed the Background in the Abstract (pg2, line 25) to “In Scotland all non-obstetric, non-psychiatric acute inpatient and day case stays are recorded by an administrative hospital discharge database, the Scottish Morbidity Record (SMR01).”*

2. It would be helpful to clarify who inputs data to the SSCA, perhaps add this detail on page 3, (this detail is described for the SMR01 on page 4, but is not presented in relation to the SSCA until the discussion page 7 line 157). Perhaps clearly outline the processes for both earlier in the paper.

   *We have inserted the following sentence (pg 4, line 71) “Ascertainment and input of stroke cases in the SSCA is a process carried out by trained but non-specialist audit co-ordinators supported by clinicians and based in stroke units within hospitals.”*

3. It would be helpful for other countries, who are developing such databases, to have some detail regarding the resourcing of the SMR01 and the SSCA? The authors could consider including some brief details.
We have inserted the following sentence (pg 4, line 70 and line 80) “It is resourced by ISD of National Services Scotland part of NHS Scotland.”

4. Perhaps add a brief description in the methods section about “linkage analysis”.

We have changed the following sentences (pg 4, line 90) “Data for 2010-2011 were requested from ISD in Scotland, where a data analyst linked records between the two data sources” and (pg5, line 101) “A data analyst within NHS Grampian Health Intelligence produced a linked dataset”.

5. Page 4, line 89, who conducted the case note review?

We have inserted the following (pg 5, line 102) “and a research fellow carried out the case-note review.”

6. I feel that the conclusion could be strengthened by adding details regarding the SSCA process in relation to trained support staff in the stroke units, the importance of documentation and coder experience.

We have changed the discussion to highlight this (pg8 line 179) “In comparison the ascertainment of stroke cases in the SSCA is a process carried out using detailed data collection forms by trained but non-specialist audit co-ordinators supported by clinicians and based in stroke units within hospitals. Quality assurance is carried out regularly by ISD staff.”

Reviewer 2:

1. The title is poorly written. It is redundant (‘data’ was used 3 times) and does not represent what this manuscript is about.

We have changed the title to “Agreement between routine electronic hospital discharges and Scottish Stroke Care Audit (SSCA) data in identifying stroke in the Scottish population.”

2. The motivation or background of this research is not really compelling. Why is it so important to identify correct number of stroke cases in general? If the limitation of hospital discharge coding data is known so well, why not use the SSCA data? Is SMR01 data cheaper to obtain?
We have inserted the following sentences (pg 3, line 62) “In Scotland, administrative hospital discharge data is used to determine the level of NHS resource needed for a specific condition. It is therefore important to determine if this contains an accurate account of stroke occurrence. We aimed to assess the discrepancies between a Scotland-wide stroke audit used to assess performance and an administrative hospital discharge database collected during 2010-2011.”

3. There is a section named ‘reference population’ under method, which is not appropriate. SSCA data are by no means ‘population’. It is just another dataset. I suggest ‘reference data’.

We have changed this to “Scottish Stroke Care Audit” (pg3, line 68).

4. This paper says SSCA data or SMR01 identified XX% of stroke cases, but how do they know the denominator, which is the true number of stroke cases? Are they assuming all stroke cases were included either in SSCA and SMR01? That is probably unrealistic and untestable assumption. Given the data, all they can say is that SSCA data or SMR01 data identified XX% or less of stroke cases, because there must be stroke cases that were not reported elsewhere.

We have changed all %’s in the paper to actual numbers to make it clearer. We have inserted the following sentences in the Discussion (pg 9, line 183) to make it clear there may be records that have not been admitted to hospital. “Records from both SSCA and SMR01 are limited to patients that are admitted to hospital having suffered a stroke. Therefore, people who have a stroke or TIA but are never admitted to hospital will be missed by both sources. Any future work in this area could use both further sources of data (such as out-of-hospital stroke mortality data) and the capture-recapture technique to provide more reliable estimates of the total number of stroke occurring in the population.”

5. The numbers in the results section do not make sense, or I am not understanding them correctly because of poor explanation. For example, the authors said there were 22416 entries in SSCA and 22000 entries in SMR01 and when those two datasets were
combined, there were 22601 entries). This can happened only when these two datasets have considerable overlap, but the next sentence says only 55.5% (n=12552) records were found in both datasets. How come? Diagrams for # of cases in each dataset will help.

We have removed the sentence “When combining both the SSCA and SMR01 datasets there were a total of 22 601 entries recorded in Scotland between 2010 and 2011.”

By changing all %’s to actual numbers we believe that the explanation of the data is easier. The numbers of cases in each dataset is shown clearly in Figure 1.

6. In the discussion, the authors said the % of strokes identified in this study was an increase from previous linkage data. It would be great if there’s any discussion why it increased.

We have added the following sentence to the Discussion (pg8, line 166) “These increases may be due to training of coding staff and an improvement in coding on hospital discharge letters. In addition the inclusion of G45 as a main diagnosis code and including all four codes in secondary diagnoses fields in this study will have contributed.”

7. The last sentence in the discussion section has nothing to do with this research.

If the referee is referring to the sentence beginning on page 9 line 195, we have added in clarification regarding the importance of differentiating transient from disabling events “Although secondary prevention is similar for both TIA and stroke, basing economic models on datasets which do not distinguish between transient and disabling events, for example, may result in inaccurate cost calculations.” If the reference is to the sentence beginning on page 9 line 203, we would respectfully disagree, as we do feel that we provide evidence to support disease registries as more robust sources of data than routine administrative coding data.
8. Overall the quality of writing is poor. There are numerous errors and grammatical mistakes.

We apologise for any grammatical errors: we have reread the manuscript and hopefully removed any ambiguities.

Additional Editorial Requests:

1. Please confirm whether informed consent, written or verbal, was obtained from all participants. If the need for consent was waived by an IRB or is deemed unnecessary according to national regulations, please clearly state this, including the name of the IRB or a reference to the relevant legislation.

We have expanded the ethical approval section (page 5 line 103): “While individual patient consent was not obtained, the Scottish Stroke Care Audit works within the “NHS Code of Practice on Protecting Patient Confidentiality” which incorporates the requirements of statute and common law including the Data Protection Act, the Human Rights Act and the Adults with Incapacity (Scotland) Act. The study was approved by Scotland A Research Ethics Committee, Ref. No.=10/MRE00/76 and the Privacy Advisory Committee of ISD, NHS Scotland, Ref 76/11.”

Please detail funding in the Acknowledgements and remove the section Sources of funding from the manuscript.

We have changed as requested.

We hope that we have addressed all reviewer and editorial requests satisfactorily.
Yours sincerely

Mary Joan Macleod on behalf of the authors