PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Differences between self-reported and verified adverse cardiovascular events in a randomized clinical trial.
AUTHORS	Bolland, Mark; Barber, Peter; Doughty, Rob; Grey, Andrew; Gamble, Greg; Reid, Ian

VERSION 1 - REVIEW

REVIEWER	Wang, Lu Brigham and Womens Hospital, Medicine No competing interests
REVIEW RETURNED	14-Dec-2012

GENERAL COMMENTS	This study examined the accuracy of adverse cardiovascular events identification in a 5-y clinical trial. The adverse events were originally identified from self-reports and national hospital discharge database. After adjudication, a low rate of verification plus a high rate of unreported events was found. The authors concluded that researchers should adjudicate self-reported events and hospital discharge codes and identify unreported events to improve adverse event report accuracy. Adverse event report is important in clinical trial for evaluation of the efficacy and safety of intervention. However, because adverse events were secondary endpoint at the most, the accuracy of events identification is always a concern. The current study addressed this common and important issue in clinical trials and provided evidence that the accuracy of self-reported events as well as hospital data-identified events is low, and many events were not reported to investigators. The manuscript is well-written, the analysis is straight forward, and data is informative. I have some suggestions that hopefully may further improve the paper. 1. The accuracy of self-reported adverse events may vary by study population. For example, well-educated individuals with medical knowledge and conscience, such as health professionals, may report disease diagnosis or event incidence more accurately. The authors may consider providing more information about the characteristics of their study population and discuss how their study findings may be generalizable to other populations. 2. When the investigators collected data on adverse events in this trial, did they ask specific questions about events, such as MI, stroke, and TIA? Or participants were asked to report any events that occurred during the course of the trial in an open-ended question? This detail needs to be clarified in Methods. 3. Did the authors consider any analysis to identify the factors that may affect the accuracy of self-reported events? The authors discussed that participants in this trial were older, had more co-

	morbidity and more cognitive impairment, which may explain the low accuracy of self-reports. Was older age associated with a lower accuracy? What co-morbidity affected the accuracy the most? 4. The authors stated that there were minor changes in relative risk for MI and stroke with calcium supplements by different event sources or level of adjudication. However, the point estimates for relative risk changed from 1.43 (the lowest) to 2.24 (the highest) and p value changed from 0.0099 to 0.27. This change is not minor or negligible. In particular, since the accuracy of self-reports and hospital discharge coding is low, and unreported event rate is high, the analysis on association between treatment and adverse events should emphasize on verified events, which indeed showed a statistically non-significant relation.

REVIEWER	Larraitz Arriola Public Health Department. Basque Government. Spain. I have no competing interests
REVIEW RETURNED	24-Dec-2012

THE STUDY	There is not standardised questionnaire for MI or stroke in this study. One of the sources compared (self report) is not described and I have serious doubts that they are really comparing self-reported questionnaires and clinical records. What was asked to the patients? How and who asked and when?
GENERAL COMMENTS	or my understanding, self report information in epidemiological studies means the information gathered from the subject with an standardised questionnaire. Asking to the participant not only if they had had the event but also when with different questions. Thai is what I understand self report information. And that is what other studies about comparision of sel report information with other sources of information do. Even the papers cited by the authors. In their study they do not ask the participants about the event under study. They just expect the participant spontaneously talk about the event. I do not think that could be compared with the information gathered from self report questionnaires.

REVIEWER	Annie Britton Senior Lecturer in Epidemiology University College London UK No competing interests
REVIEW RETURNED	06-Feb-2013

THE STUDY	I would suggest that the authors add when the trial and follow-up took place
GENERAL COMMENTS	The use of hospital records linkage to identify adverse events in clinical trials and epidemiological research is becoming increasingly important. This paper adds nicely to the literature comparing the agreement between self-report, hospital discharge records and death certification. It may help readers to know when the trial took place (I'm assuming this to be around 1999 - 2000 with follow-up till 2005, as the trial results were published in 2006?) and the authors may want to

	comment on whether the completeness and accuracy of hospital records has improved over time in New Zealand, as is thought to be the case in the UK (Burns EM, Rigby E, Mamidanna R et al J Public Health 2011). The paper does not discuss sub-typing of stroke events and this may not have been relevant for the adverse event ascertainment in this trial. We previously showed that relying on self-reported information or hospital discharge summary data is limited for sub-tying of stroke and that manual extraction of information from hospital records provides supplementary information required for sub-typing (Britton et al, BMC Medical Research Methodology 2012).

VERSION 1 – AUTHOR RESPONSE

Reviewer: Lu Wang
This study examined the accuracy of adverse cardiovascular events identification in a 5-y clinical trial. The adverse events were originally identified from self-reports and national hospital discharge database. After adjudication, a low rate of verification plus a high rate of unreported events was found. The authors concluded that researchers should adjudicate self-reported events and hospital discharge codes and identify unreported events to improve adverse event report accuracy.

Adverse event report is important in clinical trial for evaluation of the efficacy and safety of intervention. However, because adverse events were secondary endpoint at the most, the accuracy of events identification is always a concern. The current study addressed this common and important issue in clinical trials and provided evidence that the accuracy of self-reported events as well as hospital data-identified events is low, and many events were not reported to investigators. The manuscript is well-written, the analysis is straight forward, and data is informative. I have some suggestions that hopefully may further improve the paper.

1. The accuracy of self-reported adverse events may vary by study population. For example, well-educated individuals with medical knowledge and conscience, such as health professionals, may report disease diagnosis or event incidence more accurately. The authors may consider providing more information about the characteristics of their study population and discuss how their study findings may be generalizable to other populations.

Response:
The study population was healthy post-menopausal women (mean age 74) and free from other major medical conditions. The population is described briefly in the Methods section, and fully described in References 3 and 4. We think this description is reasonable for the purposes of these analyses. We already discussed that our findings were similar to those in other studies, and possible reasons for important differences in Paragraph 5 of the Discussion.

2. When the investigators collected data on adverse events in this trial, did they ask specific questions about events, such as MI, stroke, and TIA? Or participants were asked to report any events that occurred during the course of the trial in an open-ended question? This detail needs to be clarified in Methods.

Response:
Questions about events during the study were open-ended. We have added ‘or illnesses’ to the relevant sentence in the Methods.

‘Adverse events were recorded at each visit but questions about specific symptoms or illnesses were not asked.’

3. Did the authors consider any analysis to identify the factors that may affect the accuracy of self-reported events? The authors discussed that participants in this trial were older, had more co-morbidity and more cognitive impairment, which may explain the low accuracy of self-reports. Was older age associated with a lower accuracy? What co-morbidity affected the accuracy the most?

Response:
We have added the following paragraph to the Results to address this comment (and an explanatory sentence to the Methods).

‘Differences between baseline characteristics were compared using t-tests for continuous variables, and Fisher’s exact test for categorical data.’

‘Finally, we assessed possible relationships between baseline characteristics and discrepancies between self-reported and final verified events. Women in whom there were discrepancies were older ($P<0.001$), and were more likely to have reported at the baseline study visit having had a previous MI ($P=0.013$) or previous TIA or stroke ($P=0.035$) than women without such discrepancies. However, there were no differences in physical activity ($P=0.7$) or other co-morbidities between the groups ($P>0.1$ for hypertension, dyslipidemia, diabetes).’

4. The authors stated that there were minor changes in relative risk for MI and stroke with calcium supplements by different event sources or level of adjudication. However, the point estimates for relative risk changed from 1.43 (the lowest) to 2.24 (the highest) and p value changed from 0.0099 to 0.27. This change is not minor or negligible. In particular, since the accuracy of self-reports and hospital discharge coding is low, and unreported event rate is high, the analysis on association between treatment and adverse events should emphasize on verified events, which indeed showed a statistically non-significant relation.

Response:
The point we were making is that regardless of the event source or level of adjudication, the differences between the groups were sufficient to raise concerns. Thus, relative risks of clinically important adverse events of either1.4 or 2.2 are sufficient to warrant further consideration. We think that interpretation of adverse events in clinical trials should not be based solely on p-values and statistical significance for several reasons- trials are not generally powered to detect differences in adverse event rates; numerous adverse events are assessed so statistically significant differences occur commonly by chance; and finally, regardless of p-value, substantial differences raise concerns. Our view is in accordance with guidance from the European Medicines Agency: “In the case of adverse effects, p-values are of very limited value as substantial differences will raise concern” (Committee for Proprietary Medicinal Products. Points to consider on multiplicity issues in clinical trials. 2002:CPMP/EWP/908/99).

We have changed the wording in the Results from “and relatively minor changes in risk for the different event sources or levels of adjudication, although ...” to “although the relative risks changed in some cases depending upon event source or level of adjudication, and …”

The Reviewer is aware that the analyses conducted in this trial, which had limited power to detect a

modest effect of treatment on vascular risk, were not reported as definitive, but formed the basis for subsequent analyses of pooled data from RCTs of calcium supplements that confirmed the adverse vascular effects (refs 16,17 in text)

Reviewer: Larraitz Arriola
There is not standardised questionnaire for MI or stroke in this study. One of the sources compared (self report) is not described and I have serious doubts that they are really comparing self-reported questionnaires and clinical records.
What was asked to the patients? How and who asked and when?

Response:
We provided this information in the Methods. Participants were seen every six months and asked open ended questions about their health, by a member of the research team who was blinded to treatment allocation. Questions about specific symptoms were not asked.

For my understanding, self report information in epidemiological studies means the information gathered from the subject with an standardised questionnaire. Asking to the participant not only if they had had the event but also when with different questions. Thai is what I understand self report information. And that is what other studies about comparision of sel report information with other sources of information do. Even the papers cited by the authors. In their study they do not ask the participants about the event under study. They just expect the participant spontaneously talk about the event. I do not think that could be compared with the information gathered from self report questionnaires.

Response:
We agree with the Reviewer that the results from self-reports in response to open-ended questioning in a clinical trial might differ from responses to specific questions in a questionnaire, or from a mixture of open and closed questions use to generate a clinical medical record. We are not aware of any research on this topic. We have added this text to the Discussion.

“Our findings are broadly consistent with this previous research, although it is uncertain whether self-reports in response to open-ended questions in a clinical trial are comparable to responses to questionnaires or medical records based on a mixture of open- and closed-ended questions.”

Reviewer: Annie Britton
I would suggest that the authors add when the trial and follow-up took place

Response:
The trial started in 1998 and was completed in 2005. Cardiovascular event adjudication took place between 2006-7. This information has been added to the Methods.

The use of hospital records linkage to identify adverse events in clinical trials and epidemiological research is becoming increasingly important. This paper adds nicely to the literature comparing the agreement between self-report, hospital discharge records and death certification.
It may help readers to know when the trial took place (I’m assuming this to be around 1999 - 2000 with follow-up till 2005, as the trial results were published in 2006?) and the authors may want to comment on whether the completeness and accuracy of hospital records has improved over time in New Zealand, as is thought to be the case in the UK (Burns EM, Rigby E, Mamidanna R et al J Public

Health 2011).

Response:
This is an important point, but we are not aware of similar data for New Zealand. We have addressed this point in the Discussion.

“In the United Kingdom, accuracy of diagnostic coding has improved substantially in recent years,15 although no similar data are available for New Zealand. It is possible that the magnitude of differences between adjudicated events and hospital discharge coding might have changed since our study was undertaken.”

The paper does not discuss sub-typing of stroke events and this may not have been relevant for the adverse event ascertainment in this trial. We previously showed that relying on self-reported information or hospital discharge summary data is limited for sub-tying of stroke and that manual extraction of information from hospital records provides supplementary information required for sub-typing (Britton et al, BMC Medical Research Methodology 2012).

Response:
We did not specifically address this issue in our paper, but agree that it is worth raising in the Discussion.

“For sub-typing of stroke, review of hospital records adds useful data to self-reports and discharge coding.14”