

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	The effect of a Dementia Education Intervention on the confidence and attitudes of General Practitioners in Australia. A pre-test post-test study.
<b>AUTHORS</b>	Mason, Ron; Doherty, Kathleen; Eccleston, Claire; Winbolt, Margaret; Long, Marita; Robinson, Andrew

### VERSION 1 – REVIEW

<b>REVIEWER</b>	VIMALA RAMOO University of Malaya
<b>REVIEW RETURNED</b>	21-Aug-2019

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to review this article. It has been pleasure to read through the content.</p> <p>The followings are some comments/suggestions to the authors:</p> <p>Title – describes well the scope and outcomes of the study. Introduction – Clearly explained and linked to the study purposed. Methods:</p> <ul style="list-style-type: none"><li>• Study settings – were general, no details on where the study actually took place and how many sample obtained from each state and each time.</li><li>• The sample size needs justification whether it represents the study population and power of the study.</li><li>• Sampling methods used – needs more detail explanation as well. What was the response rate?</li><li>• Study instrument – too brief, although it has been reported in another paper, it needs some explanation in context of this paper for readers' understanding.</li><li>• Data collection – who collected the data; any possibility for response bias?</li><li>• Statistical test – no explanation on why non-parametric tests were used. In non-parametric tests, it would be more appropriate to use median rather than mean to describe the values of the variables.</li><li>• The effect size values (Cohen's d) in Table 2, noted most of them are larger than 1.0, would they be any specific interpretation for it.</li></ul> <p>Results:</p> <ul style="list-style-type: none"><li>• Clearly interpreted and presented.</li><li>• Some thoughts – why the authors didn't explore the potential confounders that may have influenced the outcomes of the study. It is essential to identify the possible confounding factors and control them in the analysis – to show the improvements in the outcomes are actual effects of the intervention.</li></ul> <p>Discussion:</p> <ul style="list-style-type: none"><li>• Would there be any possibilities for response bias at post intervention phase as the respondents would be aware that they need to present positive answers after the educational sessions?</li></ul>
-------------------------	---

	<p>Having some discussion on this would benefit the future researchers.</p> <ul style="list-style-type: none"> <li>• To include the limitations of the study: for e.g. design – not having a control group, self-reported answers, actual changes in their performances and knowledge were not assessed, etc.</li> </ul> <p>References</p> <ul style="list-style-type: none"> <li>• Mostly were latest, adequate and appropriate.</li> </ul> <p>Overall</p> <ul style="list-style-type: none"> <li>• Well written and reads easily.</li> <li>• To check on consistency of font size (page 7, line 13-18) and format of the text and tables.</li> </ul>
--	--

<b>REVIEWER</b>	Mieke van Driel University of Queensland, Australia
<b>REVIEW RETURNED</b>	25-Aug-2019

<b>GENERAL COMMENTS</b>	<p>This paper addresses an important issue in general practice and assessing the impact of education is always worthwhile. The paper is well written, however, some important information is missing. The following points may benefit from more clarity.</p> <ul style="list-style-type: none"> <li>- there is information missing on how the training was delivered, apparently different for GPR and SV? was it delivered face to face to all? what was the fidelity, did everyone participate?</li> <li>- what was the sample frame of this study? how many GPR and SV were eligible, how many responded to the survey? how was the survey administered?</li> <li>- the authors claim in their summarising statements that they included a representative sample of Australian GPs? what is that statement based on? these are 2 different groups with very different demographic characteristics. It is likely that those involved in GP training as supervisors are more engaged and different from their non-SV peers?</li> <li>- the title is a bit misleading as this study is about registrars and established GPs, not only about GPs</li> <li>- regarding the analysis, I am not a statistician, but it seems to me that the measured data were not 'independent' as they were in fact paired (each participant filled out the pre and post survey); how is this accounted for? How to interpret the Cohen's d?</li> <li>- the results section is very long and repeats data in tables; for ease of reading this section could be condensed a bit.</li> <li>- the authors do not discuss any limitations of their study</li> <li>- what are the implications of the study findings? where to from here? the 'so what' factor hasn't really been adequately addressed. What do we know about the link between knowledge and attitudes and actual behavior in clinical practice? there is a wealth of literature on this topic and the authors might want to provide us with some insights in that regard to give the discussion more depth and meaning. what is the take-home message for those involved in GP training or CPD?</li> </ul>
-------------------------	---

<b>REVIEWER</b>	Ivan Buljan University of Split School of Medicine
<b>REVIEW RETURNED</b>	07-Sep-2019

<b>GENERAL COMMENTS</b>	Thank you for the opportunity to review this interesting article.
-------------------------	---

	<p>In the statistical analysis part, you report on using non-parametric procedures (Mann Whitney U test and Wilcoxon sign rank test), but also present your data as means with standard deviations, which is not appropriate. I would advise to present your data as medians with corresponding 95% confidence intervals around medians. I understand that, when presented as medians, the differences will be difficult to present because the range used is very narrow, but it is the recommended way of data presentation.</p> <p>Also, I would strongly advise to present median differences between groups (pre-and post- assessment), to determine the actual size of difference, instead of relative effect size.</p> <p>Cohen's d is not an appropriate measure of effect size in non-parametric testing, but for your specific situation you could use r effect size. It is easily calculated using the JASP statistical software, which is free to use and similar to SPSS, which you are currently using in your research.</p> <p>However, if the distribution is not significantly different from normal distribution, it would be possible to use t-tests in testing differences. However, in that case you should present data with means and corresponding 95% confidence intervals, mean difference and Cohen's d effect size.</p> <p>Please format tables in your Results section (put your attention to Table 1 especially).</p> <p>Please use appropriate symbol for chi squared, instead of using x. Once again, thank you for the opportunity to review this paper.</p>
--	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name

VIMALA RAMOO

Institution and Country

University of Malaya

Thank you for the opportunity to review this article. It has been pleasure to read through the content.

The followings are some comments/suggestions to the authors:

**Title – describes well the scope and outcomes of the study.**

**Introduction – Clearly explained and linked to the study purposed.**

Methods:

**1 Study settings – were general, no details on where the study actually took place and how many samples obtained from each state and each time.**

**2 The sample size needs justification whether it represents the study population and power of the study.**

**3 Sampling methods used – needs more detail explanation as well. What was the response rate?**

*Response. Described on P7; L13-21*

### **Response to comments 1-3**

*The workshops were conducted by regional training providers in each state. We have amended the text to reflect this P6; L20-22.*

*We have included a description of the sampling method with response rates in the text. P7; 13-20*

#### **4. Study instrument – too brief, although it has been reported in another paper, it needs some explanation in context of this paper for readers' understanding.**

*Response: We have included a more detailed description of the GPACS-D p8 L9-10.*

#### **5. Data collection – who collected the data; any possibility for response bias?**

*Response: A research assistant, not associated with the workshops distributed the surveys P8; L2. Response bias might be present - this is discussed in the limitations section P21;L5-22.*

#### **6. Statistical test – no explanation on why non-parametric tests were used. In non-parametric tests, it would be more appropriate to use median rather than mean to describe the values of the variables.**

*Response. The median, while appropriate, did not clearly reflect the differences between the GP and Supervisor group. As a result of reviewer 3's comments we reviewed our analyses and decided to use t tests. Employing T tests did not affect the results. T Tests are also robust to violations of assumptions regarding normal distributions, but bootstrapped confidence intervals have also been provided. Levene's tests of homogeneity were also conducted to check other t-test assumptions. We have made changes to the tables and script in the results section to reflect this p9; L18-20. P10; L1, L3-10. P11; L8-12. P12; L4-6, L12-13, L18-20. P13; L3. P14; L1-5, L9-10, L12-13. P15; L11, L18-20, P16; L1-9, L15-16. P17; L6-7, L10-11. We have also amended the methods section to reflect changes in analytic method P8; L19-23. P9; L1-3.*

#### **7. The effect size values (Cohen's d) in Table 2, noted most of them are larger than 1.0, would they be any specific interpretation for it.**

*Response. Results larger than 1 indicates a strong effect. We have indicated this in the text. P9; L6*

*Results:*

- **Clearly interpreted and presented.**

#### **8. Some thoughts – why the authors didn't explore the potential confounders that may have influenced the outcomes of the study. It is essential to identify the possible confounding factors and control them in the analysis – to show the improvements in the outcomes are actual effects of the intervention.**

*Response. We did undertake an analysis of each of the descriptive variables against each of the subscales, and while prior training did impact on baseline scores, this was not evident post workshop. Additionally, the 'n' for those who had undertaken training was small and given this we could not be confident in the results. Similar results also emerged for having a relative with dementia, but this had no impact on pre or post workshop scores.*

*Discussion:*

#### **9. Would there be any possibilities for response bias at post intervention phase as the respondents would be aware that they need to present positive answers after the educational sessions? Having some discussion on this would benefit the future researchers.**

*Response. The potential for bias always exists. To minimise this our surveys were anonymous, and the RA made it clear that respondents could elect to refuse with no penalty or consequence. We provided no incentive for respondent bias. We have included these points in the limitations section. P21;L5-22.*

#### **10 To include the limitations of the study: for e.g. design – not having a control group, self-reported answers, actual changes in their performances and knowledge were not assessed, etc.**

*Response. A full description of the limitation of the research is discussed on P21;L5-22.*

*References*

- Mostly were latest, adequate and appropriate.

*Reviewer: 2*

*Reviewer Name*

Mieke van Driel

Institution and Country

University of Queensland, Australia

This paper addresses an important issue in general practice and assessing the impact of education is always worthwhile. The paper is well written; however, some important information is missing. The following points may benefit from more clarity.

**1. there is information missing on how the training was delivered, apparently different for GPR and SV? was it delivered face to face to all? what was the fidelity, did everyone participate?**

*Response. On P6; L22 we note that the workshops are face to face. We have amended the text to include face to face presentation for the Supervisor workshops as well as the Registrars. P7; L1-2.*

*In so far as participation is concerned, the workshop is offered as part of a suite of training modules that registrars can undertake. While Registrars are compelled to undertake further training, the modules they select are voluntary. Participation rates for the survey are contained on P9; L13-21.*

**2. what was the sample frame of this study? how many GPR and SV were eligible, how many responded to the survey? how was the survey administered?**

*Response. We have included a more detailed description of -the sampling method with response rates in the text p9; L13-21*

**3. the authors claim in their summarising statements that they included a representative sample of Australian GPs? what is that statement based on? these are 2 different groups with very different demographic characteristics. It is likely that those involved in GP training as supervisors are more engaged and different from their non-SV peers?**

*Response. Registrar data for the Registrars' Clinical Encounters in Training (ReCEnT) projects suggests that in so far as age and gender are concerned proportions are similar. This result is reflected in Australian General Practice Training (AGPT) and General Practice Supervisors Australia (GPSA) data as well. We have included text in the results section to reflect this. P10; L 1-8*

**4. the title is a bit misleading as this study is about registrars and established GPs, not only about GPs.**

*Response. We consider that the two groups are acceptably covered by the general term 'GPs'. We also note that Reviewer 1 indicated that they thought the title clearly reflected the content and scope of the article.*

**5. regarding the analysis, I am not a statistician, but it seems to me that the measured data were not 'independent' as they were in fact paired (each participant filled out the pre and post survey); how is this accounted for? How to interpret the Cohen's d?**

*We conducted two types of comparisons in our analyses. One was a comparison of the two independent workshop groups: Registrars and the Supervisors. The other was to compare results from each group between pre and post-test periods.*

**6. the results section is very long and repeats data in tables; for ease of reading this section could be condensed a bit.**

*Response. Efforts have been made to avoid repetition of data in tables. We consider the remaining explanations necessary and note that other reviewers did not highlight this as a concern.*

**7. the authors do not discuss any limitations of their study**

*Response. We agree and thank the reviewer for their advice. Limitations have been added P21;L5-22.*

**8. what are the implications of the study findings? where to from here? the 'so what' factor hasn't really been adequately addressed. What do we know about the link between knowledge and attitudes and actual behaviour in clinical practice? there is a wealth of literature on this topic and the authors might want to provide us with some insights in that regard to give the discussion more depth and meaning. what is the take-home message for those involved in GP training or CPD?**

*Response. We have now addressed the link between knowledge attitudes and confidence as well as the extent to which attitudes and confidence are sustained post workshop. Additionally, we have*

discussed the importance of attitudes in affecting changes in behaviour and the utility of the survey in providing a basis from which to develop targeted educational interventions. P20; L3-21.

Reviewer: 3

Reviewer Name

Ivan Buljan

Institution and Country

University of Split School of Medicine

**1. In the statistical analysis part, you report on using non-parametric procedures (Mann Whitney U test and Wilcoxon sign rank test), but also present your data as means with standard deviations, which is not appropriate. I would advise to present your data as medians with corresponding 95% confidence intervals around medians. I understand that, when presented as medians, the differences will be difficult to present because the range used is very narrow, but it is the recommended way of data presentation.**

**2. Also, I would strongly advise to present median differences between groups (pre-and post-assessment), to determine the actual size of difference, instead of relative effect size.**

**3. Cohen's d is not an appropriate measure of effect size in non-parametric testing, but for your specific situation you could use r effect size. It is easily calculated using the JASP statistical software, which is free to use and similar to SPSS, which you are currently using in your research.**

**4. However, if the distribution is not significantly different from normal distribution, it would be possible to use t-tests in testing differences. However, in that case you should present data with means and corresponding 95% confidence intervals, mean difference and Cohen's d effect size.**

*Response. Thank you for your comments and suggestions. We reviewed our statistical testing as a result of your comments, especially the issue of median scores and how to determine differences between scores for each group. Based on this we decided use t- tests for independent and paired samples. While our data was skewed this was within acceptable limits and given sample sizes for each group, and the robustness of t-tests to violations of assumptions of normality, we decided that the use of t-tests was an appropriate course of action. We also have also supplied bootstrapped confidence intervals.*

*We have adjusted the text in the analysis section to reflect these changes P8; L19-23, P9; L1-3. We have also made changes to the tables and script in the results section to reflect this p9; L18-20. P10; L1, L3-10. P11; L8-12. P12; L4-6, L12-13, L18-20. P13; L3. P14; L1-5, L9-10, L12-13. P15; L11, L18-20, P16; L1-9, L15-16. P17; L6-7, L10-11. We have also amended the methods section to reflect these changes P8; L 19-23, P9; L1-3.*

**5. Please format tables in your Results section (put your attention to Table 1 especially).**

*Response; Amended. P10; L13-14*

**6. Please use appropriate symbol for chi squared, instead of using x.**

*Response. Amended. P9; L20, P10; L1*

**Once again, thank you for the opportunity to review this paper.**

## VERSION 2 – REVIEW

REVIEWER	Mieke van Driel University of Queensland, Australia
REVIEW RETURNED	15-Oct-2019

<b>GENERAL COMMENTS</b>	<p>Thank you for your revision which has improved the ability to understand the study and implications. There are however a few things that are quite essential and still unclear.</p> <ul style="list-style-type: none"> <li>- On page 7 you mention a purposive sampling of registrars (without any information on how this was done and what criteria were used); then on page 8 you mention that all participants were invited?; this is confusing, please explain in more detail. I am also unsure about the 'modules mentioned on page 21, was there more to the intervention than the workshop?</li> <li>- although the intervention is described on pages 6-7, I still don't know what I should do if I would like to replicate your results; your intervention seems to be based on the one described in the Tierney study, but that was 3 hours f2f and yours is 1.5. So what did you do in that workshop with a specific focus on changing attitudes?</li> <li>- is this study a subanalysis (with different outcomes) of the Tierney publication in BMJ Open? if so, please say so to be transparent</li> <li>-I'm sorry, but I still find the results section very long and uninviting.</li> <li>-please check your referencing software as in the reference to the ReCEnT study (ref 42) the 2 final authors have been left off the list.</li> </ul>
-------------------------	--

<b>REVIEWER</b>	Ivan Buljan University of Split School of Medicine Croatia
<b>REVIEW RETURNED</b>	22-Oct-2019

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to re-review this paper. The authors have taken my suggestions into account and revised the manuscript accordingly. I have only few suggestions now, and I do not need to see the manuscript again.</p> <p>P9L54 There is a typo after chi squared (number 2 is redundant). Table 1 needs better explanation on what it presents (means, frequencies).</p> <p>Perhaps the most important comment I have is that P values cannot be presented as 0.000, but the lowest value can be <math>P &lt; 0.001</math>, so that needs to be revised in the article.</p> <p>Once again, thank you for this opportunity.</p>
-------------------------	---

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name

Mieke van Driel

Institution and Country

University of Queensland, Australia

Please state any competing interests or state 'None declared':  
none

Please leave your comments for the authors below

Thank you for your revision which has improved the ability to understand the study and implications. There are however a few things that are quite essential and still unclear.

Comment 1: On page 7 you mention a purposive sampling of registrars (without any information on how this was done and what criteria were used); then on page 8 you mention that all participants were invited; this is confusing, please explain in more detail.

Response: We have included the following text to clarify:

P7; L20-24. Lists of GPs attending the GP Registrar and Supervisor workshops were provided by each regional training organisation/provider and used as the sample frame for each region. The list comprised the GP's name and a unique id number to ensure that each pre and post survey matched with the individual.

P8; L21-23. P9; L1-3. A research assistant not associated with delivery of the workshop administered the survey. Pre-test surveys were provided to each participant as they arrived along with an information sheet about the research. Participants were informed that the survey was entirely voluntary and that completion of the survey implied consent. Participants completed the survey [38] immediately before (T1) and immediately after (T2) the workshop, with each survey matched via the unique ID for each attendee.

Comment 2: I am also unsure about the 'modules' mentioned on page 21, was there more to the intervention than the workshop?

Response: The term 'module' was used to describe other training courses available. In responding to your comment, we discovered that the statement itself was in fact ambiguous, so we have amended the sentence to rectify this and have removed the term 'module'.

We have altered the text to read:

P20; L20 . For participating registrars, the workshop was a part of their compulsory training program.

Comment 3: Although the intervention is described on pages 6-7, I still don't know what I should do if I would like to replicate your results; your intervention seems to be based on the one described in the Tierney study, but that was 3 hours f2f and yours is 1.5.

Response: Yes, the intervention is the same intervention as described in Tierney et al. We have made reference to this on P6, L8-12 and P7, L4. Further, we had described the workshop as comprising 2x1.5 hour sessions, which we accept may lead to confusion. In reality the workshop comprises a morning and afternoon session each 1.5 hours in duration and totalling 3 hours. In order to mitigate this confusion, we have changed the original text to read "a three hour face to face session" P6; L22.

We have also altered text to reflect that this intervention is the same as the one conducted by Tierney et al. The following text was included:

P8, L7-13. The workshop was evaluated using a pre-test post-test framework which employed two measures. Changes in knowledge of dementia were assessed using the Dementia Knowledge Assessment Survey (DKAS) (see Tierney et al. 2019). This paper reports the second arm of the evaluation which utilised the GPACS-D survey [38] to evaluate the impact of the workshops on confidence and attitudes.



Further clarity about the content and approach of the intervention is provided p7; L10-17 (see response to Comment 4 below).

Comment 4: So what did you do in that workshop with a specific focus on changing attitudes?

Response: Our main interest lay in whether the workshop improved knowledge and had a positive effect on attitudes and confidence. While no specific strategies were aimed at changing attitudes or increasing perceptions of self-efficacy, we did approach a number of attitude related issues in the content of the Workshop. These included a focus on a therapeutic rather than purely biomedical approach to treatment and care, emphasising the benefits of early diagnosis and the cost of not doing so. The workshop also aimed to encourage GPs to play a more active role in the treatment and management of dementia (essentially attempting to increase engagement).

We have included the following text to address this:

P7, L10-17 In an attempt to address GPs reluctance to diagnose dementia (Hansen et al. 2008), there is a strong focus on highlighting lived experience in order to situate people with dementia and their carers as central to the process, and to consider diagnosis and management through a biopsychosocial lens (Engel, D 1980; JS Jackson, TC Antonucci, E Brown, 2003). The intent is to facilitate GPs to engage with the process of diagnosis and associated management in a timely and supportive fashion.

Comment 5: is this study a sub analysis (with different outcomes) of the Tierney publication in BMJ Open? if so, please say so to be transparent.

Response: Yes, this is a sub analysis of the same workshop described in Tierney. We have added the following text to address this:

P8; L7-13. The workshop was evaluated using a pre-test post-test framework which employed two measures. Changes in knowledge of dementia were assessed using the Dementia Knowledge Assessment Survey (DKAS) (see Tierney et al. 2019). This paper reports the second arm of the evaluation which utilised the GPACS-D survey [38] to evaluate the impact of the workshops on confidence and attitudes.

Comment 6: I'm sorry, but I still find the results section very long and uninviting.

Response: We have re-examined the results section and removed text where appropriate, with associated editing for flow of ideas. Pp 10-17.

Comment 7: please check your referencing software as in the reference to the ReCEnT study (ref 42) the 2 final authors have been left off the list.

Response: We added the last 2 authors to the reference for the ReCEnT study mentioned (now ref 45).

Reviewer: 3

Reviewer Name

Ivan Buljan

Institution and Country

University of Split School of Medicine  
Croatia

Please state any competing interests or state 'None declared':  
None declared

Please leave your comments for the authors below

Thank you for the opportunity to re-review this paper. The authors have taken my suggestions into account and revised the manuscript accordingly. I have only few suggestions now, and I do not need to see the manuscript again.

Comment 1: P9L54 There is a typo after chi squared (number 2 is redundant).

Response: Rectified P10, L8.

Comment 2: Table 1 needs better explanation on what it presents (means, frequencies).

Response: We have added text to the table heading to better explain the table. P11, L9-10.

Comment 3: Perhaps the most important comment I have is that P values cannot be presented as 0.000, but the lowest value can be  $P < 0.001$ , so that needs to be revised in the article.

Response: All instances of this in Tables and text have been corrected. Pp 10-17.