PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effects of turnover and stability of health staff on quality of care in remote communities of the Northern Territory, Australia: a retrospective cohort study
AUTHORS	Jones, Michael; Zhao, Yuejen; Guthridge, Steven; Russell, Deborah; Ramjan, Mark; Humphreys, John; Wakerman, John

VERSION 1 – REVIEW

REVIEWER	Onnis, Leighann James Cook University Division of Tropical Environments and Societies, College of Business, Law & Governance
REVIEW RETURNED	15-Aug-2021

GENERAL COMMENTS	Page 9 - It would be good to have used a distance calculator other than google (it lets the paper down a little) Page 14, line 51 - there is a typo, a closing bracket after the word inhibitor without an opening bracket. Page 34, line 10 - the last 3 cells are empty Table 7 - the IV Contrasts column seems out of place as it is not mentioned in the paper and/or the blank cells need explaining if it remains in the paper

REVIEWER	Bishop, Jaclyn The Peter Doherty Institute for Infection and Immunity, National Centre for Antimicrobial Stewardship
REVIEW RETURNED	25-Aug-2021

GENERAL COMMENTS	Thank you for your submission. It is pleasing to see research focused on remote healthcare workforce, and also the publication of findings that did not align with your original hypothesis. I offer the following suggestions: Abstract: Setting: given the international audience, please add Australia into the description. (L15) Outcome measures: consider whether the quality indicators should be listed as outcomes, rather than your independent variables. (L18-20) Conclusion: I feel that the statement that 'lower staff stability and greater use of short-term staff is associated with deficits in care for some clinics but not others' is too broad and may lead the reader to overestimate the significance. Could you be more specific (e.g. about the number of clinics where there was an association and

	which/number of indicators) so that it is clear that in the majority of cases there was no association. (L47-52) Main text: Discussion: I'd like to read your view on where to next based on the findings of this research. Given the lack of association and difficulty in measurement, should research resources be focused elsewhere? You've indicated the possibility of exploring individual clinic factors more closely- given this analysis was of data 2011-15, have you begun anything to explore these findings further in the subsequent years? In the declarations you also indicate that it wasn't possible to include consumers in this study- is future research going to more actively include consumers? Your original protocol published in 2016 indicated that this study was part of a mixed-methods design that included these components - is that still the plan? Limitations: can you comment on how well this study matched your original protocol published in 2016. Were there some outcomes that you didn't/couldn't collect? Conclusion: as per the comment provided on the abstract, please consider revising the first finding. Other comments: Table 7: what is the purpose of the empty column titled IV contrasts? If this column is to be completed, I would avoid use of 'IV contrast' as the heading as this has a different meaning in medicine (an agent used in medical imaging). Also consider if some of the data is best presented as supplementary materials e.g. Table 7

REVIEWER	Martiniuk, Alexandra University of Sydney Sydney Medical School, School of Public Health
REVIEW RETURNED	30-Aug-2021

GENERAL COMMENTS	Professor Alexandra Martiniuk With Joseph Freeman (MD candidate USYD 2021) Effects of turnover and stability of health staff on quality of care in remote communities of the Northern Territory, Australia: a retrospective cohort study Thank you for asking me to review this work. Authors: are any of the authors Aboriginal or Torres Strait Islander? This would be a strength if yes Review Checklist 1. Is the research question or study objective clearly defined? The research question is clear.

	Omitted in the abstract but found at the bottom of the Introduction is a secondary objective: 'We also sought to identify factors which may counter hypothesised reduced quality of care caused by lower stability, higher turnover or higher use of short-term staff.' This secondary objective also wasn't addressed in the conclusion. 2. Is the abstract accurate, balanced and complete? Abstract: the results section would benefit from greater detail, it read as vague Strengths/limitations: similar to above, eg specifying key confounders would make this section more useful Conclusions in abstract are not entirely consistent with conclusions in body. Also, conclusion in abstract does not address the stated objectives. Abstract states that greater turnover is associated with deficits in quality of care for some clinics but not others. Conclusion should instead address the negative finding of the study. 3. Is the study design appropriate to answer the research question? This is a retrospective study using routinely collected administrative data. This is one appropriate way to answer the study question, with advantages and disadvantages. However, it is a major disadvantage to be missing data on Aboriginal health workers and AMS vs other clinic type. The Authors do discuss the importance of Aboriginal health workers and staff and present this as one of the potential reasons for differences observed between clinics. Given our existing knowledge regarding the importance of Aboriginal health workers and staff in quality of primary care for Aboriginal people, it is a shame that there are no data used regarding Aboriginal health worker presence or not. A discussion of why these data are not included would be useful. Could these data be collected and analysed along with the routine data already in the paper? 4. Are the methods described sufficiently to allow the study to be repeated? Yes, detail is adequate. Published protocol also available for reference 5. Are research ethics (e.g. participant consent, ethics approval) addressed appropriately? Yes 6. Are the outcomes clearly defined?

	Sample could be more clearly described – 48 clinics. Unclear number of patients served by each clinic. Unclear total staffing number and type by clinic. Quality of care indictors – are these the best markers of quality of care? Any indicators for patient experience with respect to quality of care from their perspective? It would be useful to understand the medicare billing schedule for these various items, funding/amount by indicator – to understand if these truly are good indicators of overall quality of care. Overall outcomes are well defined, the tables are clearly presented. 7. If statistics are used are they appropriate and described fully? Tables could be improved for clarity – titles and columns/footnotes could be improved for reader clarity (ie Tables should "stand alone" but currently do not) Should correction for multiple testing have been used in this paper? The clustering of the clinics….I understand latent class analyses were used to cluster clinics -with the purpose to better understand clinic profiles – clusters were set according to staff stability. It was unclear to me: initially 2 clusters (one large cluster, and a very small cluster of 4 clinics)? Then the 4 clinics dropped, and the large cluster further underwent latent class analyses to split further into 3 clusters? Overall the rationale, method and results of the clustering of clinics could be made clearer. 8. Are the references up-to-date and appropriate? Yes. May have considered using more Aboriginal-led work in this topic? Eg Access to primary health care services for Indigenous peoples: A framework synthesis Carol Davy1* , Stephen Harfield1, Alexa McArthur2, Zachary Munn2 and Alex Brown Farnbach S, Eades AM, Fernando JK, Gwynn JD, Glozier N, Hackett ML. The quality of Australian Indigenous primary health care research focusing on social and emotional wellbeing: a systematic review. Public Health Res Pract. 2017;27(4):e27341700. Embedding cultural safety in Australia's main health care standards Martin Laverty, Dennis R McDermott and Tom Calma

	Med J Aust 2017; 207 (1): 15-16. || doi: 10.5694/mja17.00328 Published online: 3 July 2017 Characteristics of Indigenous primary health care service delivery models: a systematic scoping review • Stephen G. Harfield, • Carol Davy, • Alexa McArthur, • Zachary Munn, • Alex Brown & • Ngiare Brown Globalization and Health volume 14, Article number: 12 (2018) Cite this article 9. Do the results address the research question or objective? Yes Aim of paper is to determine effects of stability of staff on quality of care in remote communities – An improvement might be regarding the term “quality of care” ---should be made more specific/narrow in title/aim? I am not sure that remote living Aboriginal communities would agree that those routinely collected data summarise well their quality of care. 10. Are they presented clearly? Overall writing: reducing use of acronyms will increase reading flow and clarity eg IVs, DVs….especially given other uses of these acronyms which are more common. Table 7 has a sub-title that appears to be in draft, also the IV contrast columns could be done differently – as presented this table is challenging to read/understand Unclear what the box plot is about on page 42 11. Are the discussion and conclusions justified by the results As previous, not fully. The authors conclude with ‘two very clear findings’ the first of which is that increased staff turnover correlates with decreased quality of care in some clinics but not others. However, it may be interpreted that no clear signal was found in the data for the primary objective. Thus the opening sentence of the Discussion “Overall, minimal evidence of the hypothesised negative effects of increased turnover…” is not fully accurate given the data - 12. Are the study limitations discussed adequately? Yes, but not translated into conclusions.

	An aboriginal perspective regarding the measures used appears to be lacking. A limitation is a lack of data regarding Aboriginal health workers? Or perhaps I am missing this somewhere? 13. Is the supplementary reporting complete (e.g. trial registration; funding details; CONSORT, STROBE or PRISMA checklist)? STROBE item 15 (b) asks for category boundaries – should they be reported in text for the 3 Clusters? The description of clusters needs to be clarified as mentioned previously. 14. To the best of your knowledge is the paper free from concerns over publication ethics (e.g. plagiarism, redundant publication, undeclared conflicts of interest)? Yes 15. Is the standard of written English acceptable for publication? Yes, generally -the writing style is clear and adequate – except as specified above.

VERSION 1 – AUTHOR RESPONSE

Response to editors' and reviewers' comments: bmjopen-2021-055635

Comment	Response
Reviewer 1: Dr. Leighann Onnis, JCU Division of Tropical Environments and Societies	
Page 9 - It would be good to have used a distance calculator other than google (it lets the paper down a little)	We respectfully disagree. Given the very large distances involved – clinics are located an average of 304 km from Alice Springs or Darwin (whichever is nearest) – and the ability on google maps to zoom in to accurately pinpoint a location, we feel that google maps is more than adequate for calculating distances for this study.
Page 14, line 51 - there is a typo, a closing bracket after the word inhibitor without an opening bracket.	Thanks for picking this typo up. We found the same typo in a number of places throughout the paper. These have all been corrected.

Page 34, line 10 - the last 3 cells are empty	We think that this comment refers to Page 35, line 10, Table 5. These blank cells for the number of clinic-months are not applicable for line 10, Number of clinics. We have made this clearer for the reader by inserting n/a (not applicable) in these cells. We have also revised Table 5 to make it clear that the number of clinic-months was also not applicable for the traits section of the table.
Table 7 - the IV Contrasts column seems out of place as it is not mentioned in the paper and/or the blank cells need explaining if it remains in the paper	Thankyou for bringing our attention to the difficulty for the reader in interpreting the information presented in the columns previously labelled as IV Contrasts. We have removed these columns and added the asterisk indicating statistical significance to the specific statistic to which it refers. Please also see our response below to Reviewer 2's query about these columns in Table 7.
Reviewer 2: Dr. Jaclyn Bishop, The Peter Doherty Institute for Infection and Immunity	
ABSTRACT Setting: given the international audience, please add Australia into the description. (L15)	This has been updated. Thankyou for identifying this issue and drawing our attention to it.
Outcome measures: consider whether the quality indicators should be listed as outcomes, rather than your independent variables. (L18-20)	Thankyou for the suggestion. While there are probably too many quality indicators to list individually, given the constraints of abstract length, we have modified the text to give a broad indication of the different types of quality indicators (outcome measures) used. This now reads: Associations between independent variables (Resident Remote Area Nurse (RAN) and Aboriginal Health Practitioner (AHP) turnover rates, stability rates and the proportional use of agency nurses) and indicators of health service quality in child and maternal health, chronic disease management and preventive health activity were tested using linear regression, adjusting for community and clinic size. Latent class modelling was used to investigate between-clinic heterogeneity.

Conclusion: I feel that the statement that 'lower staff stability and greater use of short-term staff is associated with deficits in care for some clinics but not others' is too broad and may lead the reader to overestimate the significance. Could you be more specific (e.g. about the number of clinics where there was an association and which/number of indicators) so that it is clear that in the majority of cases there was no association. (L47-52)	We appreciate the desire for a more tangible sense of how much variation in what direction than is provided by the standard deviation (SD) columns of Table 4. The practical challenge is that there are just too many results to do this for every one. However we had included on set of boxplots to illustrate for a simple example of just how much variation and in what direction, for agency nurse proportion and antenatal care (Figure 1). At the reviewer's suggestion we have now added the distribution of association sizes classified as negative, effectively zero or positive on a standardized basis. We hope this adds a level of concrete interpretation.
Main text: Discussion: I'd like to read your view on where to next based on the findings of this research. Given the lack of association and difficulty in measurement, should research resources be focused elsewhere? You've indicated the possibility of exploring individual clinic factors more closely- given this analysis was of data 2011-15, have you begun anything to explore these findings further in the subsequent years?	Thankyou for raising this question. Despite the lack of association between workforce stability and quality of care, and difficulty in measurement, we do not think that research resources should be focused elsewhere. It remains critically important to better understand the substantial between-clinic variation in quality of care that this study revealed, as this may be highly relevant to attaining more equitable health outcomes for Aboriginal and Torres Strait Islander populations living in remote Australia (and for Indigenous populations living in remote communities elsewhere in the world). We have added a paragraph to the discussion which addresses your query. It reads: Given the high health care needs of Aboriginal and Torres Strait Islander populations living in remote communities, and the importance of ensuring equitable access to high quality primary health care, future research is warranted to explore whether and how the range of factors postulated as possible explanations are indeed associated with the substantial between-clinic variation in quality of primary care in remote clinics. The authors are currently undertaking some of this work by examining whether similar patterns exist in Aboriginal Community Controlled Clinics in the same jurisdiction and will be updating analyses with NT Department of Health data to try and gain a better understanding of the extent to which factors such as cultural safety, employment of Aboriginal staff and acceptability of the health service to community members are associated with variability in quality of care.

In the declarations you also indicate that it wasn't possible to include consumers in this study- is future research going to more actively include consumers? Your original protocol published in 2016 indicated that this study was part of a mixed-methods design that included these components - is that still the plan?	This study was part of a broader project utilising a mixed-methods design that included consumers, as per our published protocol. However, the particular study reported in this paper did not include analysis of qualitative data collected from consumer interviews and focus groups. Nevertheless, the broader project's qualitative data analyses informed the discussion section of this paper. We have amended the sub-section Methods/Patient and public involvement as follows: It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of the research reported here. Nevertheless, this study formed part of a broader project which used mixed-methods, including interviews and focus group discussions with clinic users. The results of the broader project were disseminated to participants and the analysis of the broader project informed the interpretation of this study's findings and discussion.
Limitations: can you comment on how well this study matched your original protocol published in 2016. Were there some outcomes that you didn't/couldn't collect?	This study closely reflected the outcomes stated in our original protocol published in 2016. No data were available on the proportion of patients with cardiovascular disease on acetylsalicylic acid, so this indicator was not used. We also did not use the quality indicator proportion of known hypertensives with controlled blood pressure. As indicated in the main text our aim was to examine a sufficient range of indicators of quality of care to ensure that a reasonable spectrum of PHC activities was covered by the indicators, rather than to be exhaustive. We did have access to BP measures but concluded that other measures had greater utility. We have added an explanation in the limitations section: Additionally, no data were available on the proportion of patients with cardiovascular disease on acetylsalicylic acid, so even though measurement of this outcome was described in the study protocol, this indicator was not used.
Conclusion: as per the comment provided on the abstract, please consider revising the first finding	The conclusion has been revised accordingly.

Other comments: Table 7: what is the purpose of the empty column titled IV contrasts? If this column is to be completed, I would avoid use of 'IV contrast' as the heading as this has a different meaning in medicine (an agent used in medical imaging).	Please also see our response above to Reviewer 1's query about the columns previously labelled IV contrasts (now removed) in Table 7. The purpose of these 3 columns was to show statistical significance of associations between cluster number and each of the respective quality indicators. Only 2 of the quality indicators were statistically significant at $p<0.05$, and these cells had asterisk in them: cluster 1 was significantly associated with having any antenatal visit; cluster 1 was also significantly associated with eligible women having had a pap smear in the past 2 years. Blank cells indicated that there was no statistically significant association between cluster number and the respective quality indicator. We have now added the asterisk to the two cells where the associations for these clusters and stability metric are reported.
Also consider if some of the data is best presented as supplementary materials e.g. Table 7	The information presented in Table 7 relates to how clustering of clinics (clustered by workforce metrics) is not associated with quality of care. These data, which are not as hypothesised, represent an important albeit negative finding of the study and our preference is to retain the table in the main text.
Reviewer 3: Prof. Alexandra Martiniuk, University of Sydney Sydney Medical School With Joseph Freeman (MD candidate USYD 2021)	
Authors: are any of the authors Aboriginal or Torres Strait Islander? This would be a strength if yes	While none of the authors of this particular paper are Aboriginal or Torres Strait Islander, the broader project, of which this study is a part, included several Aboriginal and Torres Strait Islander researchers. Input and feedback was sought from all team members as the manuscript was developed and authorship offered to all team members who met the ICMJE authorship criteria.
The research question is clear. Omitted in the abstract but found at the bottom of the Introduction is a secondary objective: 'We also sought to identify factors which may counter hypothesised reduced quality of care caused by lower stability, higher turnover or higher use of short-term staff.' This secondary objective also wasn't addressed in the conclusion.	The secondary objective is now included in the abstract: A secondary objective was to identify factors which may counter hypothesised reduced quality of care associated with lower stability, higher turnover or higher use of short-term staff.

Abstract: the results section would benefit from greater detail, it read as vague	Please see comment to Editor, above, regarding the difficulty in adequately summarising such a large number of results numerically. Nevertheless, we have attempted to provide greater detail in the manuscript, where possible, so that the results are less vague. The amended abstract results section now reads: The proportion of resident Aboriginal clients receiving high quality care as measured by various quality indicators varied considerably across indicators and clinics. Higher quality care was more likely to be received for management of chronic diseases such as diabetes and least likely to be received for general/preventive adult health checks. Many indicators had target goals of 0.80 which were mostly not achieved. The evidence for associations between decreased stability measures or increased use of agency nurses and reduced achievement of quality indicators was not supported as hypothesised. For the majority of associations, the overall effect sizes were small (close to zero) and failed to reach statistical significance. Where statistically significant associations were found, they were generally in the hypothesised direction.
Strengths/limitations: similar to above, eg specifying key confounders would make this section more useful	Key confounders, listed in the methods section (other independent variables) are now explicitly identified in the main text as potential confounders and also in this sub-section: Analyses adjusted for key potential confounders, including remote community population size, average number of employees, Euclidean distance in kilometres to the major centres of Darwin or Alice Springs (whichever was closest) and Euclidean distance in kilometres to the nearest of the five NT hospitals;

Conclusions in abstract are not entirely consistent with conclusions in body. Also, conclusion in abstract does not address the stated objectives. Abstract states that greater turnover is associated with deficits in quality of care for some clinics but not others. Conclusion should instead address the negative finding of the study.	We have amended the abstract conclusion so that it is clearer that the negative findings of the study are clearly addressed and so that the stated objectives are also addressed: Overall, minimal evidence of the hypothesised negative effects of increased turnover, decreased stability and increased reliance on temporary staff on quality of care was found. Substantial variations in clinic-specific estimates of association were evident, suggesting that clinic-specific factors may counter any potential negative effects of decreased staff employment stability. Investigation of clinic-specific factors using latent class analysis failed to yield clinic characteristics that adequately explain between-clinic variation in associations. Understanding the reasons for this variation would significantly aid the provision of clinical care in remote Australia.

This is a retrospective study using routinely collected administrative data. This is one appropriate way to answer the study question, with advantages and disadvantages. However, it is a major disadvantage to be missing data on Aboriginal health workers and AMS vs other clinic type. The Authors do discuss the importance of Aboriginal health workers and staff and present this as one of the potential reasons for differences observed between clinics. Given our existing knowledge regarding the importance of Aboriginal health workers and staff in quality of primary care for Aboriginal people, it is a shame that there are no data used regarding Aboriginal health worker presence or not. A discussion of why these data are not included would be useful. Could these data be collected and analysed along with the routine data already in the paper?	All clinics in this study were NT Department of Health remote clinics. No clinics were AMSs/Aboriginal Community Controlled Health Services. There are no missing data about clinic type. The study design (retrospective cohort study) is also not the reason that specific data on Aboriginal health workers were not included. These data were available and could have been used as independent variables. However, it was necessary to select only a small number of key workforce metrics, so it was inevitable that some indicators, even those of hgh interest, were omitted. The metrics chosen were guided by previous research on the use of workforce indicators in rural and remote Australia.[1] We acknowledge that the body of research regarding appropriate health workforce indicators in rural and remote Australia is currently weak. Several of this study's authors are co-investigators on an Aboriginal-led study which aims to prioritise a range of non-clinical indicators, including workforce indicators, for use in remote Aboriginal health services. The pared-back list of workforce indicators currently sits at 44, but hopefully by the end of 2021 we will have a better idea of which workforce indicators matter most in this context. At the time of these analyses, however, we used the best information available to guide our selection of indicators. We consider it likely that the number of Aboriginal staff (not just of Aboriginal health workers) may be identified as a key workforce indicator, however this work is ongoing and this is by no means a certainty. We have added a further limitation to the manuscript which reads: A further limitation related to our choice of workforce indicators. Given the dearth of literature describing health workforce metrics specific to the Australian remote Aboriginal and Torres Strait Islander context, our choice of health workforce indicators was guided by literature taken from the rural Australian context.[1] The metrics we chose were unable to capture all important facets of the remote health workforce, including, for example, employment of local Aboriginal staff.
Sample could be more clearly described – 48 clinics. Unclear number of patients served by each clinic. Unclear total staffing number and type by clinic.	We have now ennumerated the total number of patients served by the included clinics. (Abstract/Participants section; also Methods/Data sub-section)

Quality of care indictors – are these the best markers of quality of care? Any indicators for patient experience with respect to quality of care from their perspective? It would be useful to understand the medicare billing schedule for these various items, funding/amount by indicator – to understand if these truly are good indicators of overall quality of care.	The markers of quality of care used were the best available at the time. The development of the NT AHKPIs was guided by the Northern Territory (NT) Aboriginal Health Forum. More information on the NT AHKPI indicators, including background to their development, governance and use of NT AHKPIs etc. is available at: https://health.nt.gov.au/professionals/aboriginal-health-key-performance-indicator/nt-ahkpi-project-background Unfortunately, Medicare billing does not necessarily or accurately reflect clinical activity in remote government-run health clinics, which most frequently are nurse-led rather than doctor-led. We are not aware of any indicators of patient experience with respect to quality of care received by remote residents.
Overall outcomes are well defined, the tables are clearly presented.	Thankyou.
Tables could be improved for clarity – titles and columns/footnotes could be improved for reader clarity (ie Tables should “stand alone” but currently do not)	We have done as suggested and improved clarity of the tables, including by providing greater description of each table in its title and more detail in footnotes where needed.
Should correction for multiple testing have been used in this paper?	We appreciate the suggestion but feel that this paper, being the first of its kind in this environment, was exploratory in nature and therefore suggests evidence of effects rather than proving them. For this reason we have been cautious in our interpretation of the numerous statistical hypothesis tests reported and feel that correct for multiple testing is not necessary. Given the reviewer’s question however we had added an explicit comment on the exploratory nature of the study in the statistical methods section.

The clustering of the clinics….I understand latent class analyses were used to cluster clinics -with the purpose to better understand clinic profiles – clusters were set according to staff stability. It was unclear to me: initially 2 clusters (one large cluster, and a very small cluster of 4 clinics)? Then the 4 clinics dropped, and the large cluster further underwent latent class analyses to split further into 3 clusters? Overall the rationale, method and results of the clustering of clinics could be made clearer.	You have understood correctly. The initial 2 solution LCA revealed one large cluster and a very small cluster. The very small cluster was characterised by workforce metrics and quality indicators but ultimately dropped from further analysis as this cluster comprised four outlier clinics and outliers are known to unduly influence the cluster solution, as explained in the methods. The 4 clinics were therefore dropped, and the large cluster further underwent latent class analyses to split further into 3 clusters. This procedure resulted in a more robust solution of 3 clusters. We have reworded the methods/statistical approach sub-section which explains the rationale, method and results, so that it is clearer: The analysis was then repeated on the forty-four remaining clinics. The four clinics omitted from the repeated latent class analysis (due to them being outliers that could unduly influence further analysis) were statistically significantly associated with lower stability ($p=0.04$), higher turnover ($p=0.002$) and lower proportional use of agency nurses ($p=0.004$).

References: May have considered using more Aboriginal-led work in this topic? Access to primary health care services for Indigenous peoples: A framework synthesis Carol Davy1* , Stephen Harfield1, Alexa McArthur2, Zachary Munn2 and Alex Brown Farnbach S, Eades AM, Fernando JK, Gwynn JD, Glozier N, Hackett ML. The quality of Australian Indigenous primary health care research focusing on social and emotional wellbeing: a systematic review. Public Health Res Pract. 2017;27(4):e27341700. Embedding cultural safety in Australia's main health care standards Martin Laverty, Dennis R McDermott and Tom Calma Med J Aust 2017; 207 (1): 15-16. || doi: 10.5694/mja17.00328 Published online: 3 July 2017 Characteristics of Indigenous primary health care service delivery models: a systematic scoping review • Stephen G. Harfield, • Carol Davy, • Alexa McArthur, • Zachary Munn, • Alex Brown & • Ngiare Brown Globalization and Health volume 14, Article number: 12 (2018) Cite this article	Thankyou for your suggestions about Aboriginal-led work in this area. We have found two of these papers particularly useful and have cited them in the additional discussion inserted in the Discussion section that describes future areas of research.

9. Do the results address the research question or objective? Yes Aim of paper is to determine effects of stability of staff on quality of care in remote communities – An improvement might be regarding the term “quality of care” --- should be made more specific/narrow in title/aim? I am not sure that remote living Aboriginal communities would agree that those routinely collected data summarise well their quality of care.	We acknowledge that the quality indicators used in this study may not adequately reflect clinic users’ perceptions of quality of their care. We have amended the manuscript (Discussion/Limitations) so that it is clear to the reader that the study lacked indicators measuring certain important non-clinical aspects of quality of care. As previously indicated, the work to identify the most appropriate non-clinical indicators to use in remote Aboriginal health services in this jurisdiction is only now underway. Future studies, however, could use a broader range of non-clinical quality indicators.
Overall writing: reducing use of acronyms will increase reading flow and clarity eg IVs, DVs….especially given other uses of these acronyms which are more common.	Thankyou. We have removed these acronyms and spelled them out in full.
Table 7 has a sub-title that appears to be in draft, also the IV contrast columns could be done differently – as presented this table is challenging to read/understand Unclear what the box plot is about on page 42	We have removed the draft sub-title for Table 7. We have also removed the IV contrast columns in Table 7 as this was confusing for all reviewers. The boxplot on page 42 is Figure 1. The formatting when submitting files to meet BMJ requirements meant that its title was listed on page 41 (so it is not clear that the boxplot is Figure 1). The boxplot shows the distribution of slopes (β coefficients) for the 48 included clinics for the association between proportional use of agency nurses and each of three sub-indicators relating to provision of antenatal care. For example, if you look at one of the indicators eg. antenatal care at 13 weeks, you can see that there is a range of positive and negative associations depending on the clinic (3 clinics are outliers: 2 have stronger negative associations – high proportional use of agency nurses is associated with low proportion receiving antenatal care by 13 weeks; 1 clinic has strong positive association – high proportional use of agency nurses is associated with high proportion receiving antenatal care by 13 weeks).

11. Are the discussion and conclusions justified by the results As previous, not fully. The authors conclude with 'two very clear findings' the first of which is that increased staff turnover correlates with decreased quality of care in some clinics but not others. However, it may be interpreted that no clear signal was found in the data for the primary objective. Thus the opening sentence of the Discussion "Overall, minimal evidence of the hypothesised negative effects of increased turnover…" is not fully accurate given the data -	As above, the conclusion has been modified to better reflect the results.
12. Are the study limitations discussed adequately? Yes, but not translated into conclusions.	Now translated into conclusions, as above.
An aboriginal perspective regarding the measures used appears to be lacking.	The NT Aboriginal Health Key Performance Indicators (NT AHKPIs) that we used in this study as quality measures were developed by the Northern Territory (NT) Aboriginal Health Forum, which comprises of the Commonwealth Department of Health, Aboriginal Medical Services Alliance Northern Territory and Northern Territory Department of Health. The Aboriginal Medical Services Alliance Northern Territory is the peak body in Northern Territory for Aboriginal Community Controlled Health Organisations, and brings a strong Aboriginal perspective to the development and use of these clinical indicators. Only now is the Aboriginal Medical Services Alliance Northern Territory leading research on what are appropriate non-clinical indicators across a number of domains which reflect quality health care. The research is still incomplete and was not available at the time that this research was undertaken.
A limitation is a lack of data regarding Aboriginal health workers? Or perhaps I am missing this somewhere?	As discussed above, this is a limitation of the workforce indicators we selected. This limitation is now addressed in the Discussion/Limitations section.

13. Is the supplementary reporting complete (e.g. trial registration; funding details; CONSORT, STROBE or PRISMA checklist)? STROBE item 15 (b) asks for category boundaries – should they be reported in text for the 3 Clusters? The description of clusters needs to be clarified as mentioned previously.	The equivalent to category boundaries for latent class models is the highest predicted cluster membership probability, and this has now been added into the statistical methods section on page 11.

1. Russell, D.J., J.S. Humphreys, and J. Wakerman, *How best to measure health workforce turnover and retention: Five key metrics.* Aust Health Rev, 2012. **36**(3): p. 290-295.