PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effect of remote patient monitoring for patients with chronic kidney disease who perform dialysis at home: a systematic review
AUTHORS	Nygård, Henriette; Nguyen, Lien; Berg, Rigmor C

VERSION 1 – REVIEW

REVIEWER	Richard Bodington Northern General Hospital, Sheffield Kidney Institute
REVIEW RETURNED	05-Mar-2022

GENERAL COMMENTS	The question of whether RPM is effective in CKD is very important as interventions may be inserted into clinical practice without sufficient evidence to support their use. The systematic review is perfect way of assessing this. Overall I feel the paper is thorough in it's approach and balanced in it's reasoning. My major recommendation would be to edit the tables. Table 2 is difficult to understand. IRR is an unusual statistic to use and could cause some confusion. Also you switch between description and various RR in the results column. In Table 3 many of your rows don't fit with your column headings. Overall tables 2 and 3 are not intuitive and require several reads and use of the text. Including the name of the QoL tool in the table would be good also. Table 1 could be improved to make it clearer exactly what the intervention was. Following on from this, one of the limitations of the study is that you are comparing often very different interventions under the heading of RPM- I don't think this is explicitly acknowledged. Another limitation that is not stated is that the vast majority of your study population comes from two studies, and greater than half from a single study. You appear to discuss all of the studies equally. Some other small points: - In the abstract conclusion- you state RPM is associated with health benefits- is this too vague? -Page 6, line 5 - there is a half sentence that makes no sense -Your outcomes more or less align with the SONG-PD and SONG-HD outcomes but you don't explicitly mention this- maybe you should for increased generalisability. - P14 line 9- do you mean 'infections', rather than 'infections not requiring hospitalisation'? - P14 line 60- the data suggests or hints at these conclusions, it wouldn't say indicate. - P15 line 26- you state QoL improved- can you quantify this? - P15 line 35- you saw no difference in QoL- however here you seem to be suggesting that there is a link to less travel- what is this based on? Finally I am unable to comment on the appropriateness of the

	statistics you have used and this needs further review.

REVIEWER	Shrikant D. Pande Changi General Hospital
REVIEW RETURNED	18-Mar-2022

GENERAL COMMENTS	Knowing the limited studies done on this subject in the past, the authors have done a good job to review previous literature. I have no additional suggestions or comments due to the nature of the topic and past studies.

REVIEWER	William Levack University of Otago, Medicine
REVIEW RETURNED	31-Mar-2022

GENERAL COMMENTS	BMJ Open Peer Review 1 April 2022 Effect of remote patient monitoring for patients with chronic kidney disease who perform dialysis at home: a systematic review Thank you for the opportunity to review this paper. Overall, I think this review is well conducted and worthy of publication, although I think the clinical meaningfulness of some of the findings have been overstated (see below). As just an observation: the lack of high-quality clinical trials in this area of practice is perhaps an indication that the conduct of this review was a little premature – more meaningful conclusions may be reached once further studies have been conducted (as the authors point out). Nonetheless, I still think this review is a worth sharing with the international community as a first foray into exploring the evidence in this area of clinical practice. I drew on AMSTAR 2 when reviewing this systematic review. (See: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.) Overall, this review scores well on many of the criteria in AMSTAR-2, and certainly all the ones that I (personally) think really matter in terms of minimising major source of bias in a review. However, to improve the reporting of this review the authors may wish to consider adding the following: - A list of potentially relevant studies read in full but excluded as an Appendix, with reason for the exclusion of each. - Information about the sources of funding for each of the studies included in the review (as another possible source of bias). - A commentary on the possible impact of publication bias on the conclusions of this review. Because there were so few trials included in this review, it would not be possible to statistically evaluate publication bias. However, it would be possible to include a commentary on the potential for publication bias to have impacted on the findings from this review in the discussion section, and a recommendation to include an evaluation of publication bias in any further review on this topic that includes more studies. My main criticism of this paper relates to the interpretation and reporting of the review findings. In general, I believe that the international consensus is that it is misleading to report an outcome being in favour of an intervention when the measure of precision of the estimate of effect (i.e. the 95% CI) includes values that favour a control or comparison group. So, for example, I would not recommend reporting an outcome favouring

	RPM follow-up (e.g. less all-cause hospitalisations) when the precision of that estimate of effect (i.e. the 95% CI) is very wide and includes outcome that favour the control group. I also would recommend not reporting outcomes favouring PRM when the estimate of that effect is reported to be non-significant. This occurs throughout the results section and is used to draw conclusions regarding improved clinical outcomes for people who receive RPM follow-up compared to standard care in the discussion section. The same is true for the data on risk of infection and quality of life. The one outcome where I think there is clear evidence in favour of RPM follow-up is the risk of technical failure, where the meta-analysis shows a positive effect in favour of RPM follow-up, and where the 95% CI does not provide any data conflicting with this conclusion. My recommendation would be to revise the results, discussion, summary of findings table, and abstract in a way that does not overstate results where there is no clear effect in favour of RPM. Incidentally, the findings on hospitalisation only seem to be in favour of RPM for adjusted IRR, but I'm unclear how an IRR (a measure of rate of incidents of a particular event; e.g. hospitalisation) should be interpreted for count data, where each 'count' is not a separate event (e.g. hospital days). It is also not clear to me what the adjusted IRR data is adjusted for. It may be that some of my criticism above could be addressed by addressing this issue first. Regarding the summary of findings table: It is not clear to me why an assumed risk with a control intervention and with RPM cannot be reported for hospitalisation data, infection data, or QOL data, if data on these outcomes have been reported to draw conclusions about effect sizes in favour (or not in favour) of RPM. A few more recommendations: - I suggest that the authors add reference numbers to study IDs in the Tables. It's difficult to match the data in the Table to the information reported in the results section without these additional references. - The results section in the abstract should contain some statement regard the effect size of RPM follow-up if the conclusion in the abstract reads "there may be positive effects", regardless of whether there was low to very low evidence supporting this finding. Alternatively, the conclusion should be "there is insufficient evidence to draw a conclusion about the benefits of RPM follow-up".

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Dr. Richard Bodington (Northern General Hospital)	
Overall I feel the paper is thorough in it's approach and balanced in it's reasoning.	Thank you
My major recommendation would be to edit the tables. Table 2 is difficult to understand. IRR is an unusual statistic to use and could cause some confusion. Also you switch between description and various RR in the results column.	Thank you. We agree with regard to IRR. However, the authors of the three studies that presented IRR did not reply to our requests for more information/data and IRR is the best/most complete outcome data we have from these studies. To avoid confusion, we inserted a short text that explains IRR. Table 2 has the best/most complete data for each study, which regretfully varies across studies.

	The presentation reflects this variation and is indeed not ideal. Without the possibility to covert the varied results to one common scale (e.g. RR), improving the content of the table was difficult. We are happy to reconfigure tables 1-2 if that would be of interest.
In Table 3 many of your rows don't fit with your column headings.	Thank you. Recall that we used GRADE, which is a framework for developing and presenting summaries of evidence. It has a standard table format (with six columns) which we follow (see GRADE handbook). We've checked and cannot see that we have make any mistakes in the table set-up or content. Also, most of the outcomes could not be metaanalysed and we therefore followed the SWiM reporting guideline (we have now added the reference to this guideline).
Overall tables 2 and 3 are not intuitive and require several reads and use of the text.	Thank you. We would be happy to change table 2, but we are uncertain how we could make it more intuitive. We report the results for each outcome from each study with the most complete data/information that is provided in the included study reports – regretfully, reporting in the publications was both poor and inconsistent, and despite our requests, we received no additional data from the study authors. With respect to table 3, see the comment above.
Including the name of the QoL tool in the table would be good also.	Thank you. We included the name of the QoL tool in the table as suggested.
Table 1 could be improved to make it clearer exactly what the intervention was.	Thank you. We added some information about the intervention to table 1 as suggested.
Following on from this, one of the limitations of the study is that you are comparing often very different interventions under the heading of RPM- I don't think this is explicitly acknowledged.	Thank you. As suggested, we added some text to acknowledge the heterogeneity of the interventions.
Another limitation that is not stated is that the vast majority of your study population comes from two studies, and greater than half from a single study. You appear to discuss all of the studies equally.	Thank you. This is a good point. As suggested, we added some text to acknowledge the imbalance in study sizes across the studies.
In the abstract conclusion- you state RPM is associated with health benefits- is this too vague?	Thank you. We revised the text in the abstract as suggested. We present the result for hospitalisations (a result that had low certainty of evidence). Regrettably, we could not conduct a metaanalysis on hospitalisations, so there is no pooled effect estimate for hospitalisations. According to the GRADE

	handbook, effect estimates should not be highlighted for outcomes with very low certainty evidence, so we do not write effect estimates for the other outcomes in the abstract. Also, the word limit for the abstract prevents us from presenting results in more detail in the abstract.
Page 6, line 5 - there is a half sentence that makes no sense	We checked and we believe there are no half sentences in the manuscript, but we are happy to revise any sentences that may be incomplete.
Your outcomes more or less align with the SONG-PD and SONG-HD outcomes but you don't explicitly mention this- maybe you should for increased generalisability.	Thank you. This is a good point. We revised as suggested.
P14 line 9- do you mean 'infections', rather than 'infections not requiring hospitalisation'?	Thank you. We changed to ‘infections’.
P14 line 60- the data suggests or hints at these conclusions, it wouldn't say indicate.	We are unsure what Dr. Bodington means and have not revised the text, but we are happy to do so once we understand what change is suggested.
P15 line 26- you state QoL improved- can you quantify this?	Thank you. As suggested, we added text that quantifies the improvement in QoL.
P15 line 35- you saw no difference in QoL- however here you seem to be suggesting that there is a link to less travel- what is this based on?	Thank you. This is the discussion section, and we are referring to other research – not our study – that suggests there *could* be a link between health-related quality of life and time patients use for travel. We provide the references to this research.
I am unable to comment on the appropriateness of the statistics you have used and this needs further review.	Ok.
Reviewer 2: Dr. Shrikant D. Pande (Changi General Hospital)	
Knowing the limited studies done on this subject in the past, the authors have done a good job to review previous literature. I have no additional suggestions or comments due to the nature of the topic and past studies.	Thank you, we appreciate the positive feedback.
Reviewer 3: Dr. William Levack (University of Otago)	
Overall, I think this review is well conducted and worthy of publication, although I think the clinical	Thank you.

meaningfulness of some of the findings have been overstated (see below). As just an observation: the lack of high-quality clinical trials in this area of practice is perhaps an indication that the conduct of this review was a little premature – more meaningful conclusions may be reached once further studies have been conducted (as the authors point out). Nonetheless, I still think this review is a worth sharing with the international community as a first foray into exploring the evidence in this area of clinical practice.	
I drew on AMSTAR 2 when reviewing this systematic review. Overall, this review scores well on many of the criteria in AMSTAR-2, and certainly all the ones that I (personally) think really matter in terms of minimising major source of bias in a review. However, to improve the reporting of this review the authors may wish to consider adding the following: - A list of potentially relevant studies read in full but excluded as an Appendix, with reason for the exclusion of each.	Thank you. This is a good suggestion. As suggested, we provide a list of all relevant studies read in full text but excluded, with reasons for the exclusion (in appendix).
- Information about the sources of funding for each of the studies included in the review (as another possible source of bias).	Thank you. This is a good suggestion. As suggested, we extracted information about the sources of funding for each included study and give this info.
- A commentary on the possible impact of publication bias on the conclusions of this review. Because there were so few trials included in this review, it would not be possible to statistically evaluate publication bias. However, it would be possible to include a commentary on the potential for publication bias to have impacted on the findings from this review in the discussion section, and a recommendation to include an evaluation of publication bias in any further review on this topic that includes more studies.	Thank you. As suggested, we comment on the possible impact of publication bias.
My main criticism of this paper relates to the interpretation and reporting of the review findings. In general, I believe that the international consensus is that it is misleading to report an outcome being in favour of an intervention when the measure of precision of the estimate of effect (i.e. the 95% CI) includes values that favour a control or comparison group. So, for example, I would not recommend reporting an outcome favouring RPM follow-up (e.g. less all-cause hospitalisations) when the precision of that estimate of effect (i.e. the 95% CI) is very wide	Thank you. We agree that results should not be overstated. Unfortunately, we do not understand the suggestion “to revise the results, .. summary of findings table in a way that does not overstate results”. In the results section, we simply state the results as they are in numbers. We do not understand what the reviewer means with the suggestion to revise these numeric results – we cannot state other numbers than the ones we have extracted from the studies or calculated in our

and includes outcome that favour the control group. I also would recommend not reporting outcomes favouring PRM when the estimate of that effect is reported to be non-significant. This occurs throughout the results section and is used to draw conclusions regarding improved clinical outcomes for people who receive RPM follow-up compared to standard care in the discussion section. The same is true for the data on risk of infection and quality of life. The one outcome where I think there is clear evidence in favour of RPM follow-up is the risk of technical failure, where the meta-analysis shows a positive effect in favour of RPM follow-up, and where the 95% CI does not provide any data conflicting with this conclusion. My recommendation would be to revise the results, discussion, summary of findings table, and abstract in a way that does not overstate results where there is no clear effect in favour of RPM.	metaanalyses. How to express the results in *words* when the CI crosses the line of no difference (which can vary from barely to considerably) -- taking into account also what constitutes appreciable benefit and harm, the magnitude of effect, and optimal information size -- is an ongoing discussion among statisticians, methodologists and others (see e.g. Cochrane handbook and GRADE guideline – which we followed). By way of background: There were 17 individual study outcomes, of which 12 showed no significant difference between the groups (but effect estimate was 11 favouring RPM and 2 favouring control), and 5 showed a statistically significant difference between the groups favouring RPM. Regrettably, it was only possible to pool two outcomes, of which one showed no statistically significant difference (effect estimate favoured RPM) and one was statistically significant favouring RPM. When reporting the synthesised results (pooled, either narratively or statistically, as shown in table 3) in words, we find that we are cautious in our statements. We state in the abstract "indicate there may be positive effects of RPM follow-up", in the discussion "there is no conclusive evidence, but that positive effects of RPM are indicated", and in the conclusion "low to very low evidence that indicate there may be positive effects of RPM follow-up…appears to be safe and provide health benefits". Also, in the results section, we state for the metanalysis "technical failure: result implies benefit of RPM (0.78, 95% CI 0.66, 0.92)". We find that we are fair in our reporting, and conscious not to overstate the results, referring frequently to our GRADE low/very low certainty of evidence to ensure that readers are aware to read the results with caution.
Incidentally, the findings on hospitalisation only seem to be in favour of RPM for adjusted IRR, but I'm unclear how an IRR (a measure of rate of incidents of a particular event; e.g. hospitalisation) should be interpreted for count data, where each 'count' is not a separate event (e.g. hospital days).	Thank you. See our comment to reviewer 1 regarding explanation of IRR. Also, all four studies that reported on hospitalisations – except the small study by Milan 2020 which found RR 1.33 (0.63-2.81) – had effect estimates that were in favour of RPM.

It is also not clear to me what the adjusted IRR data is adjusted for. It may be that some of my criticism above could be addressed by addressing this issue first.	Thank you. In our first submission, we did not state what each analysis in each of the included studies was adjusted for (it varies across the studies), but we have now included this information in the appendix.
Regarding the summary of findings table: It is not clear to me why an assumed risk with a control intervention and with RPM cannot be reported for hospitalisation data, infection data, or QOL data, if data on these outcomes have been reported to draw conclusions about effect sizes in favour (or not in favour) of RPM.	Thank you. This is because we were unable to pool the results for these outcomes, which in turn was because of incomplete and inconsistent outcome reporting in the included studies. Assumed risk cannot be calculated in numbers without the crude data from the included studies.
I suggest that the authors add reference numbers to study IDs in the Tables. It's difficult to match the data in the Table to the information reported in the results section without these additional references.	Thank you. This is a good suggestion and we added reference numbers to the studies in the tables.
The results section in the abstract should contain some statement regard the effect size of RPM follow-up if the conclusion in the abstract reads "there may be positive effects", regardless of whether there was low to very low evidence supporting this finding. Alternatively, the conclusion should be "there is insufficient evidence to draw a conclusion about the benefits of RPM follow-up".	Thank you. There were five outcomes, but we could only calculate a pooled effect estimate on two outcomes, both of which had very low certainty of evidence (shown in table 3). Both the Cochrane handbook and GRADE guideline recommend against showing effect estimates from outcomes with very low certainty in summaries/abstracts. Thus, in the abstract, we offer more information on the outcome which has low certainty of evidence, but there is no pooled effect estimate for this outcome. As explained in these handbooks and guidelines, and in other documents explaining systematic review methodology, it is common to state direction of effect also in the absence of (pooled) effect sizes.