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

TITLE (PROVISIONAL)	Is telehealthcare for heart failure patients cost-effective? An economic evaluation alongside the Danish TeleCare North Heart Failure trial.
AUTHORS	Vestergaard, Anne Sig; Hansen, Louise; Sørensen, Sabrina Storgaard; Jensen, Morten Berg; Ehlers, Lars Holger

VERSION 1 – REVIEW

REVIEWER	H. Bavafa UW-Madison
REVIEW RETURNED	24-Jun-2019

GENERAL COMMENTS	Summary:
	The authors use a randomized controlled trial to evaluate the cost- effectiveness of telemedicine use for heart failure (HF) patients. Using data from 274 patients (134 treatment, 140 control), the authors show that the usage of telemedicine reduces total healthcare costs by 35% (5,668 off a base of 16,241 British pounds). They find no significant impact on patients' health-related quality of life.
	Assessment:
	The paper is a joy to read; the analysis and writing is comprehensive, and the work is at a mature stage. Identifying conditions that can benefit from telemedicine is important as they help with future targeting of such interventions. I share my comments below.
	1. I suggest the authors further clarify their cost calculations / provide an alternative set based on the sample studied. Currently, they divide the costs of telemedicine by the expected number of HF patients in the North Denmark Region (6,700 patients). At its current form the extension from a smaller sample to the full population is a bit confusing without further clarification. E.g., full telemedicine adoption is a strong assumption; can the authors relax it to obtain a more reasonable set of estimates?
	2. Instead of Figure 2, I suggest the authors show their main effect on healthcare costs in a figure, but leave this decision to the editor/authors.
	3. The authors may wish to connect their research to prior papers that have also examined telemedicine-type interventions (e.g., e-visits) on healthcare costs and quality.

4. I suggest the authors report the cost savings in percentage terms (e.g., the 35% in my summary) to help the reader better
interpret the results.

REVIEWER	Alexis Foster
	University of Sheffield, UK
REVIEW RETURNED	12-Jul-2019
GENERAL COMMENTS	This is an excellent paper, well written, communicates well quite technical issues and is an important area of research.
	There are just some minor comments to address:
	Abstract- In the result section NMB and QALYs should not be abbreviations. If short of words just do NNB as that will be less wellknown than QALYS.
	Concluding sentence in the abstract does not seem to link up with the rest of the abstract or the content of the manuscript so maybe rewrite the last sentence to reflect your concluding remarks.
	Introduction Last sentence- 'which is likely to' rather than 'likely is' may read better.
	The second paragraph might be best cut into to, such as starting a new paragraph when introduce the 'In 2016, the European Society'
	Methods First paragraph- I am not sure what is meant by somatic in and outpatient treatments- this just needs clarifying or a different word used.
	P8, Line 51- You may not want to repeat again the 30 days after study commencing as the reader has already read it earlier in the paragraph and its confusing reading it again as makes you question whether this is different or the same as earlier in the paragraph whereas if it was not there the reader just presumes its the same scenario.
	P10- Line 4- I am not sure why suddenly talk about costs of COPD as until then only spoke about HF. Please either explain why bringing in COPD costs or is it a mistake?
	p10- Line 33- Maybe rewrite to say a 'prepaid return envelope was included' just to make it flow better.
	Results section p13, Line 18- When you refer to municipalities it would be useful to just put in brackets a few words to remind the reader what they are or a footnote as its not a phrase/service term used must and by this point the reader needs a reminder.
	p13, Line 19- It would be useful to bring attention that in relation to municipalties was not significant as the authors gloss over that. There's some explanation of the costs but would be useful to link it to the P value not being stat sig.

P16- Line 6- Sense of security- This phrase needs clarification of which is meant by this as you discuss it a couple of times but there's no explanation of what this means.
Throughout you refer to appendix but there's actually a few so need to number your appendicies and refer to the relevant one in the text.
I am not a health economist and have recommended that it is reviewed by a health economist to but I think as a non health economist it was an excellent paper to read.

REVIEWER	Meng Li
	University of Southern California Leonard D. Schaeffer Center for
	Health Policy & Economics, USA
REVIEW RETURNED	15-Sep-2019
	· ·
GENERAL COMMENTS	This study evaluated the cost and outcome of a telehealth program in heart failure patients in Denmark. The methods were appropriate and the writing was clear. The study found that the telehealth program was associated with cost-saving in one year, relative to usual care, and had no adverse effects on quality of life outcomes. This result is not surprising, especially in a short period of time among a group of patients with relatively stable disease. I think this study can benefit from additional insights on the mechanisms and drivers for cost-saving. My specific comments are:
	1. In table 2, instead of the mean cost, I think it would make more sense to include the utilization of different types of service for the treatment and the control groups. The unit cost of teach type of service can then be included in the appendix. Telehealth program can save money in the short term, because it replaces some of the primary care visits and hospitalizations, and its cost is lower. I think showing the utilization of different types of service in table 2 will give the readers, especially those from other countries, a better idea of how much less resource the telehealth group used.
	2. How much did the patients in the telehealth arm use the technology? In another word, what is their adherence? Do we see greater cost-saving in those who used telehealth more? This is relevant because the implementation of a program can affect its outcomes, which in this case is the cost-effectiveness.
	3. Please report the mortality and utility outcomes of the treatment and the control groups, not just the incremental QALY gained.
	4. Please provide the rationale for using seemingly unrelated regression in table 3. Why this specific empirical approach? Also please clarify if the same set of covariates were adjusted for both costs and QALY.
	5. I think the study can benefit from some stratified analyses by NYHA class. Is the cost-saving from less severe or more severe patients? For patients with mild HF, telehealth might be a lower- cost alternative for them to monitor their disease. However, for those with moderate to severe HF, telehealth may lower cost in the short term because it replaces more expensive primary care visits and hospitalizations, but may increase cost in the long term

because some of the early signs for disease progression and complications may be overlooked in telehealth.
6. The study can benefit from some one-way sensitivity analyses examining the main drivers of the results. Are the results driven by reducing hospitalizations, primary care visits, or outpatient visits?
7. To me, it's not surprising at all that telehealth can lead to some cost-saving in the short term, because it replaces some of the more expensive in-person contacts with healthcare providers. The question is, is this going to adversely affect health outcomes? The time horizon of this study is almost too short to answer this, given that the majority of the study population have NYHA class I-II HF. Please discuss this limitation, and the limitation of conducting a trial-based CEA (rather than a model-based CEA).
8. Given the very small difference between treatment and control in quality of life and that generic instruments may not be sensitive enough, please discuss the possibility of using disease-specific measures such as major cardiovascular events as outcomes for this analysis.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: H. Bavafa Institution and Country: UW-Madison Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below Summary:

The authors use a randomized controlled trial to evaluate the cost-effectiveness of telemedicine use for heart failure (HF) patients. Using data from 274 patients (134 treatment, 140 control), the authors show that the usage of telemedicine reduces total healthcare costs by 35% (5,668 off a base of 16,241 British pounds). They find no significant impact on patients' health-related quality of life.

Assessment:

The paper is a joy to read; the analysis and writing is comprehensive, and the work is at a mature stage. Identifying conditions that can benefit from telemedicine is important as they help with future targeting of such interventions. I share my comments below.

1. I suggest the authors further clarify their cost calculations / provide an alternative set based on the sample studied. Currently, they divide the costs of telemedicine by the expected number of HF patients in the North Denmark Region (6,700 patients). At its current form the extension from a smaller sample to the full population is a bit confusing without further clarification. E.g., full telemedicine adoption is a strong assumption; can the authors relax it to obtain a more reasonable set of estimates?

Answer. We agree that full telemedicine adoption is a strong assumption. This is however, a national decision in Denmark. Denmark is a special country concerning telehealthcare where it is a central part of the Danish Digitalisation Strategy and a specific issue in the Finance Act. The decision is that all patients should have the offer – so full adoption is not that every patient actually end up using the Telekit, but rather that all patients should have the offer to use it. In this way, we believe full 'adoption' is the correct assumption for the evaluation.

Please also note that the methods used in this paper is the same as the method we used regarding the national roll out of telemedicine for COPD in Denmark (Udsen et al. BMJ Open. 2017 May 17;7(5):e014616. doi: 10.1136/bmjopen-2016-014616.)

We also plan to do a separate paper on the budget impact of a national roll out in Denmark. The preliminary results will be presented at the ISPOR conference in November 2019.

2. Instead of Figure 2, I suggest the authors show their main effect on healthcare costs in a figure, but leave this decision to the editor/authors.

Answer. Thank you for the suggestion. According to international guidelines for economic evaluation a scatterplot or CEAC is suggested as illustration of the main result. We therefore want to keep Figure 2 in the paper. We believe the main effect on healthcare costs are presented in detail in Table 2.

3. The authors may wish to connect their research to prior papers that have also examined telemedicine-type interventions (e.g., e-visits) on healthcare costs and quality.

Answer. Thank you, -this is an interesting suggestion. There are many different types of telemedicine, and a more broad view on cost-effectiveness could indeed be interesting. We have chosen, however, to refer only to a few selected studies on telemedicine due to lack of space. Due to the similarity to the former TeleCare North COPD study we prefer to compare with the results from that particular Danish study. Furthermore, we only address a few other studies on telehealthcare in HF. We cannot find any more recent studies that we should have included.

4. I suggest the authors report the cost savings in percentage terms (e.g., the 35% in my summary) to help the reader better interpret the results.

Answer. This seems to be a good idea. We have inserted savings in percentage terms in the abstract and inserted your summary in the main results.

Reviewer: 2

Reviewer Name: Alexis Foster Institution and Country: University of Sheffield, UK Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below This is an excellent paper, well written, communicates well quite technical issues and is an important area of research.

There are just some minor comments to address:

Abstract- In the result section NMB and QALYs should not be abbreviations. If short of words just do NNB as that will be less wellknown than QALYS.

Answer. Thank you for the suggestion. We have re-written the abstract and do not use abbreviations without explanation.

Concluding sentence in the abstract does not seem to link up with the rest of the abstract or the content of the manuscript so maybe rewrite the last sentence to reflect your concluding remarks.

Answer. Thank you for the suggestion. We have deleted this last sentence. Instead, we have added a limitation in the 'Strength and limitations' section about the issue of the lack of understanding of the mechanisms of savings.

Introduction

Last sentence- 'which is likely to' rather than 'likely is' may read better.

Answer. OK. This is now changed.

The second paragraph might be best cut into to, such as starting a new paragraph when introduce the 'In 2016, the European Society....'

Answer. Thank you for the suggestion. This is now changed.

Methods

First paragraph- I am not sure what is meant by somatic in and outpatient treatments- this just needs clarifying or a different word used.

Answer. OK. This is now changed. We have inserted another explanation.

P8, Line 51- You may not want to repeat again the 30 days after study commencing as the reader has already read it earlier in the paragraph and its confusing reading it again as makes you question whether this is different or the same as earlier in the paragraph whereas if it was not there the reader just presumes its the same scenario.

Answer. Thank you for the suggestion. This is now deleted.

P10- Line 4- I am not sure why suddenly talk about costs of COPD as until then only spoke about HF. Please either explain why bringing in COPD costs or is it a mistake?

Answer. This is not a mistake. We have rephrased it to other patients using the regional telehealth system.

p10- Line 33- Maybe rewrite to say a 'prepaid return envelope was included' just to make it flow better.

Answer. OK. This is now changed.

Results section

p13, Line 18- When you refer to municipalities it would be useful to just put in brackets a few words to remind the reader what they are or a footnote as its not a phrase/service term used must and by this point the reader needs a reminder.

Answer. Good point. We have deleted the sentence about municipalities here. But we have inserted in brackets in the methods section for a better explanation of the municipalities as '(administrative units for tax-financed local health and social care)'.

p13, Line 19- It would be useful to bring attention that in relation to municipalties was not significant as the authors gloss over that. There's some explanation of the costs but would be useful to link it to the P value not being stat sig.

Answer. Thank you for the suggestion. We have deleted the sentence here about municipalities

P16- Line 6- Sense of security- This phrase needs clarification of which is meant by this as you discuss it a couple of times but there's no explanation of what this means.

Answer. Thank you. We have re-written this. Now we state it more broadly as changed opinions and beliefs.

Throughout you refer to appendix but there's actually a few so need to number your appendicies and refer to the relevant one in the text.

Answer. Good point. We have split it into three parts. Appendix A provides detailed information regarding the Danish TeleCare North Heart Failure Trial. Appendix B gives information on the cost estimates associated with the telehealthcare solution. Appendix C is for the description of imputation approach. We have changed references in the main text accordingly.

I am not a health economist and have recommended that it is reviewed by a health economist to but I think as a non health economist it was an excellent paper to read.

Reviewer: 3

Reviewer Name: Meng Li

Institution and Country: University of Southern California Leonard D. Schaeffer Center for Health Policy & Economics, USA Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below This study evaluated the cost and outcome of a telehealth program in heart failure patients in Denmark. The methods were appropriate and the writing was clear. The study found that the telehealth program was associated with cost-saving in one year, relative to usual care, and had no adverse effects on quality of life outcomes. This result is not surprising, especially in a short period of time among a group of patients with relatively stable disease. I think this study can benefit from additional insights on the mechanisms and drivers for cost-saving. My specific comments are:

1. In table 2, instead of the mean cost, I think it would make more sense to include the utilization of different types of service for the treatment and the control groups. The unit cost of teach type of service can then be included in the appendix. Telehealth program can save money in the short term, because it replaces some of the primary care visits and hospitalizations, and its cost is lower. I think showing the utilization of different types of service in table 2 will give the readers, especially those from other countries, a better idea of how much less resource the telehealth group used.

Answer. Thank you for the suggestion -we agree. We have looked into the each type of service in order to try to give the readers a better understanding of how telehealthcare may have changed the utilisation of the different types of healthcare services. The majority of inpatient services were heart-related (approximately 60-65% in the before period in both groups; and approximately 40-50% in the after period in both groups). The distribution between types of outpatient services (disease-specific, procedures and telemedicine) were similar at baseline between groups. Telehealthcare apparently led to a different distribution of service utilization between groups, with disease-specific and procedures decreasing in the intervention groups and telemedicine services increasing in the intervention group.

We have chosen not to go further into this area in this paper. It is a very difficult question that should be addressed in a separate paper. We will seriously consider doing this paper. Instead, we have addressed this limitation in the 'Strength and limitations' section.

2. How much did the patients in the telehealth arm use the technology? In another word, what is their adherence? Do we see greater cost-saving in those who used telehealth more? This is relevant because the implementation of a program can affect its outcomes, which in this case is the cost-effectiveness.

Answer. Thank you for the suggestion, we agree. After submission, we gained access to tele monitoring data for all 145 patients originally included the intervention group. All patients except five were active users of the Telekit equipment. One was excluded from our analysis due to lack of baseline EQ-5D data. The other four patients left the trial after a short period of time (between 1 - 5 month). The median number of day the patients use the equipment was below 20% of the days of study, but the majority of the patients seems to use the equipment more often that they were supposed to (they are asked to report data only once a week or every 14^{th} day). Please remember, there were no financial incentives for the patients to use or report tele data. We do not wish to add these data to this study, but maybe in a follow up paper (see our answer above).

3. Please report the mortality and utility outcomes of the treatment and the control groups, not just the incremental QALY gained.

Answer. Thank you for the suggestion. This has been added to the results section. 'The adjusted baseline utility score was similar across the two groups (0.7079 for control and 0.7075 for intervention). The mortality was similar between both groups, with 5 deaths in the control group and 7 deaths in the intervention group.'

4. Please provide the rationale for using seemingly unrelated regression in table 3. Why this specific empirical approach? Also please clarify if the same set of covariates were adjusted for both costs and QALY.

Answer. We have now added the explanation. Seemingly unrelated regression is a recommended and widely used method because cost and HRQoL normally is correlated. There is a difference between the variables because EQ-5D has a ceiling effect that we need to control for. We believe this is standard knowledge for health economists performing economic evaluations.

5. I think the study can benefit from some stratified analyses by NYHA class. Is the cost-saving from less severe or more severe patients? For patients with mild HF, telehealth might be a lower-cost alternative for them to monitor their disease. However, for those with moderate to severe HF, telehealth may lower cost in the short term because it replaces more expensive primary care visits and hospitalizations, but may increase cost in the long term because some of the early signs for disease progression and complications may be overlooked in telehealth.

Answer. We have checked the robustness in the sensitivity analyses by excluding the 10% most severe patients and by including NYHA class I. This did not change the general picture of substantial savings. The idea of stratified analyses is good. However, this is a small study with only 134 patients in the intervention group (and e.g. only 14 patients in NHYA 4 in the intervention arm and 8 patient in the control arm) and therefore a stratified analysis was considered unrealistic. We further agree, that we need to investigate long term cost-effectiveness in another study. This is explicitly stated in the 'Strengths and limitations' section.

6. The study can benefit from some one-way sensitivity analyses examining the main drivers of the results. Are the results driven by reducing hospitalizations, primary care visits, or outpatient visits?

Answer. We believe this is already made explicit in text and table 2. The savings are mainly driven by cheaper hospitalizations.

7. To me, it's not surprising at all that telehealth can lead to some cost-saving in the short term, because it replaces some of the more expensive in-person contacts with healthcare providers. The question is, is this going to adversely affect health outcomes? The time horizon of this study is almost too short to answer this, given that the majority of the study population have NYHA class I-II HF. Please discuss this limitation, and the limitation of conducting a trial-based CEA (rather than a model-based CEA).

Answer. We agree. This is only a short term (one-year time horizon) study of cost-effectiveness. The conclusion may change over a longer time horizon. This is explicitly stated as a limitation of the study design.

8. Given the very small difference between treatment and control in quality of life and that generic instruments may not be sensitive enough, please discuss the possibility of using disease-specific measures such as major cardiovascular events as outcomes for this analysis. Answer. We believe this is already included in the discussion. We have already made reference to another published study from the TeleCare North Heart Failure Trial showing no effect on the HF disease-specific questionnaire Kansas City Cardiomyopathy Questionnaire 12 (KCCQ12).

VERSION 2 – REVIEW

REVIEWER	Alexis Foster
	University of Sheffield, UK
REVIEW RETURNED	22-Oct-2019

GENERAL COMMENTS	Many thanks for addressing the reviewers' comments. The re- written abstract is great and you have taken on board the comments. I am satisfied the changes I suggested had been made (I was reviewer 2). It is more important that Reviewer 3 feels their comments have been addressed as related to technical details.
------------------	---

REVIEWER	Meng Li University of Southern California, USA
REVIEW RETURNED	20-Oct-2019

GENERAL COMMENTS	 Several of my comments were overlooked by the authors. 1. I still think the study can benefit from some one-way sensitivity analyses to understand the potential range of cost savings. Please discuss the limitation of conducting a trial-based CEA (rather than a model-based CEA). I etill thick its insertent to include utilization data, anther them
	3. I still think it's important to include utilization data, rather than just costs, either in Table 2 or in Appendix in this type of analysis. Costs are notoriously nontransferable across borders, and it is hard for researchers from outside of Denmark to gauge how much cost saving would be achieved if such telehealth program is to be implemented in their country. Perhaps the authors wrote this paper only for researchers in Denmark, but that would limit the significance and generalizability of the findings. However, I'll leave this for the editor to decide.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name: Alexis Foster

Institution and Country: University of Sheffield, UK Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below Many thanks for addressing the reviewers' comments. The re-written abstract is great and you have taken on board the comments. I am satisfied the changes I suggested had been made (I was reviewer 2). It is more important that Reviewer 3 feels their comments have been addressed as related to technical details.

Reviewer: 3

Reviewer Name: Meng Li Institution and Country: University of Southern California, USA Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below Several of my comments were overlooked by the authors.

1. I still think the study can benefit from some one-way sensitivity analyses to understand the potential range of cost savings.

Answer: We are sorry if we did not provide good enough answers in our last reply. We most certainly did not overlook any of your comments. We agree with your general view that one-way sensitivity

analyses are beneficial and can help readers understand the potential for cost savings. However, as we also tried to explain earlier, we cannot do more than we have already done. The reason for this is the following: The data that is available for our analyses does not give us the same possibilities to do one-way sensitivity analyses as e.g. a modeling study. We do not have detailed billing information like in the US either. We have acquired Danish register data from four different registries. Information on hospital costs, for instance, were taken from the Danish National Patient Registry and the costing of hospital services is based on Danish DRG case-mix system. This method has been used for economic evaluation alongside trials many times before, but the limitations to this kind of data might not be common knowledge. We are sorry for this confusion, and we should have explained it better before. For example, we typically have one DRG tariff for the total cost per patient contact (i.e. one figure per patient per hospital stay or per outpatient visit) and we are quite sure this figure is unbiased and a good estimate of the cost of the total hospital service. However, we cannot see how much of the cost is due to length of inpatient stay or due to medicine, procedures etc. Thus, we do not have information on resource usage or data to do this kind of detailed one-way sensitivity analyses that you request. The type of information on patient specific data that we have from the Danish registers and the use of the Danish DRG system has been show on many previous occasions to provide an unbiased estimate of the mean and total cost, but unfortunately, there are many limitations with this kind of register data. We hope this answer gives you a better understanding of the limitations in our dataset. As we argued previously, it will require a different type of data set (or a decision model) to do the one-way sensitivity analyses that you ask for. We find your comments relevant, but we are unable to provide more detailed analyses than we have already performed. Please note that we have done other types of sensitivity analyses e.g. a probabilistic sensitivity analysis using the kind of data we have to a maximum. If we had made a model-based analysis instead of the cost-effectiveness analysis alongside the trial, we could have made a Tornado diagram (or similar overview of all oneway sensitivity analyses). We hope this is a satisfactory answer.

2. Please discuss the limitation of conducting a trial-based CEA (rather than a model-based CEA).

Answer: We agree. We have added this as an extra limitation in the short list of strength and limitations: "Trial-based economic evaluations are limited by truncated time horizons, difficulty in generalizing to other settings, and failure to incorporate evidence from other trials or observational studies."

Because of the confusion about the Danish register data and the pros and cons of this type of data, we have also deleted the term "micro costing" in the strength and limitation section. We still believe we have used micro costing in other parts of our evaluation, however, DRG tariffs are used for costing of hospital services and this is strictly speaking a case-mix costing system.

3. I still think it's important to include utilization data, rather than just costs, either in Table 2 or in Appendix in this type of analysis. Costs are notoriously nontransferable across borders, and it is hard for researchers from outside of Denmark to gauge how much cost saving would be achieved if such telehealth program is to be implemented in their country. Perhaps the authors wrote this paper only for researchers in Denmark, but that would limit the significance and generalizability of the findings. However, I'll leave this for the editor to decide.

Answer: We agree. It is a good idea to include an overview of the differences in resource utilization between the two groups in the trial. The inclusion of utilization data would provide valuable insight for readers and perhaps give a better understanding of the differences in costs between the groups and the possible savings. However, we cannot include this. As we have explained above, this is unfortunately not possible with the type of data that is available for this analysis. We are actually not certain either, that it will be possible at all to conduct such an in debt analysis because this kind of data is not collected in a systematical way in Danish Registers. For this economic evaluation, we have

accessed four different Danish databases. We agree that it would be interesting to try to explain the mechanisms of savings better. One way to try to do this is to look further into the differences in the patient pathways and resource usage between the intervention and control group. This would require access to patient records for all patients, and this would be a new study.

VERSION 3 – REVIEW

REVIEWER	Alexis Foster
	University of Sheffield, UK
REVIEW RETURNED	20-Dec-2019
GENERAL COMMENTS	Many thanks for addressing the comments.
REVIEWER	Meng Li
	University of Southern California, USA
REVIEW RETURNED	25-Dec-2019
GENERAL COMMENTS	The authors satisfactorily addressed my comments.