PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Economic evaluation of a phase III international randomised controlled trial of very early mobilisation after stroke (AVERT)
AUTHORS	Gao, Lan; Sheppard, Lauren; Wu, O; Churilov, Leonid; Mohebbi, Mohammadreza; Collier, Janice; Bernhardt, Julie; Ellery, Fiona; Dewey, Helen; Moodie, Marj

VERSION 1 - REVIEW

REVIEWER	jim Burke University of Michigan, US
REVIEW RETURNED	19-Sep-2018

GENERAL COMMENTS	Gao et al present the results of a cost-utility analysis of the AVERT trial comparing very early mobilization to usual care. They found that VEM (as with the main paper) resulted in a slight reduction in QALYs, and slight (but highly non-significant) reductions in cost. While this paper appears to have been well executed , I struggled with the very notion of performing a cost effectiveness analysis on a negative trial. Yet, I think there is quite a bit of interesting and potentially important data here. This is a large, recent and reasonably representative sample of stroke stroke patients with detailed measures of functional status, quality of life and self-reported costs (including indirect costs). These data seem very well positioned to build updated models of costs (particularly the indirect costs) of stroke. Major Issue: 1. Motivation for the paper – My biggest concern is with the theoretical setup of this paper. A cost-utility analysis of AVERT would have been of great importance IF the study had been positive. But, for an intervention that does not have net benefit, what is the relevance of studying the either the utility or the costs? The introduction tries to set this question up my arguing about the importance of long-term outcomes. While, in general, I certainly agree that long-term outcomes are important, there is little reason to believe they're relevant in this case. Via what mechaism would AVERT have improved long-term outcomes without improving short-term outcomes? If the authors' had found improved long term outcomes combined with a trend towards worse short term outcomes, wouldn't the most likely interpretation be that this was a chance finding?

	In general, I can see how this study made a lot of sense before the main result of AVERT was known. But, I don't see what it tells us once the main result is known. Minor Issues: 1. The relationship between how MRS and QALYs were modeled in this study is confusing. Initially, I wondered why the MRS was reported at all. If QALYs are measured by the aqol-4d instrument, what does the long-term MRS tell us? Yet, I was concerned that there would too much noise in the aqol-4d measurements to make any interpretation of differences between groups. Ultimately, it seems that this is what the authors found. For this manuscript, I'd suggest streamlining these methods to describe what was ultimately done. Separately, it may be of interest to describe the relationship between MRS and QALYS, with a particular emphasis on factors that predict their divergence. 2. How reliable are the self-reported utilization measures that drive the cost measures? I can imagine that, given the complexity of care for many of these patients, that reliability may be suboptimal. Validation of these measures would greatly strengthen the cost portions of the manuscript. The comprehensiveness of the cost measures is a strength, particularly on the indirect costs. It would have been helpful to see a cost/utilization break down in the main manuscript. 3. How was PT time measured? How should we interpret differences in costs between groups, other than PT time given that VEM doesn't change outcomes — as PT time seems as if it is the only measure that was designed to differ between groups. 4. The cost of a rehab admission in Australia seems implausibly low. In general, the variation between countries and between cost centers in stroke costs in the supplemental table 1 is fascinating. The discussion briefly alludes to some of these differences (e.g. length of stay), but i suspect even more could be said. 5. I'm not sure that i followed the description on the regional variation in cost effectiveness in the discussion — were cost/qalys statistically significant and different across countries? My interpretation was that while differences existed (necessarily), that they were fairly small and non-significant, but i may have been misinterpreting coefficients. 6. I'm not clear on the timecourse of the cost-utility analysis. 7. The confidence intervals on the productivity cost differences between groups seem implausibly narrow and the mean effects seem implausibly small. Can you give some context on why this might make sense? 8. If the authors are going to talk about the probability of cost-effetiveness, it'd be helpful to contextualize that finding by discussing about the probability of net effectiveness, first. 9. Is it believable that VEM has lower net costs, without improving outcomes? My intuition is that its not very plausible and suggests that the finding is noise? What was the prior probability that VEM would not improve MRS or QALYs, but would reduce costs? It had to be extremely low.

	10. In general, I got the sense that the manuscript often referred to non-significant mean differences as “differences”. However, isn’t it possible that all of these “differences” (particularly in costs) are just random error? If so, should they be described as “differences” at all/

REVIEWER	Roberta Longo University of Leeds, UK
REVIEW RETURNED	25-Oct-2018

GENERAL COMMENTS	The paper is a thorough investigation on the cost-effectiveness of very early intervention for stroke patients alongside an international trial. The challenges posed by the design of the trial, with centres across the world, are very well handled and any limitations acknowledged. Methods are robust and results well presented. My only comment would be to proof-read the manuscript making sure that acronyms are defined before they are first used to aid comprehension, e.g. NIHSS and LOS are not defined the first time around

REVIEWER	Paolo Landa university of Exeter, United Kingdom
REVIEW RETURNED	04-Dec-2018

GENERAL COMMENTS	The paper is well reported and the supplementary material is big enough to answer any question of the reader. I ask you few modifications: 1) some acronyms are not explained in the first time they appear (e.g. Los, NIHSS) 2) in the recent literature on rehabilitation after stroke you can include: Dean et al. (2018) Community-based rehabilitation training after stroke: results of a pilot randomised controlled trial (ReTrain) investigating acceptability and feasibility February 2018BMJ Open 8(2):e018409 Norris et al. (2018) Acceptability and experience of a functional training programme (ReTrain) in community-dwelling stroke survivors in South West England: A qualitative study July 2018BMJ Open 8(7):e02217 3) some sections are too brief, you need to provide more information and more detail to help the reader, for example the outcomes and the costs. 4) about the date of currency conversion, it is not reported, there is only the year, but in a year you need a specific average value. 5) in the last part of outcomes , you report that you had inherent difficulties. can you explain in more detail? 6) when you write that there were 56 stroke units across five geographical jurisdiction, can you say the distribution of the units? The overall work is well reported

VERSION 1 – AUTHOR RESPONSE

Table 1 Response to reviewers' comment

Reviewer' comment	Response	Revision
Reviewer No. 1		
1. Motivation for the paper – My biggest concern is with the theoretical setup of this paper. A cost-utility analysis of AVERT would have been of great importance IF the study had been positive. But, for an intervention that does not have net benefit, what is the relevance of studying the either the utility or the costs? The introduction tries to set this question up my arguing about the importance of long-term outcomes. While, in general, I certainly agree that long-term outcomes are important, there is little reason to believe they're relevant in this case. Via what mechanism would AVERT have improved long-term outcomes without improving short-term outcomes? If the authors' had found improved long term outcomes combined with a trend towards worse short term outcomes, wouldn't the most likely interpretation be that this was a chance finding? In general, I can see how this study made a lot of sense before the main result of AVERT was known. But, I don't see what it tells us once the main result is known.	Regardless of benefit, neutral or harm outcomes, researchers need to publish their per protocol research results. We appreciate these comments as we needed to justify the paper given trial results, ensure the paper is of interest, and we have reworked our introduction for this paper. In the field of stroke research, there has been a general consensus that the end point is measured at 3-month without any long-term follow up. As a consequence, there is a paucity of long-term costing data for patients post stroke, especially for the costs relating to rehospitalisation, rehabilitation service, community services, and costs in non-health sectors (i.e. informal care, productivity loss). From a costing perspective, even if the intervention did not contribute to improved outcomes for patients, given the size of the patient sample recruited and its broader representativeness (across five countries from developing to developed world), there is the potential to add to knowledge around the cost of care for stroke, especially in the long-term. In our opinion, this is the primary motivation to pursue the economic evaluation in this paper. We have revised the introduction to reflect this shift in focus.	Page 7-8 of the revised manuscript (the highlighted are newly added). In the research field of stroke, primary endpoint is usually assessed at month 3 after stroke[11-14], which means there is a paucity of data in terms of long-term resource use and cost of care for patients with stroke. Given AVERT provided a longer-term (i.e. 12 months) comprehensive measurement of costs relating to stroke care (i.e. direct medical, direct non-medical, and indirect costs), and the broader representativeness of patients across countries and regions (>2000 patients were recruited from both developing and developed world), together with the implications of stroke economic burden sustained beyond the acute phase (i.e., 3 months), holistically examining the cost of stroke care that falls within health and non-health sectors could potentially advance understanding of pattern of resource use post stroke and identify any gaps to improve care for stroke and chances to curb the increasing economic burden of disease. This examination also benefits healthcare funders (i.e. governments, insurance companies) and the public with addition of substantial knowledge of long-term rehabilitation cost for stroke. This economic evaluation, which was part of the registered trial protocol (Australian New Zealand Clinical Trials Registry, ACTRN12606000185561) and planned prior to knowledge of outcomes, was conducted alongside the Phase III RCT,.[10] The aim of this paper is to assess the cost-effectiveness of very early mobilisation within 24 hours

		after stroke in terms of improving patient outcomes at 12-months, in comparison to usual care (UC), with a particular focus on examining the resource use and cost of care after stroke.
2. The relationship between how MRS and QALYs were modeled in this study is confusing. Initially, I wondered why the MRS was reported at all. If QALYs are measured by the aqol-4d instrument, what does the long-term MRS tell us? Yet, I was concerned that there would too much noise in the aqol-4d measurements to make any interpretation of differences between groups. Ultimately, it seems that this is what the authors found. For this manuscript, I'd suggest streamlining these methods to describe what was ultimately done. Separately, it may be of interest to describe the relationship between MRS and QALYS, with a particular emphasis on factors that predict their divergence.	The economic evaluation was undertaken in accordance with the published economic evaluation protocol (the protocol set out mRS and QALY as the outcomes; Sheppard L, Dewey H, Bernhardt J, et al. Economic Evaluation Plan (EEP) for A Very Early Rehabilitation Trial (AVERT): An international trial to compare the costs and cost-effectiveness of commencing out of bed standing and walking training (very early mobilization) within 24h of stroke onset with usual stroke unit care. Int J Stroke 2016;11(4):492-94. doi: 10.1177/1747493016632254).To avoid being selective in reporting, these are the outcomes are presented. The relationship between mRS and QALY was comprehensively investigated in a separate paper (accepted by International Journal of Stroke on 10th Dec 2018). As per reviewer's comment, we have removed the details of mRS outcome to the appendix.	Page 8 of the revised manuscript (the highlighted are newly added). Due to the inherent difficulties of administering the AQoL instrument to acute stroke patients (i.e. most of patients were not able to respond to these questions at baseline), the mRS score at baseline10 was used as a surrogate measure of patient utility during the acute phase. The detailed methods of this work are reported elsewhere[19] and a brief description is supplied in the online Supplementary document 1. A comprehensive exploration in the relationship between mRS score and AQoL-4D is presented separately to this paper[20]. The outcome of mRS has been moved to the Supplementary document 5 of the revised online supplementary material.
3. How reliable are the self-reported utilization measures that drive the cost measures? I can imagine that, given the complexity of care for many of these patients, that reliability may be suboptimal. Validation of these measures would greatly strengthen the cost portions of the manuscript. The comprehensiveness of the cost measures is a strength, particularly on the indirect costs. It would have been helpful to see a cost/utilization break down in the main manuscript.	The cost-CRF was completed by a trained team member either via interviewing the patient or the patient's carer/immediate relative at months 3 and 12 respectively. It is believed that in this way, the data obtained achieved maximum accuracy as opposed to a self-filled resource use questionnaire or diary. Some of the outcomes reported in the cost-CRF was cross-validated by the medical records extracted from participating hospitals (e.g. length of stay for the index hospitalisation and discharge place etc.). Given a lack of reliable sources, some of the other utilisation outcomes were unable to be verified (e.g. change in productivity). As per reviewer's comments, the details on cost and	Pages 9 to 11 of the revised manuscript (Please note the following paragraphs have been moved from the Supplementary document 3 of the previous version of the Online Supplementary Material to reflect the shift in the focus of the paper). Resource use All resource use during the study period was electronically collated using a validated Cost Case Report Form (Cost CRF) administered and recorded by trained staff at 3- and 12-months using face to face assessments with patients and carers, and medical records. Cost CRF used in Australia is supplied as an

	resource utilisation are moved to the main manuscript to reflect the focus on resource use and cost.	example (Supplementary document 2). Cost CRF from other participating countries could be requested from corresponding author. Healthcare resource use The quantity of resources used for the following health care resource items was recorded: number of ambulance transfers (emergency and non-emergency), acute hospitalisation (including length of stay, LoS), rehospitalisation (number of occasions and LoS for each occasion), rehabilitation hospital admission (number of occasions and LoS for each occasion), outpatient rehabilitation program (number of occasions and number of days for each occasion), rehabilitation provided at home/nursing facility (number of occasions and number of sessions for each occasion), private physiotherapy (number of sessions), respite care (number of sessions) and individual outpatient (including physiotherapy, occupational therapy, and speech and language therapy) visits (service type and number of sessions) for patients from United Kingdom, Singapore and Malaysia only. Non-healthcare resource use The quantity of resources used was recorded for the following non-heath care resource items: accommodation move due to stroke (location moved to and date of move), community service (type of service use and number of service used both for prior to and post-stroke), home modification (type of modification, supplier and cost), special equipment and aids (type of equipment/aids and quantity consumed), informal care (purpose of the care and hours used), live-in maids (number of maids prior to and post

		stroke) (for Singapore and Malaysia only), changes to employment (employment status and weekly hours of working both prior to and post-stroke). Resource use reported at 3 (i.e. resources used between 0 and 3 months) and 12 (i.e. resources used between 4 and 12 months) months was used to calculate the total annual resource use for each participant. Generally, where patients were still using a particular resource at the time of 12-month data collection, the last day of 12 months' follow-up (calculated from the day of index stroke) was used to estimate the duration of that resource utilisation. In the event of a patient dying, resource use data for the period prior to death was ascertained from their carer and medical records, wherever possible. The revised Tables 1 and 2 are appended to this table for reviewer's reference.
4. How was PT time measured? How should we interpret differences in costs between groups, other than PT time given that VEM doesn't change outcomes — as PT time seems as if it is the only measure that was designed to differ between groups.	PT time was recorded by a log documented by each participating PT. As the design of the VEM intervention was to promote earlier and more mobilisation after stroke, the PT time in the VEM was supposed to be greater than the control. And the results in terms of both PT time and cost reflected that.	No change made
5. The cost of a rehab admission in Australia seems implausibly low. In general, the variation between countries and between cost centers in stroke costs in the supplemental table 1 is fascinating. The discussion briefly alludes to some of these differences (e.g. length of stay), but i suspect even more could be said.	The unit cost of rehabilitation hospital admission for Australia provided in Supplementary document 3: Table I is a daily cost (there is a footnote underneath that table to indicate such). As shown in Supplementary document 6: Table III, generally, the rehab admission median costs were comparable between Australia ($13,134) and New Zealand ($11,262).	No change made

6. I'm not sure that I followed the description on the regional variation in cost effectiveness in the discussion — were cost/qalys statistically significant and different across countries? My interpretation was that while differences existed (necessarily), that they were fairly small and non-significant, but Imay have been misinterpreting coefficients.	The country-specific ICERs were based on point estimates in terms of differences in cost and QALY, so they were different across countries. But reviewer is right in that if the 95%CI is taken into account, none of the differences are significant. To avoid confusion and over-interpretation, we have deleted these discussions.	Page 19 of the revised manuscript (the highlighted are newly added). This multinational trial also revealed that in managing patients post-stroke, practice of stroke care varied from country to country. Although 100% of patients with stroke were hospitalised for the initial acute care, the LoS differs significantly greatly, ranging from 4 days (Malaysia) to 25 days (New Zealand), which might be attributable to the different severity of stroke and/or differences in clinical practice care processes. Moreover, in Malaysia, patients tended to receive rehabilitation services in an outpatient (i.e. up to 52% of patients received the outpatients rehabilitation program services) rather than inpatient (i.e. only up to 2% patients were admitted to rehabilitation hospital) setting; and patients were less likely to utilise ambulant transfer and apply home modifications, as compared to participants from other countries. This might be a signal for future study around stroke care in Malaysia, research potentially could be helpful to improve the service delivery for outpatient rehabilitation program. Patients from western countries consumed more community services and rehabilitation services that provided at home/nursing home than their Asian counterparts, which reflects the difference in social welfare and healthcare systems.
7. I'm not clear on the timecourse of the cost-utility analysis.	Conventionally, the QALY gain for each group should be calculated based on area under curve approach by using utility weight measured at various time points (in this case, baseline, Month 3 and Month 12). However, as discussed on page 12, due to the absence of baseline utility measurement and unreliable estimates from mapping method, we were only able to estimate the between-group difference using the Month 12 utility outcome (but adjusted for baseline	No change is made

	mRS score and NIHSS, which represent the severity of stroke).	
8. The confidence intervals on the productivity cost differences between groups seem implausibly narrow and the mean effects seem implausibly small. Can you give some context on why this might make sense?	For the productivity cost, in Supplementary document 6: Table III, the mean and median costs were $46 (SD246), $0 (IQR 0-0) and $58 (SD312), $0 (IQR 0-0) in the VEM and UC groups respectively. In Supplementary document 7, the proportion of patients worked prior to stroke was 27% in the VEM and 26% in the UC groups while this proportion reduced to 14% in the VEM and 12% in the UC groups after stroke. From all these estimates, we believe there is no strong evidence to show a significant between-group difference in terms of the productivity cost.	No change is made
9. If the authors are going to talk about the probability of cost-effectiveness, it'd be helpful to contextualize that finding by discussing about the probability of net effectiveness, first.	The probability of cost-effective (i.e. ICER<$50,000/QALY) was calculated based on 2000 bootstrapped iterations of ICER. For each iteration, between-group difference in both cost and QALY was sampled together to generate an ICER. And then the proportion of iterations with ICER that fell below $50,000/QALY threshold was calculated, which is the probability of being cost-effective. Given the VEM intervention is unlikely to be cost-effective, the probability estimates are removed in the revised manuscript.	Page 17 of the revised manuscript. The 12 months within-trial cost-effectiveness analysis showed that VEM was unlikely to be cost-effective than UC in patients with stroke. Between-group differences in costs and benefits (probability of achieving a favourable outcome of mRS and differences in QALYs) over the one year study period were not significant from a health sector perspective.
10. Is it believable that VEM has lower net costs, without improving outcomes? My intuition is that it's not very plausible and suggests that the finding is noise? What was the prior probability that VEM would not improve MRS or QALYs, but would reduce costs? It had to be extremely low.	Thanks for reviewer's comments. The results in Table 1 of the main body showed that if the healthcare system perspective was taken, VEM was associated with higher cost than that in the UC group (some of the point estimates showed the VEM has lower cost, but non-significant). We have revised the manuscript thoroughly to avoid using any expression that may suggest VEM associated with less net cost.	Due to the substantial changes made, please refer to the revised manuscript for the changes.
11. In general, I got the sense that the manuscript often referred to non-significant mean differences as "differences". However, isn't it possible that all of these	If only taking point estimates into account, there were some between-group differences, but when the 95% CI was considered in addition to the point estimate, it rendered the	Due to the substantial changes made, please refer to the revised manuscript for the changes.

"differences" (particularly in costs) are just random error? If so, should they be described as "differences" at all/	differences insignificant. In order to address reviewer's concern, we have revised the manuscript accordingly to minimise the use of 'non-significant difference'.	
Reviewer No. 2		
12. The paper is a thorough investigation on the cost-effectiveness of very early intervention for stroke patients alongside an international trial. The challenges posed by the design of the trial, with centres across the world, are very well handled and any limitations acknowledged. Methods are robust and results well presented. My only comment would be to proof-read the manuscript making sure that acronyms are defined before they are first used to aid comprehension, e.g. NIHSS and LOS are not defined the first time around	Great thanks to the reviewer's positive comments. We have revised the manuscript as per the comments.	The following acronyms are now defined when first used: VEM (very early mobilisation), UC (usual care), NIHSS (National Institute of Health Stroke Severity), LoS (length of stay).
Reviewer No. 3		
13. Some acronyms are not explained in the first time they appear (e.g. Los, NIHSS)	We have revised the manuscript as per the comments.	The following acronyms are now defined when first used: VEM (very early mobilisation), UC (usual care), NIHSS (National Institute of Health Stroke Severity), LoS (length of stay).
14. In the recent literature on rehabilitation after stroke you can include: Dean et al. (2018) Community-based rehabilitation training after stroke: results of a pilot randomised controlled trial (ReTrain) investigating acceptability and feasibility February 2018BMJ Open 8(2):e018409 Norris et al. (2018) Acceptability and experience of a functional training programme (ReTrain) in community-dwelling stroke survivors in South West England: A qualitative study July 2018BMJ Open 8(7):e02217	Thanks to the reviewer for providing us with the latest references. We have incorporated these references in the revised manuscript.	References added pages 7 and in reference list (references 8 and 9).

15. Some sections are too brief, you need to provide more information and more detail to help the reader, for example the outcomes and the costs.	In consideration of the word limit, we had to place some of the contents in the online supplementary document. However, as per the reviewer's comment, some of the contents have been moved back to the main document.	See additions to the main text on pages 9 to 13.
16. About the date of currency conversion, it is not reported, there is only the year, but in a year you need a specific average value.	The specific month for the currency conversion rate is provided in the revised manuscript (the average from that month was used in the estimation).	Page 11 of the revised manuscript. All costs are expressed in Australian dollars (AUD) for the 2015 reference year value and can be converted to United States dollar (USD) using the Purchasing Power Parity rate 1 USD=1.463 AUD[21] (December 2015).
17. In the last part of outcomes , you report that you had inherent difficulties. Can you explain in more detail?	At baseline, most of patients were unable to complete a quality of life instrument due to their health emergency. This is added in the revised manuscript.	Page 9 of the revised manuscript. Due to the inherent difficulties of administering the AQoL instrument to acute stroke patients (i.e. most of patients were not able to respond to these questions at baseline), the mRS score at baseline[10] was used as a surrogate measure of patient utility during the acute phase.
18. When you write that there were 56 stroke units across five geographical jurisdiction, can you say the distribution of the units?	There were 24 sites in Australia, 29 sites in UK, 1 site in NZ, 1 site in Malaysia, and 1 site in Singapore. The number of sites in each country has been added to the revised manuscript.	Page 13 of the revised manuscript (the highlighted are newly added). Between July 2006 and October 2014, 2,104 patients (VEM 1,054; UC 1,050) were recruited across 58 sites from Australia (N=1,054, 24 sites), New Zealand (N=189, 1 site), United Kingdom (N=610, 29 sites), Singapore (N=128, 1 site) and Malaysia (N=123, 1 site).

VERSION 2 – REVIEW

REVIEWER	Jim Burke University of Michigan, USA
REVIEW RETURNED	12-Feb-2019

GENERAL COMMENTS	I think that this revision very fairly addresses my initial critique and my most substantive concern — that these data may be misinterprted as a “positive” cost effectiveness analysis has been completely eliminated. This is a well-written paper and reports the results of a considerable amount of primary data collection as well as CEA analysis. My only residual concern is more editorial than scientific. The motivation to execute and report a pre-planned analysis is obviously sound. However, in terms of constructing a manuscript its less obvious that the pre-planned analysis should be the center of the paper. This approach buries the lede. Based on the main trial results, the outcome of the CEA can be predicted with great confidence. As a potential reader of this paper, I may have glanced at the abstract and never looked a the paper in detail. Yet, this study isn’t merely a CEA based on existing data and estimates, but rather it is both a CEA and a detailed study on primary cost collection. While the CEA results are obvious, the primary cost collection data are less obvious and are extremely valuable to researchers performing subsequent CEAs. Yet, I suspect I never would have found these data had I not been asked to review this paper! I would have looked at the abstract, concluded that the conclusion was obvious and never read the paper. Thus, I’d suggest considering a reframing of this paper (perhaps just the abstract — the added sentence probably wouldn’t have been enough to catch my eye) that highlights both the CEA results as well (or even secondarily to) as the primary cost data collection.

REVIEWER	Paolo Landa University of Exeter Medical School United Kingdom
REVIEW RETURNED	10-Feb-2019

GENERAL COMMENTS	The overall work improved and the review process filled the empty or weak parts of the paper. The paper is ready to be published

VERSION 2 – AUTHOR RESPONSE

Response reviewer’s comments

Reviewer: 1

Comment

Please leave your comments for the authors below I think that this revision very fairly addresses my initial critique and my most substantive concern — that these data may be misinterprted as a “positive” cost effectiveness analysis has been completely eliminated. This is a well-written paper and reports the results of a considerable amount of primary data collection as well as CEA analysis.

My only residual concern is more editorial than scientific. The motivation to execute and report a pre-planned analysis is obviously sound. However, in terms of constructing a manuscript its less obvious that the pre-planned analysis should be the center of the paper. This approach buries the lede. Based on the main trial results, the outcome of the CEA can be predicted with great confidence. As a potential reader of this paper, I may have glanced at the abstract and never looked a the paper in detail. Yet, this study isn’t merely a CEA based on existing data and estimates, but rather it is both a

CEA and a detailed study on primary cost collection. While the CEA results are obvious, the primary cost collection data are less obvious and are extremely valuable to researchers performing subsequent CEAs. Yet, I suspect I never would have found these data had I not been asked to review this paper! I would have looked at the abstract, concluded that the conclusion was obvious and never read the paper. Thus, I'd suggest considering a reframing of this paper (perhaps just the abstract — the added sentence probably wouldn't have been enough to catch my eye) that highlights both the CEA results as well (or even secondarily to) as the primary cost data collection.

Response

Thank you for the comment. We have revised the abstract as per your suggestion (please refer to the highlighted sentences in the abstract).

Revision

Abstract

Objectives While Very Early Mobilisation (VEM) intervention for stroke patients was shown not to be effective at 3 months, 12 -month clinical and economic outcomes remain unknown. The aim was to assess cost-effectiveness of a VEM intervention within a Phase III randomised controlled trial (RCT).

Design An economic evaluation alongside a RCT, and detailed resource use and cost analysis over 12-months post-acute stroke.

Setting Multi-country RCT involved 58 stroke centres.

Participants 2104 patients with acute stroke who were admitted to a stroke unit.

Intervention A very early rehabilitation intervention within 24 hours of stroke onset

Methods Cost-utility analyses were undertaken according to pre-specified protocol measuring VEM against usual care (UC) based on 12 -month outcomes. The analysis was conducted using both health sector and societal perspectives. Unit costs were sourced from participating countries. Details on resource use (both health and non-health) were sourced from Cost Case Report Form. Dichotomised Modified Rankin Scale (mRS) scores (0-2 vs 3-6) and Quality Adjusted Life Years (QALYs) were used to compare the treatment effect of VEM and UC. The base case analysis was performed on an Intention-To-Treat (ITT) basis and 95% confidence intervals (CI) for cost and QALYs were estimated by bootstrapping. Sensitivity analysis were conducted to examine the robustness of base case results.

Results VEM and UC groups were comparable in the quantity of resource use and cost of each component. There were no differences in the probability of achieving a favourable mRS outcome (0.030, 95%CI: -0.022 to 0.082), QALYs (0.013, 95%CI: -0.041 to 0.016) and cost (AUD1082, 95%CI: -$2520 to $4685) from a health sector perspective; or AUD$102, 95%CI: -$6907 to $7111, from a societal perspective including productivity cost). Sensitivity analysis achieved results with mostly overlapped CIs.

Conclusions VEM and UC were associated with comparable costs, mRS outcome and QALY gains at 12 months. Compared with to UC, VEM is unlikely to be cost-effective. The long-term data collection during the trial also informed resource use and cost of care post-acute stroke across five participating countries.

VERSION 3 - REVIEW

REVIEWER	Jim Burke University of Michigan
REVIEW RETURNED	13-Mar-2019

GENERAL COMMENTS	The authors have reasonably addressed my small residual concern.