

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)	The involvement of physician assistants in inpatient care in hospitals in the Netherlands: a cost-effectiveness analysis
AUTHORS	Timmermans, Marijke; van den Brink, Geert; van Vught, Anneke; Adang, Eddy; van Berlo, Charles; van Boxtel, Kim; Braunius, Weibel; Janssen, Loes; Venema, Alyssa; van den Wildenberg, Frits; Wensing, Michel; Laurant, Miranda

VERSION 1 - REVIEW

REVIEWER	Roderick S. Hooker Retired American
REVIEW RETURNED	20-Feb-2017

GENERAL COMMENTS	<p>This is the second part of a Dutch PA employment in hospital services study. The first was a qualitative examination and this is an economic evaluation. I write my review from an American perspective, but I am aware this is a British journal, read widely by Europeans and the authors Dutch. While it is a global audience who reads this growing PA literature, the American audience may be the larger set of readers who may take the most interest. I apologize if my views appear Americocentric.</p> <p>Quality Adjusted Life Years (QALY) is an economic evaluation for assessing medical interventions and expressed in cost terms. Less than a dozen studies have been done on interventions that include PAs or NPs and the work by Timmermans and colleagues is the latest. To this reviewer this is the essence of this work – quantifying the intervention of PAs in roles once solely occupied by physicians. I suggest this utility analysis asset and assessment is lost in the manuscript and needs to be emphasized more (especially in the Introduction and the abstract). The reason is that economists know what QALY is all about but the average reader (policy maker, student, manager, etc.) may not. For PA and NP behavioral scholars this work will be a model for more refined work on the growing subject of physician substitution and task transfer. A short paragraph as to what it is and why it is being used will inform a wider range of readers. It is a utility theory being applied here and that theory should be stated as such. Incidentally, one suggestion is to change the title to a “Cost-Utility Analysis” which would make this unique study stand out more and more likely to be read.</p> <p>The other asset to this piece, from an international perspective, is the EQ-SD-3L. Europeans may know this instrument but North Americans may not - many were taught the SF 36 (RAND). The EQ-SD is European in use (although an American version exists). My argument for being more descriptive about this instrument is the</p>
-------------------------	--

same as above – North American sociologists know these instruments but the readers I mentioned may not know them (as such) and find them arcane. It is a testimony to the Dutch they are tuned up to these instruments and use such validated survey questionnaires but the >100,000 North American PAs (who should read this literature) will not be as informed. A brief explanation is needed in the Introduction and abstract.

Another comment for general readers is the type of hospital OECD countries use (versus American ones). These European settings are often large medical centers and may include both acute and chronic care patients under one roof. That nuance gets lost on American readers (sorry) since they are primarily used to the 4500 acute care hospitals (approaching 1 million acute care beds) that are big and small, urban and rural. The growing acute care hospital beds in the US are being staffed more and more by PAs. Few PAs are involved with chronic care hospitals (although NPs may fill this role where there are shortages of physicians). Providing a sentence or two about Dutch hospitals will inform the readers more about the role of PAs.

Medical Specialty (Table 1). The medical specialties in this study are a range of surgical and internal medicine subspecialties (not general medicine). In many American hospitals it is the “hospitalists” (albeit “generalists”) who assume the care for most inpatients and intensive care specialists for intensive care patients. This is mention because this study is one of a range of specialties and not general medicine. There is no criticism but just a nudge to mention these specialties as a sentence or two in the manuscript. See:

- Hefter, Y., Madahar, P., Eisen, L. A., & Gong, M. N. (2016). A time-motion study of ICU workflow and the impact of strain. *Critical care medicine*, 44(8), 1482-1489;
- Beresford JV, Hooker RS. The physician assistant hospitalist: A time-motion study. *Journal of Hospital Administration*. 2015; 4(5): 61-66;
- Kartha, A., Restuccia, J. D., Burgess, J. F., Benzer, J., Glasgow, J., Hockenberry, J., ... & Kaboli, P. J. (2014). Nurse practitioner and physician assistant scope of practice in 118 acute care hospitals. *Journal of Hospital Medicine*, 9(10), 615-620;
- Cawley, J. F., & Hooker, R. S. (2006). The Effects of Resident Work Hour Restrictions on Physician Assistant Hospital Utilization. *The Journal of Physician Assistant Education*, 17(3), 41-43; others.

Charlson Index: The authors use the Charlson comorbidity scoring system. That choice of comorbidity scores is fine and one of 6 health status indices. However, in my review of papers that use comorbidity indices the Elixhauser Comorbidity Index seems to be replacing the Charlson index. This newer index can be used to measure the comorbidity of individual episodes of hospital care or it can be used to create a longitudinal index based on multiple hospital episodes and physician visits over time, if so desired, depending on the purpose of the specific research project. It uses a weighted algorithm based on the association between comorbidity and death, in order to produce an overall score for the Elixhauser Index. van Walraven et al developed a composite score (VW score) for the Elixhauser comorbidities by modeling in-hospital mortality with inpatient admission data (Ottawa Hospital, Ontario, Canada, 1996–2008). (van Walraven et al., 2009).

Measure of central tendency: the data is presented in means but I would like to know if this is the best measure of probability distribution if the distribution is not symmetrical. Then the question arises if there are cohort differences in outcomes. Just a sentence as to why the 'mean' was chosen over 'median' please.

Imputation: Was the imputation technique embedded in the statistical package? Please state if this was so or what technique was used. Just a few words within a sentence.

Personnel costs (Table S1): I think the term "personnel costs" used is fine. However, perhaps a brief clarification that this is the labour cost in terms of wages and does not include benefits and bonuses. Or does it? The importance of this topic is because the Dutch may have one of the shorter workweeks among OECD countries and this observation should catch the eye of PA/NPs and physicians in other countries. The wage ratio is roughly 1:2.9 – an important difference between MD and PA.

Non-elective readmission is a significant cost and a concern to hospital systems globally (and affects insurance reimbursement rates in the US). That this was not found statistically significant across the various models in this study is important to note. It is a topic, I suggest, that needs emphasis – as an outcome of care. Coupled with the reduced labour cost of a PA (regardless of country) and the growing interest in this topic policy makers should be aware of this finding.

Hospital wards versus hospitals: are you able to state how many hospitals were involved? I read the list of hospitals at the end but would like to see this raw number inserted into the text. Some information about their bed size of the hospital would be helpful for comparison purposes (if known).

Table 2m Utilities at admission ... and QALY gained: Could this be better visualized as a graph?

Some of the references used may be supplanted with possibly better ones. I suggest the following three be replaced:

1. Mittman et al (2002). OUTDATED. Suggest: Cawley JF, Hooker RS. Physician assistants in American medicine: the half century mark. *American Journal of Managed Care*. 2013; 19(10); e333-e341.

2. Iannuzzi (2015). The Iannuzzi study has been DISCOUNTED and I suggest does not serve this study well. Please see for explanation as to why: Cawley JF, Hooker RS. Letter to the Editor re Iannuzzi et al Comparing hospitalist-resident to hospitalist midlevel practitioner team performance on length of stay and direct patient care cost. *Journal of Graduate Medical Education*. 2015; 7(6): 689. Suggest Kartha, A., Restuccia, J. D., Burgess, J. F., Benzer, J., Glasgow, J., Hockenberry, J., ... & Kaboli, P. J. (2014). Nurse practitioner and physician assistant scope of practice in 118 acute care hospitals. *Journal of Hospital Medicine*, 9(10), 615-620.

3. Miller (1998). OUTDATED. Suggest: Oswanski, M. F., Sharma, O. P., & Raj, S. S. (2004). Comparative review of use of physician assistants in a level I trauma center. *The American Surgeon*, 70(3), 272.

The rest of the references are fine. Additional references that may support (or inform) this Timmerman et al study are listed below (but not intended that they be used):

James Van Rhee, P. A. C., Ritchie, J., & Eward, A. M. (2002). Resource use by physician assistant services versus teaching services. *Journal of the American Academy of Physician Assistants*, 15(1), 33-42.

Victorino, G. P., & Organ Jr, C. H. (2003). Physician assistant influence on surgery residents. *Archives of Surgery*, 138(9), 971-976.

Nyberg, S. M., Keuter, K. R., Berg, G. M., Helton, A. M., & Johnston, A. D. (2010). A national survey: Acceptance of physician assistants and nurse practitioners in trauma centers. *Journal of the American Academy of Physician Assistants*, 23(1), 35-41.

Althausen, P. L., Shannon, S., Owens, B., Coll, D., Cvitash, M., Lu, M., ... & Bray, T. J. (2016). Impact of hospital-employed physician assistants on a level II community-based orthopaedic trauma system. *Journal of Orthopaedic Trauma*, 30, S40-S44.

Kulkarni, N., & Cardin, T. (2014). Hospital medicine workforce: The impact of nurse practitioner and physician assistant providers. *Journal of Hospital Medicine*, 9(10), 678-679.

Dies, N., Rashid, S., Shandling, M., Swallow, C., Easson, A. M., & Kennedy, E. (2016). Physician assistants reduce resident workload and improve care in an academic surgical setting. *Journal of the American Academy of Physician Assistants*, 29(2), 41-46.

Carpenter, D., Bush, B., Watson, H., & Coopersmith, C. (2014). 139: Physician Assistant Demographics in the Intensive Care Unit. *Critical Care Medicine*, 42(12), A1393-A1394.

Carpenter, D., Bush, B., Watson, H., & Coopersmith, C. (2014). 134: Physician Assistant Billing And Procedures in the Intensive Care Unit. *Critical Care Medicine*, 42(12), A1392.

Costa, D. K., Wallace, D. J., Barnato, A. E., & Kahn, J. M. (2014). Nurse practitioner/physician assistant staffing and critical care mortality. *CHEST Journal*, 146(6), 1566-1573.

Jones, C. D., Cumbler, E., Honigman, B., Burke, R. E., Boxer, R. S., Levy, C., ... & Wald, H. L. (2017). Hospital to Post-Acute Care Facility Transfers: Identifying Targets for Information Exchange Quality Improvement. *Journal of the American Medical Directors Association*, 18(1), 70-73.

Gillard, J. N., Szoke, A., Hoff, W. S., Wainwright, G. A., Stehly, C. D., & Toedter, L. J. (2011). Utilization of PAs and NPs at a level I trauma center: effects on outcomes. *Journal of the American Academy of Physician Assistants*, 24(7), 34-43.

REVIEWER	Zachary Hartsell Wake Forest baptist Medical Center, USA
REVIEW RETURNED	06-Mar-2017

GENERAL COMMENTS	<p>Page 3 line 44. Isn't one of the objective to also determine differences in quality of care using QALY?</p> <p>Page 5 Line 93. This sentence needs to be cited.</p> <p>Page 5 line 95. In the US PAs often change careers at least three times.</p> <p>Page 5 line 96-98. Are the US studies applicable to the Netherlands model? Important to note that all previous studies of this have been in the US experience.</p> <p>Page 6 line 115, I suggest physicians are "involved with" not in charge of.</p> <p>Page 6 line 116. What level of residents?</p> <p>Page 6 line 119, specialist physicians, you mean attending physicians (not trainees).</p> <p>Page 7 line 32, what qualified a patient as terminally ill?</p> <p>Page 7 line 33, In the US Daycare is called "observation status"</p> <p>Page 8 line 152. Why one month after discharge? Is this the standard time allotment used for this type of research?</p> <p>Page 8 line 162, is 10% standard for a random sample in this kind of work?</p> <p>Page 8 line 168, are the supervision costs the same for residents and PAs?</p> <p>Page 11, line 248. Are patient surveys accurate way to know what they used? are patient reliable about reporting services?</p> <p>Page 14, line 320, in the US PAs make significantly more than residents.</p> <p>Otherwise well written. Will add meaningfully to the PA literature.</p>
-------------------------	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Roderick S. Hooker

Institution and Country: Retired, American Competing Interests: None declared.

This is the second part of a Dutch PA employment in hospital services study. The first was a qualitative examination and this is an economic evaluation. I write my review from an American perspective, but I am aware this is a British journal, read widely by Europeans and the authors Dutch. While it is a global audience who reads this growing PA literature, the American audience may be the larger set of readers who may take the most interest. I apologize if my views appear Americocentric.

Quality Adjusted Life Years (QALY) is an economic evaluation for assessing medical interventions and expressed in cost terms. Less than a dozen studies have been done on interventions that include PAs or NPs and the work by Timmermans and colleagues is the latest. To this reviewer this is the essence of this work – quantifying the intervention of PAs in roles once solely occupied by physicians. I suggest this utility analysis asset and assessment is lost in the manuscript and needs to be emphasized more (especially in the Introduction and the abstract). The reason is that economists know what QALY is all about but the average reader (policy maker, student, manager, etc.) may not. For PA and NP behavioral scholars this work will be a model for more refined work on the growing subject of physician substitution and task transfer. A short paragraph as to what it is and why it is being used will inform a wider range of readers. It is a utility theory being applied here and that theory should be stated as such. Incidentally, one suggestion is to change the title to a “Cost-Utility Analysis”

which would make this unique study stand out more and more likely to be read.

Response: Thank you very much for your advice. Based on your comment, we added a short explanation on the QALY to the introduction section: 'Costs concerned all direct healthcare costs from day of admission until one month after discharge. Health outcome concerned quality-adjusted life years (QALYs), which is a composite measure of effectiveness consisting of quality of life and life years gained.' (line 107-110)

The other asset to this piece, from an international perspective, is the EQ-SD-3L. Europeans may know this instrument but North Americans may not - many were taught the SF 36 (RAND). The EQ-SD is European in use (although an American version exists). My argument for being more descriptive about this instrument is the same as above – North American sociologists know these instruments but the readers I mentioned may not know them (as such) and find them arcane. It is a testimony to the Dutch they are tuned up to these instruments and use such validated survey questionnaires but the >100,000 North American PAs (who should read this literature) will not be as informed. A brief explanation is needed in the Introduction and abstract.

Response: Thank you for this comment. We agree that it is important to give a brief explanation on the EQ-SD-3L. However, this is already stated in the method section, line 156-160: 'QALYs were derived using the EuroQoL-5D questionnaire (EQ-5D-3L)¹⁵, which is a widely used validated patient questionnaire comprising five domains: mobility, self-care, usual activities, pain, and anxiety/depression. Each domain has three possible levels indicating; no problems, moderate problems or severe problems. The EQ-5D-3L was assessed at three time points: at admission, discharge and one month after discharge.' In our opinion, this should be written in the method section and not in the introduction, as this concerns information on how the QALY is measured. Based on your comment, we added the fragment 'which was measured with the EuroQoL-5D questionnaire' to the abstract, as it is indeed important to state which questionnaire was used to measure the QALYs.

Another comment for general readers is the type of hospital OECD countries use (versus American ones). These European settings are often large medical centers and may include both acute and chronic care patients under one roof. That nuance gets lost on American readers (sorry) since they are primarily used to the 4500 acute care hospitals (approaching 1 million acute care beds) that are big and small, urban and rural. The growing acute care hospital beds in the US are being staffed more and more by PAs. Few PAs are involved with chronic care hospitals (although NPs may fill this role where there are shortages of physicians). Providing a sentence or two about Dutch hospitals will inform the readers more about the role of PAs.

Response: Thank you for this suggestion. We have now added a short explanation to the introduction section in line 101-104: 'Besides, all studies were conducted in the United States, where most hospitals involving PAs concern only acute care. In the Netherlands, most hospitals include both acute and chronic care under one roof'.

Medical Specialty (Table 1). The medical specialties in this study are a range of surgical and internal medicine subspecialties (not general medicine). In many American hospitals it is the "hospitalists" (albeit "generalists") who assume the care for most inpatients and intensive care specialists for intensive care patients. This is mentioned because this study is one of a range of specialties and not general medicine. There is no criticism but just a nudge to mention these specialties as a sentence or two in the manuscript. See:

- Hefter, Y., Madahar, P., Eisen, L. A., & Gong, M. N. (2016). A time-motion study of ICU workflow and the impact of strain. *Critical care medicine*, 44(8), 1482-1489;
- Beresford JV, Hooker RS. The physician assistant hospitalist: A time-motion study. *Journal of*

Hospital Administration. 2015; 4(5): 61-66;

- Kartha, A., Restuccia, J. D., Burgess, J. F., Benzer, J., Glasgow, J., Hockenberry, J., ... & Kaboli, P. J. (2014). Nurse practitioner and physician assistant scope of practice in 118 acute care hospitals. *Journal of Hospital Medicine*, 9(10), 615-620;
- Cawley, J. F., & Hooker, R. S. (2006). The Effects of Resident Work Hour Restrictions on Physician Assistant Hospital Utilization. *The Journal of Physician Assistant Education*, 17(3), 41-43; others.

Response: Based upon your comment, we have clarified this briefly in line 136-138: 'Control wards were matched with the intervention wards on the basis of hospital type (i.e. academic or non-academic) and medical specialty (i.e. a range of surgical and medical specialties). No wards with general medicine were involved.'

Charlson Index: The authors use the Charlson comorbidity scoring system. That choice of comorbidity scores is fine and one of 6 health status indices. However, in my review of papers that use comorbidity indices the Elixhauser Comorbidity Index seems to be replacing the Charlson index. This newer index can be used to measure the comorbidity of individual episodes of hospital care or it can be used to create a longitudinal index based on multiple hospital episodes and physician visits over time, if so desired, depending on the purpose of the specific research project. It uses a weighted algorithm based on the association between comorbidity and death, in order to produce an overall score for the Elixhauser Index. van Walraven et al developed a composite score (VW score) for the Elixhauser comorbidities by modeling in-hospital mortality with inpatient admission data (Ottawa Hospital, Ontario, Canada, 1996–2008). (van Walraven et al., 2009).

Response: Thank you very much for this information. It is not possible for us to change the comorbidity scoring, as this was coded while reviewing the patient records. For future studies, we will consider using the Elixhauser Comorbidity Index instead of the Charlson index.

Measure of central tendency: the data is presented in means but I would like to know if this is the best measure of probability distribution if the distribution is not symmetrical. Then the question arises if there are cohort differences in outcomes. Just a sentence as to why the 'mean' was chosen over 'median' please.

Response: The distribution of costs is indeed skewed. However, especially with regard to the outcome 'cost' it is common to report means instead of medians to enhance interpretability, as median cost are considered less informative. This is extensively discussed with experts in the field of economic evaluation, of which one is a coauthor of this study (Eddy Adang).

Imputation: Was the imputation technique embedded in the statistical package? Please state if this was so or what technique was used. Just a few words within a sentence.

Response: The imputation technique (multiple imputation) was indeed embedded in the statistical package. We have added this to line 224-225.

Personnel costs (Table S1): I think the term "personnel costs" used is fine. However, perhaps a brief clarification that this is the labour cost in terms of wages and does not include benefits and bonuses. Or does it? The importance of this topic is because the Dutch may have one of the shorter workweeks among OECD countries and this observation should catch the eye of PA/NPs and physicians in other countries. The wage ratio is roughly 1:2.9 – an important difference between MD and PA.

Response: Bonuses are indeed not included. We have now clarified this in supplemental table 1.

Non-elective readmission is a significant cost and a concern to hospital systems globally (and affects

insurance reimbursement rates in the US). That this was not found statistically significant across the various models in this study is important to note. It is a topic, I suggest, that needs emphasis – as an outcome of care. Coupled with the reduced labour cost of a PA (regardless of country) and the growing interest in this topic policy makers should be aware of this finding.

Hospital wards versus hospitals: are you able to state how many hospitals were involved? I read the list of hospitals at the end but would like to see this raw number inserted into the text. Some information about their bed size of the hospital would be helpful for comparison purposes (if known).

Response: In total, 34 wards were included across 23 different hospitals. We have added this to line 244-246: 'Between April 2013 and May 2015 we included 1,021 patients spread over 17 hospital wards in the intervention group, and 1,286 patients spread over 17 hospital wards in the control group. In total, 23 hospitals across the Netherlands were involved.'

Table 2m Utilities at admission ... and QALY gained: Could this be better visualized as a graph?

Response: As the differences between the study arms are very small, it is difficult to make the differences visible in a graph. Besides, we believe that the presentation of the results in a table is more informative as numbers can be clearly presented.

Some of the references used may be supplanted with possibly better ones. I suggest the following three be replaced:

1. Mittman et al (2002). OUTDATED. Suggest: Cawley JF, Hooker RS. Physician assistants in American medicine: the half century mark. *American Journal of Managed Care*. 2013; 19(10); e333-e341.

2. Iannuzzi (2015). The Iannuzzi study has been DISCOUNTED and I suggest does not serve this study well. Please see for explanation as to why: Cawley JF, Hooker RS. Letter to the Editor re Iannuzzi et al Comparing hospitalist-resident to hospitalist midlevel practitioner team performance on length of stay and direct patient care cost. *Journal of Graduate Medical Education*. 2015; 7(6): 689. Suggest Kartha, A., Restuccia, J. D., Burgess, J. F., Benzer, J., Glasgow, J., Hockenberry, J., ... & Kaboli, P. J. (2014). Nurse practitioner and physician assistant scope of practice in 118 acute care hospitals. *Journal of Hospital Medicine*, 9(10), 615-620.

3. Miller (1998). OUTDATED. Suggest: Oswanski, M. F., Sharma, O. P., & Raj, S. S. (2004). Comparative review of use of physician assistants in a level I trauma center. *The American Surgeon*, 70(3), 272.

The rest of the references are fine. Additional references that may support (or inform) this Timmerman et al study are listed below (but not intended that they be used):

James Van Rhee, P. A. C., Ritchie, J., & Eward, A. M. (2002). Resource use by physician assistant services versus teaching services. *Journal of the American Academy of Physician Assistants*, 15(1), 33-42.

Victorino, G. P., & Organ Jr, C. H. (2003). Physician assistant influence on surgery residents. *Archives of Surgery*, 138(9), 971-976.

Nyberg, S. M., Keuter, K. R., Berg, G. M., Helton, A. M., & Johnston, A. D. (2010). A national survey: Acceptance of physician assistants and nurse practitioners in trauma centers. *Journal of the American*

Academy of Physician Assistants, 23(1), 35-41.

Althausen, P. L., Shannon, S., Owens, B., Coll, D., Cvitash, M., Lu, M., ... & Bray, T. J. (2016). Impact of hospital-employed physician assistants on a level II community-based orthopaedic trauma system. *Journal of Orthopaedic Trauma*, 30, S40-S44.

Kulkarni, N., & Cardin, T. (2014). Hospital medicine workforce: The impact of nurse practitioner and physician assistant providers. *Journal of Hospital Medicine*, 9(10), 678-679.

Dies, N., Rashid, S., Shandling, M., Swallow, C., Easson, A. M., & Kennedy, E. (2016). Physician assistants reduce resident workload and improve care in an academic surgical setting. *Journal of the American Academy of Physician Assistants*, 29(2), 41-46.

Carpenter, D., Bush, B., Watson, H., & Coopersmith, C. (2014). 139: Physician Assistant Demographics in the Intensive Care Unit. *Critical Care Medicine*, 42(12), A1393-A1394.

Carpenter, D., Bush, B., Watson, H., & Coopersmith, C. (2014). 134: Physician Assistant Billing And Procedures in the Intensive Care Unit. *Critical Care Medicine*, 42(12), A1392.

Costa, D. K., Wallace, D. J., Barnato, A. E., & Kahn, J. M. (2014). Nurse practitioner/physician assistant staffing and critical care mortality. *CHEST Journal*, 146(6), 1566-1573.

Jones, C. D., Cumbler, E., Honigman, B., Burke, R. E., Boxer, R. S., Levy, C., ... & Wald, H. L. (2017). Hospital to Post-Acute Care Facility Transfers: Identifying Targets for Information Exchange Quality Improvement. *Journal of the American Medical Directors Association*, 18(1), 70-73.

Gillard, J. N., Szoke, A., Hoff, W. S., Wainwright, G. A., Stehly, C. D., & Toedter, L. J. (2011). Utilization of PAs and NPs at a level I trauma center: effects on outcomes. *Journal of the American Academy of Physician Assistants*, 24(7), 34-43.

Response: Thank you very much for your comments about our references and suggestions for additional references. We've deleted the outdated references and replaced them by the references you suggested.

Reviewer: 2

Reviewer Name: Zachary Hartsell

Institution and Country: Wake Forest baptist Medical Center, USA Competing Interests: None Declared

Page 3 line 44. Isn't one of the objective to also determine differences in quality of care using QALY?

Response: The objective of the described study was to determine differences in costs as well as QALY. But the latter is part of the 'effectiveness' in the formulated aim: 'Objective. To investigate the cost-effectiveness of substitution of inpatient care from medical doctors (MDs) to physician assistants (PAs).' We as well measured quality of care, but those results are described in another paper.

Page 5 Line 93. This sentence needs to be cited.

Response: We agree with your comment, and added a reference to this sentence.

Page 5 line 95. In the US PAs often change careers at least three times.

Page 5 line 96-98. Are the US studies applicable to the Netherlands model? Important to note that all previous studies of this have been in the US experience.

Response: Thank you for this suggestion. In line 309-312 we had already briefly discussed the generalizability of the US studies: 'These studies can however hardly be compared with our study, because different methods to estimate costs were used and the settings were different. In addition, most of these studies compared a hospitalist/PA model with the traditional resident-based model, while hospitalists were not part of the models we used'. Based upon your comment and the comment of the other reviewer, we added the following sentences to line 101-104: 'Besides, all studies were conducted in the United States, where most hospitals involving PAs concern only acute care. In the Netherlands, most hospitals include both acute and chronic care under one roof.'

Page 6 line 115, I suggest physicians are "involved with" not in charge of.

Response: We have rephrased the sentence (line 120): 'In the MD model, only MDs provide medical care for admitted patients at a specific hospital department.'

Page 6 line 116. What level of residents?

Response: This can concern both junior and senior residents. We clarified this in line 121: 'Most of them are junior or senior residents.'

Page 6 line 119, specialist physicians, you mean attending physicians (not trainees).

Response: Based upon your comments we changed medical specialist into attending physician (line 124): 'The residents are supervised by attending physicians'.

Page 7 line 33, In the US Daycare is called "observation status"

Response: Thank you for this information. We've added this to line 151-152: 'Daycare was defined as hospital admissions that were intended to last 24 hours or less (observation status).'

Page 8 line 152. Why one month after discharge? Is this the standard time allotment used for this type of research?

We believe it is important to include a fixed period after discharge in the timeframe, as relatively low use of resources during admission and a relatively short length of stay does not per se reflect adequate medical care. Patients can for example be readmitted because they were discharged too soon. We chose for 1 month after discharge, as events happened after that period are less likely to be related to the initial admission period (Halfon et al, 2002). We have added this to line 191-193: 'We chose for 1 month after discharge, as events happened after that period are less likely to be related to the initial admission period. 18 '

Page 8 line 162, is 10% standard for a random sample in this kind of work?

Response: As far as we know there is no general rule on how many records have to be reassessed, but we believe that the reassessment of 10% of the patients records (in total 231 records) is adequate to control validity of the data. In case of an inter-rater agreement of less than 95%, the records of the total sample were reassessed. It turned out that this was necessary only at one hospital ward, for a single outcome measure.

Page 8 line 168, are the supervision costs the same for residents and PAs?

Response: Supervision is provided by attending physicians. Costs per hour of supervision by the attending physician are independently of the involved PA or residents.

Page 11, line 248. Are patient surveys accurate way to know what they used? are patient reliable about reporting services?

We chose to use patients' questionnaires instead of medical patient records, as after a pilot study the patient records turned out to be incomplete regarding the concerning items (readmission, presentation at emergency department, general practitioner and home care). For example only readmissions were reported which occurred at the same hospital wards, but readmissions on other wards or in other hospitals were not registered. Although we realize that the use of patients questionnaires has disadvantages as well, we believe that in this particular case this is the best choice.

Page 14, line 320, in the US PAs make significantly more than residents.
In contrast to the US, in the Netherlands the salaries of residents and Pas are comparable. To clarify this, we added changed the concerning sentence into: 'Since in the Netherlands the salary of PAs is comparable to the salary of residents (table S1),...'(line 33-334)

Otherwise well written. Will add meaningfully to the PA literature.
Thanks for you kind comment. We have further improved the paper based on reviewers comments.

VERSION 2 – REVIEW

REVIEWER	Roderick S. Hooker Institution: Retired Country: USA
REVIEW RETURNED	24-Apr-2017

GENERAL COMMENTS	The authors have addressed all of my comments. The revised manuscript has been adjusted to my satisfaction. This study as a publication will make a meaningful contribution to the growing economic literature on physician substitution and task transfer. The authors are commended for their fine work. I would welcome an invitation to write an accompanying editorial.
-------------------------	---