

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)	Financial Performance of English NHS Trusts and Variation in Clinical Outcomes: A Longitudinal Observational Study
AUTHORS	Nagendran, Myura; Kiew, Grace; Raine, Rosalind; Atun, Rifat; Maruthappu, Mahiben

VERSION 1 – REVIEW

REVIEWER	Mark Dusheiko Institute of Social and Preventive Medicine. University Hospital of Lausanne and University of Lausanne. And Faculty of Business, economics and Commerce, University of Lausanne.
REVIEW RETURNED	05-Apr-2018

GENERAL COMMENTS	<p>Summary of paper: The paper 'Financial Performance of English NHS Trusts and Variation in Clinical Outcomes: A Longitudinal Observational Study' investigates the trends over time in hospital financial efficiency - operating margins and its association with indicators of care quality at acute NHS Trust level over recent years. The motivation of the study was to present an exploratory analysis, primarily looking at associations between quality of care and poor financial performance. The main focus was to see whether worse financial performance was associated with lower measured quality at acute trust level over time in the NHS in England. Data for 6 years covering the period 2011-2016 were analysed. The main quality outcomes assessed to be associated with hospital financial performance included readmission rates, patient inpatient satisfaction, A&E 4 hour waiting time excess waits, delayed discharge days in hospital, and cancer 2 week waiting time and 62 week treatment waiting time targets. Simple descriptive time trends and differences in outcomes by deciles (pooled over all periods, and stratified by years) of financial performance were presented. A linear regression model pooling data across all time periods and controlling for characteristics of the hospitals was estimated.</p> <p>The study presents clear evidence of significant variation in financial performance as measured by operating margins. Around a quarter of trusts experienced negative and large operating margins between 2011-2013. There was however a marked decline in financial performance across all acute NHS trusts from 2013 onwards with nearly 90% of trusts experiencing negative operating margins by 2016. This marked downturn in the financial efficiency of trusts was also accompanied by worsening trends in some of process measures of care quality: mainly A&E waiting time increases, delayed patient discharges and increased cancer assessment and treatment waiting times, although readmission rates and patient elective treatment satisfaction remained stable. Cross-sectional patterns confirmed this with the 10% worst financially performing trusts having significantly worse process indicators of quality compared to the 10% best performing trusts, but over time this gap in performance narrowed (as all trusts financial performance deteriorated). Linear regression analysis of the relationship between the quality indicators and financial operating margins, that conditioned on characteristics of the hospital trusts (namely volume of inpatient and outpatient activity, foundation trust status, teaching status, geographic and socio-economic characteristics of region,</p>
-------------------------	---

expenditures on agency staff and consulting services, as well as bed capacity and occupancy rates) confirmed that A&E waiting time breaches and delayed transfers were significantly associated with worse financial performance.

The authors acknowledge the complexity of understanding the relationship between financial performance and quality of care, and clearly identify the limitations with their study. However conclude that this is an important area of interest. The paper was reasonably well written although the description and presentation of the statistical results was not always very clear and detailed definition of the variables and full regression results would have been appreciated. I found the topic of the research of interest with a high degree of policy relevance. The underlying trends in hospital financial performance in the English NHS, associated characteristics of poor and well performing trusts as well as associated differences of trends in indicators of quality made for an interesting and thought provoking read.

The main weakness of the paper is the fact that the study design, which is primarily descriptive and observational. The theoretical and institutional considerations of how financial performance and quality may be related, and the factors driving the results was not thought through in a systematic way. This was reflected in the selection of indicators, explanation of the observed associations, and description of the financial and institutional operating environments. A lot of the observed associations appeared more or less mechanical in nature due to the reimbursement system and external changes in the health sector beyond the influence of the acute hospital sector. Given the focus of the paper was investigating the effects of financial difficulties on quality, rather than in understanding the mechanisms generating both financial problems and how they would impact or be aligned with quality; then in my view a deeper consideration of these factors is required in order to better interpret the findings and guide the research.

I have some more general comments below followed by some specific comments that are more minor:

General comments:

The authors need to better consider and explain what determines financial performance as measured by operating margins in acute (public) hospital sector in the English NHS, and how this relates to quality.

In particular this has be precisely related to the reimbursement system, regulation policy and health care market circumstances (wider health and social care environment).

For instance, if you consider the measure of financial performance, explain what drives revenues and costs of care and how this would be related to quality. From what I understand the operating margin of the hospital is calculated as $(\text{Revenue} - \text{Costs})/\text{Revenue}$. This can be expressed in more detail as roughly $1 - \frac{\sum_i(\text{FC}_i + \text{VC}_i * \text{Q}_i)}{\sum_i(\text{P}_i * \text{Q}_i + \text{R})}$. Where FC_i is the fixed costs of providing treatments i , VC_i are the variable costs associated with treatment i (can also be considered as the marginal costs of providing each additional treatment Q_i) of providing Q_i treatments, P_i is the reimbursement price of the i _th treatment and Q_i is the amount of the i _th type of treatment. R are additional revenues for instance from research, or from bonus payments related to quality improvement or other revenues earned outside of providing treatments to patients. Summing the costs and revenues over the i treatments gives you total costs and revenues. If total costs (numerator) equals total revenue (denominator) then net operating margins 0. If the denominator exceeds the numerator then operating margins are positive and if the numerator i.e. costs exceeds the denominator i.e. revenue then operating margins are negative.

Looking at the denominator, it is clear that reimbursement prices are crucial in determining financial performance. In the English NHS these are fixed at the average cost across all trusts providing a given treatment $P_i = [(FCi_mean + VCi_mean * Q_mean) / (Q_mean) * MFF]$, adjusted for the market forces factor that impacts local costs i.e. wage differences and other price differences by region. Hence it would appear that hospital trusts that treat higher volumes of patients would benefit from economies of scale and therefore have lower average costs than the mean hospital. For instance their fixed costs would be spread over many more patients. It is also likely that their marginal costs would fall as patient volumes increase (up to a certain point) as higher volumes would improve treatment efficiency. Hence one would expect to observe economies of scale and possibly scope with larger hospitals enjoying better financial performance. We also know from the literature that higher volumes and concentration of services as well as expertise can increase quality of care. The higher volumes and revenues can also justify more investment in new technologies. Hence this would suggest that better financially performing trusts would have better quality (through economies of scale and scope and higher volumes as well as specialisation).

The reimbursement price mechanism would also directly (mechanically) relate operating margins to quality. For instance reimbursement prices for specialised services are significantly higher and have been shown to be well above average costs of treatment (see papers by Street et al. CHE RP 103, 2014 on costs of specialised care). Hence hospitals providing specialised services will likely have higher revenues (above average costs) and are normally provide better quality of care – through specialisation, concentration of expertise and teaching status etc. The reimbursement system also penalises lower quality directly with no reimbursement for readmissions after 30 days. There are also best practice tariffs (BPTs) rewarding more cost effective treatment interventions such as day case surgery (lower risk, more efficient), which were associated with quality improvements (see Sutton et al, Health Economics) and CEQUIN quality improvement payments that reward up to 2.5% in additional performance payments compliance with quality targets. That is the R component in the formula would directly improve profit margins and be associated with better quality almost by definition. So from a revenue perspective would be a strong expectation to see better quality related to better financial performance.

There was also the introduction of a 30% marginal payment for emergency admissions in excess of 2010 levels that would penalise hospitals financially for admitting more emergency patients above an expected level, and no payment for never events. Emergency admissions are also more costly with lower margins and disrupt the ability of hospitals to undertake elective treatments which are more profitable. The authors also did not mention the efficiency savings target initiatives, where effectively all Pbr tariffs were reduced annually (by around 4-5%) since 2011. This would cut all hospital revenues and likely explains a lot of the observed reductions in financial performance. More emergency care, complex patients, volatile demand would reduce profits, and lower volumes. Again reinforcing a positive finance/quality relationship.

On the cost side the expected relationship between costs and quality is less clear. The main concern would be that cost savings, for instance from having lower staff to patient and bed ratios, employing less well qualified staff (cheaper) e.g. lower grade nurses and doctors; having fewer senior staff to junior staff ratios, skill mix substitution with specialist nurses or junior doctors doing the work of more experienced and qualified staff, and lower investment in new and better quality equipment could lower quality and productivity. Hence this could lead to better financial performance but at the expense of quality. However in terms of efficiency and productivity

employing less skilled staff and using lower quality and cheaper technology could reduce efficiency of care e.g. increase length of stays, increase complications and treatment costs, lower productivity and reduce profitability as well as quality. Also lower volumes would lead to higher average and marginal costs and may be associated with lower quality of care.

There is evidence that lower staff ratios due reduce quality. Also cutting down on management and administrative expenses can lead to lower quality of care and worse financial performance in terms of lost productivity, particularly if it leads clinicians to allocate more time to administration and management and the expense of improving quality of care or productivity. Hospital costs will also be a function of patient severity with more complex patients costing more and more likely to experience adverse outcomes, and be associated with worse quality as measured by outcomes.

The study also neglects to mention the influence of patient choice and the use of private sector providers, in particular independent sector treatment centres who are focused on specific elective procedures that are fairly high volume, routine and standardised. They are focused factories providing efficient treatment at good quality with low waiting times, and have attracted increased market share during the period of study. They also tend to select less complex patients, hence lower cost patients.

This has implications for the authors analysis as hospital trusts facing greater competition from ISTCs will lose market share and revenues these providers, have lower revenues and likely be left with treating more complex patients. Hence hospital trusts facing greater competition from ISTCs would likely be in a worse financial position apparently have worse quality indicators.

The introduction of patient choice and patient reported outcomes and information on hospital waiting times, mortality rates and volumes by procedure has also been shown to have lead to significant changes in patient flows, with patients switching to better quality hospitals (see for instance research by Propper et al. Gutdeker, Gravelle, Moschini and Sicliani etc). Hence this would reinforce the relationship between better quality and financial performance (in terms of revenues from attracting more patients, and scale economies). The impact of competition on quality and financial performance should therefore be mentioned and discussed. The positive aspects would be that competition on quality would drive up quality and reward hospitals with greater volumes of patients and revenue. However greater competition would also see some hospitals lose out to more focused providers and face higher costs, lower volumes and revenues. In order to compete would have to improve quality to attract patients. Patient choice and selection – evidence that patients switched for elective treatment to better quality providers, lower waiting times, better outcomes. This would have increased profits at better quality hospitals, more elective care and volumes.

Also Trusts operating in more competitive markets may have had to invest more in quality to attract patients and to be more efficient, but may have had to be operating on lower margins. Hence competitive forces may have been sufficient to sustain quality, but may also have contributed to lower margins.

The study also needs to consider the effect of the imposed nurse staffing ratios on costs and quality. This was mandated and is imposing additional costs on trusts where shortages existed or where attracting regular nursing staff is problematic (high demand for agency staffing etc). It is important to understand that this was imposed, creating considerable additional costs that would lower margins, but may be improving quality (hence leading to a negative relationship between margins and quality as trust efficiency reduced, but quality improves).

The paper does not consider the issues related to foundation trusts. They have stronger incentives to reduce costs as well as increase margins. Likely would also look to have higher revenues by attracting more patients through quality and doing high profit margin procedures. I would consider separating out foundation and non-foundation trusts in the analysis to see if margins are different and relationship between quality and lower margins also differ.

Finally the study needs to mention the other issues in health and social care impacting hospital quality and financial performance, namely primary and community care as well as social care. Namely investment in primary care has stagnated over the period of the study with number of GPs not keeping pace with population needs. Evidence that access and quality of GP care reduces secondary care costs through reduced emergency admissions (see Asaria and Cookson). Hence poor financial performance and lower quality process measures may be observed in acute trusts due to rising emergency admissions, which may have nothing to do with quality of secondary care. Similarly access to community care, tertiary and social care has been squeezed, this will impact discharge delay, readmissions and hence hospital costs and quality (see Gaugan, Siciliani and Gravelle for papers on discharge delay and availability of social care). However it is totally beyond the control of the NHS Trusts.

The discussion may want to elaborate on these points. In particular, that to improve quality and financial performance the trusts may need to integrate care and align resources more closely with primary and tertiary care. Allocating more resources to outpatient services in the community, early diagnosis, primary and secondary prevention as well as discharge management and patient follow up as well as rehabilitation. Current reimbursement and funding structures with payment on for inpatient and hospital based outpatient activity does not encourage this and if the integration and quality of care in the ambulatory sector is not improved hospital financial positions may not get any better. Hence new funding and payment models may be needed to improve financial performance and quality (for instance bundled payments for care across care sectors).

These points above should be considered in the motivation of the paper and specification of the analysis. For instance, in explaining the specification of your longitudinal regression model and selection of control variables. In my opinion, if the paper wants to make a more significant contribution you should identify a set of quality indicators more directly under the control of the hospital (inpatient waiting times, standardised hospital and 30 day mortality rates, hospital acquired infection rates, adverse hospital events, staff satisfaction and recommendations). I would then look to exploit the longitudinal data and timing of external changes in the reimbursement and health care environment that has put trusts under financial pressure, and look to see how quality had responded to these financial pressures. Hence you could condition on baseline quality, observe trusts who experienced significant changes in their financial situation due to the changes in reimbursement reforms or external situations (such as increased A&E consultations, Reduced supply of primary and tertiary care, reductions in reimbursement rates etc. or increased ISTC competition, imposition of minimum nursing and other staff ratios plus the costs of hiring extra agency staff). These would be exogenous external shocks impacting operating margins and then one could look to see over time following this externally imposed financial pressures, how quality has responded.

Specific comments:

The panel data model is very simplistic and full results are not

	<p>provided. Did you estimate a GLS random effects model? Did you consider a log-linear specification, also was it possible to condition on baseline outcomes? May want to look at analysis that stratifies by foundation or non-foundation trust status. Did you include a time trend in the analysis?</p> <p>I would try to look to exploit a before and after comparison of quality in response to specific exogenous shocks that impacted hospital margins exogenously (and possible differentially by hospital).</p> <p>Paper needs to provide more background and detail on reimbursement system, competitive environment and incentives to provide high quality care and efficient care: For instance the reduced payments for readmissions, emergency treatment marginal payment reductions, no payment for never events, PbR efficiency saving tariff reductions. CEQUIN and Best Practice Tariffs. Introduction of minimum nursing staff ratios.</p>
--	--

<p>REVIEWER</p>	<p>Adrián Villaseñor Centre for Health Economics University of York United Kingdom</p> <p>I have attempted to give my unbiased and honest review of this paper, however, I would like to state that: I am currently working on a very similar research article. The outcome variables, interest variables, and sources of data are very closely related. This could have influenced the way in which I frame my comments.</p>
<p>REVIEW RETURNED</p>	<p>10-Apr-2018</p>

<p>GENERAL COMMENTS</p>	<p>Thanks to the authors for addressing this very important issue with clarity and rigour. I would like, however, raise a few points of concern that are likely to improve your paper.</p> <ol style="list-style-type: none"> 1. I think the paper could benefit from a theoretical or conceptual discussion, with appropriate references, of why these concepts (financial performance and their outcomes) might be correlated. Without any explanation of the mechanisms behind the correlation between two variables, any statistical analysis can be seen as spurious. 2. The claim on page 4, lines 46-51, should be backed up with appropriate references. 3. I would suggest the authors state the financial periods clearly, and in full (e.g. 2010-2011) as it can be confusing for the reader (see page 2, line 12 vs page 5, lines 13-17). 4. I think the paper would benefit from further justifying your measure of financial performance. For Foundation trusts I think it is fine to use "Table ID 1, Subcode 100" as a measure of Turnover, but I would ask you to justify why you are using "Table ID1, Subcode 160" as a measure of Surplus/Deficit. Same point applies to Surplus/Deficit measures for Non-Foundation trusts. These measures of surpluses/deficits include financial items that trusts can use to 'adjust' their surplus/deficit. The use of these financial devices might also be related to trusts outcomes, which will make your sample biased. 5. Moreover, since you are not doing the analysis separately by foundation / non foundation trusts, I think there is a discussion missing of whether the two concepts, "SURPLUS/ (DEFICIT) FOR THE YEAR" for foundation trusts, and "Adjusted retained surplus/(deficit)" for Non-Foundation trust are comparable. 6. Your main measure of Financial Performance is "operating margin" as a percentage of turnover. Could you please provide a further explanation of why this might be an appropriate measure that
--------------------------------	---

	<p>takes into account differences in trusts size?</p> <p>7. A point related to point 4 and 6 then would be, is it appropriate to call your measure "operating margin" (page 2, line 6). Please justify further.</p> <p>8. Tables should be able to be understood even in the absence of the article's text. Please consider 1) Formatting your tables in a way that are visually appealing, and 2) add any table footnotes that allow the reader to understand the table.</p> <p>9. Table 4 uses the word "impact", which you can't infer from your statistical analysis. Please avoid any causal language throughout your paper.</p> <p>10. I find Figure 1 uninformative. Please consider doing histograms per time period.</p> <p>11. Please justify further why you did not add year dummies into your regression analysis and add a footnote to Table 4 stating what control variables you used for each regression.</p> <p>12. I do not see how the analysis presented in page 11, lines 1-11 with its accompanied figures, add anything to the analysis. It seems to come out from nowhere and makes the analysis hard to follow.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Mark Dusheiko

Institution and Country: Institute of Social and Preventive Medicine. University Hospital of Lausanne and University of Lausanne. And Faculty of Business, economics and Commerce, University of Lausanne.

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below Summary of paper:

The paper 'Financial Performance of English NHS Trusts and Variation in Clinical Outcomes: A Longitudinal Observational Study' investigates the trends over time in hospital financial efficiency - operating margins and its association with indicators of care quality at acute NHS Trust level over recent years. The motivation of the study was to present an exploratory analysis, primarily looking at associations between quality of care and poor financial performance. The main focus was to see whether worse financial performance was associated with lower measured quality at acute trust level over time in the NHS in England. Data for 6 years covering the period 2011-2016 were analysed. The main quality outcomes assessed to be associated with hospital financial performance included readmission rates, patient inpatient satisfaction, A&E 4 hour waiting time excess waits, delayed discharge days in hospital, and cancer 2 week waiting time and 62 week treatment waiting time targets. Simple descriptive time trends and differences in outcomes by deciles (pooled over all periods, and stratified by years) of financial performance were presented. A linear regression model pooling data across all time periods and controlling for characteristics of the hospitals was estimated.

The study presents clear evidence of significant variation in financial performance as measured by operating margins. Around a quarter of trusts experienced negative and large operating margins between 2011-2013. There was however a marked decline in financial performance across all acute NHS trusts from 2013 onwards with nearly 90% of trusts experiencing negative operating margins by 2016. This marked downturn in the financial efficiency of trusts was also accompanied by worsening trends in some of process measures of care quality: mainly A&E waiting time increases, delayed patient discharges and increased cancer assessment and treatment waiting times, although readmission rates and patient elective treatment satisfaction remained stable. Cross-sectional patterns confirmed this with the 10% worst financially performing trusts having significantly worse process indicators of quality compared to the 10% best performing trusts, but over time this gap in performance narrowed (as all trusts financial performance deteriorated). Linear regression analysis of the relationship between the quality indicators and financial operating margins, that conditioned on characteristics of the hospital trusts (namely volume of inpatient and outpatient activity, foundation trust status, teaching

status, geographic and socio-economic characteristics of region, expenditures on agency staff and consulting services, as well as bed capacity and occupancy rates) confirmed that A&E waiting time breaches and delayed transfers were significantly associated with worse financial performance.

The linear regression analyses only adjusted for number of beds available at the Trust and year (please see 'Methods', sub-section 'Statistical analysis' paragraph 2).

The authors acknowledge the complexity of understanding the relationship between financial performance and quality of care, and clearly identify the limitations with their study. However conclude that this is an important area of interest. The paper was reasonably well written although the description and presentation of the statistical results was not always very clear and detailed definition of the variables and full regression results would have been appreciated.

I found the topic of the research of interest with a high degree of policy relevance.

Thank you for your appreciation of the relevance of this work.

The underlying trends in hospital financial performance in the English NHS, associated characteristics of poor and well performing trusts as well as associated differences of trends in indicators of quality made for an interesting and thought provoking read.

Thank you for your kind comments.

The main weakness of the paper is the fact that the study design, which is primarily descriptive and observational. The theoretical and institutional considerations of how financial performance and quality may be related, and the factors driving the results was not thought through in a systematic way.

We agree that the study design is a weakness of the paper, though one not easily adjusted. This is fully acknowledged in the limitations section of the discussion, specifically that we are unable to demonstrate causal inferences.

This was reflected in the selection of indicators, explanation of the observed associations, and description of the financial and institutional operating environments. A lot of the observed associations appeared more or less mechanical in nature due to the reimbursement system and external changes in the health sector beyond the influence of the acute hospital sector.

The selection of indicators was driven by logistical constraints and is acknowledged in 'Methods', sub-section 'Outcome measures', paragraph 1 as well as 'Discussion', sub-section 'Limitations', paragraph 3. We have acknowledged the potential role of external factors such

as primary care that would be beyond the influence of the acute hospital sector in 'Discussion, sub-section 'Limitations', paragraph 5.

Given the focus of the paper was investigating the effects of financial difficulties on quality, rather than in understanding the mechanisms generating both financial problems and how they would impact or be aligned with quality; then in my view a deeper consideration of these factors is required in order to better interpret the findings and guide the research.

Our main focus in the paper was on the association with, rather than effect of, financial difficulties on quality. We have attempted to consider these factors in a deeper way in the discussion section, taking into account the remit of the paper and logistical constraints such as word count.

I have some more general comments below followed by some specific comments that are more minor:

General comments:

The authors need to better consider and explain what determines financial performance as measured by operating margins in acute (public) hospital sector in the English NHS, and how this relates to quality. In particular this has been precisely related to the reimbursement system, regulation policy and health care market circumstances (wider health and social care environment).

A full description of this could add significantly to the length of the discussion section. We would be happy to oblige if the editors feel it would improve the manuscript.

For instance, if you consider the measure of financial performance, explain what drives revenues and costs of care and how this would be related to quality. From what I understand the operating margin of the hospital is calculated as $(\text{Revenue} - \text{Costs})/\text{Revenue}$. This can be expressed in more detail as roughly $1 - \frac{\sum_i (FC_i + VC_i * Q_i)}{\sum_i (P_i * Q_i + R)}$. Where FC_i is the fixed costs of providing treatments i , VC_i are the variable costs associated with treatment i (can also be considered as the marginal costs of providing each additional treatment Q_i) of providing Q_i treatments, P_i is the reimbursement price of the i -th treatment and Q_i is the amount of the i -th type of treatment. R are additional revenues for instance from research, or from bonus payments related to quality improvement or other revenues earned outside of providing treatments to patients. Summing the costs and revenues over the i treatments gives you total costs and revenues. If total costs (numerator) equals total revenue (denominator) then net operating margins 0. If the denominator exceeds the numerator then operating margins are positive and if the numerator i.e. costs exceeds the denominator i.e. revenue then operating margins are negative.

This is an accurate characterisation of the factors driving revenue and care costs.

Looking at the denominator, it is clear that reimbursement prices are crucial in determining financial performance. In the English NHS these are fixed at the average cost across all trusts providing a given treatment i $P_i = \frac{(FCi_{\text{mean}} + VCi_{\text{mean}} * Q_{\text{mean}})}{(Q_{\text{mean}}) * MFF}$, adjusted for the market forces factor that impacts local costs i.e. wage differences and other price differences by region. Hence it would appear that hospital trusts that treat higher volumes of patients would benefit from economies of scale and therefore have lower average costs than the mean hospital. For instance their fixed costs would be spread over many more patients. It is also likely that their marginal costs would fall as patient volumes increase (up to a certain point) as higher volumes would improve treatment efficiency. Hence one would expect to observe economies of scale and possibly scope with larger hospitals enjoying better financial performance. We also know from the literature that higher volumes and concentration of services as well as expertise can increase quality of care. The higher volumes and revenues can also justify more investment in new technologies. Hence this would suggest that better financially performing trusts would have better quality (through economies of scale and scope and higher volumes as well as specialisation).

We have added a sentence to 'Discussion', sub-section 'Limitations', paragraph 5 to acknowledge the economies of scale that can result from high volume patient turnover.

The reimbursement price mechanism would also directly (mechanically) relate operating margins to quality. For instance reimbursement prices for specialised services are significantly higher and have been shown to be well above average costs of treatment (see papers by Street et al. CHE RP 103, 2014 on costs of specialised care). Hence hospitals providing specialised services will likely have higher revenues (above average costs) and are normally provide better quality of care – through specialisation, concentration of expertise and teaching status etc. The reimbursement system also penalises lower quality directly with no reimbursement for readmissions after 30 days. There are also best practice tariffs (BPTs) rewarding more cost effective treatment interventions such as day case surgery (lower risk, more efficient), which were associated with quality improvements (see Sutton et al, Health Economics) and CEQUIN quality improvement payments that reward up to 2.5% in additional performance payments compliance with quality targets. That is the R component in the formula would directly improve profit margins and be associated with better quality almost by definition. So from a revenue perspective would be a strong expectation to see better quality related to better financial performance.

We have added two sentences to ‘Discussion’, sub-section ‘Limitations’, paragraph 4 to acknowledge reimbursement prices for specialised services and best practice tariffs. This paragraph already makes clear that our inability to assess proportion of activity subject to a tariff is a limitation of the work.

There was also the introduction of a 30% marginal payment for emergency admissions in excess of 2010 levels that would penalise hospitals financially for admitting more emergency patients above an expected level, and no payment for never events. Emergency admissions are also more costly with lower margins and disrupt the ability of hospitals to undertake elective treatments which are more profitable. The authors also did not mention the efficiency savings target initiatives, where effectively all Pbr tariffs were reduced annually (by around 4-5%) since 2011. This would cut all hospital revenues and likely explains a lot of the observed reductions in financial performance. More emergency care, complex patients, volatile demand would reduce profits, and lower volumes. Again reinforcing a positive finance/quality relationship.

We have added a sentence to ‘Discussion’, sub-section ‘Conclusions and policy implications’, paragraph 1 to acknowledge financial penalties and negatively reinforcing reimbursement mechanisms.

On the cost side the expected relationship between costs and quality is less clear. The main concern would be that cost savings, for instance from having lower staff to patient and bed ratios, employing less well qualified staff (cheaper) e.g. lower grade nurses and doctors; having fewer senior staff to junior staff ratios, skill mix substitution with specialist nurses or junior doctors doing the work of more experienced and qualified staff, and lower investment in new and better quality equipment could lower quality and productivity. Hence this could lead to better financial performance but at the expense of quality. However in terms of efficiency and productivity employing less skilled staff and using lower quality and cheaper technology could reduce efficiency of care e.g. increase length of stays, increase complications and treatment costs, lower productivity and reduce profitability as well as quality. Also lower volumes would lead to higher average and marginal costs and may be associated with lower quality of care.

We agree with the possible impacts these factors can have and specifically that better financial performance may occur at the expense of quality. This has already been acknowledged briefly in ‘Discussion’, sub-section ‘Conclusions and policy implications’, paragraph 1 by considering the opposite scenario: *“Trusts with financial deficits may be spending more than they can afford ... and one could therefore argue that higher quality should be expected for this extra financial outlay.”*

There is evidence that lower staff ratios due reduce quality. Also cutting down on management and administrative expenses can lead to lower quality of care and worse financial performance in terms of lost productivity, particularly if it leads clinicians to allocate more time to administration and management and the expense of improving quality of care or productivity. Hospital costs will also be a function of patient severity with more complex patients costing more and more likely to experience adverse outcomes, and be associated with worse quality as measured by outcomes.

We agree and have acknowledged the impact that lower staff ratios can have in ‘Discussion’, sub-section ‘Comparison with other studies’, paragraph 2. ‘Discussion’, sub-section ‘Conclusions and policy implications’, paragraph 1 to acknowledge the point about cutting down on management.

The study also neglects to mention the influence of patient choice and the use of private sector

providers, in particular independent sector treatment centres who are focused on specific elective procedures that are fairly high volume, routine and standardised. They are focused factories providing efficient treatment at good quality with low waiting times, and have attracted increased market share during the period of study. They also tend to select less complex patients, hence lower cost patients.

We have added a sentence to 'Discussion', sub-section 'Limitations', paragraph 5 to acknowledge this point.

This has implications for the authors analysis as hospital trusts facing greater competition from ISTCs will lose market share and revenues these providers, have lower revenues and likely be left with treating more complex patients. Hence hospital trusts facing greater competition from ISTCs would likely be in a worse financial position apparently have worse quality indicators.

We have added a sentence to 'Discussion', sub-section 'Limitations', paragraph 5 to acknowledge this point.

The introduction of patient choice and patient reported outcomes and information on hospital waiting times, mortality rates and volumes by procedure has also been shown to have lead to significant changes in patient flows, with patients switching to better quality hospitals (see for instance research by Propper et al. Gutdeker, Gravelle, Moschini and Sicliani etc). Hence this would reinforce the relationship between better quality and financial performance (in terms of revenues from attracting more patients, and scale economies). The impact of competition on quality and financial performance should therefore be mentioned and discussed. The positive aspects would be that competition on quality would drive up quality and reward hospitals with greater volumes of patients and revenue. However greater competition would also see some hospitals lose out to more focused providers and face higher costs, lower volumes and revenues. In order to compete would have to improve quality to attract patients. Patient choice and selection – evidence that patients switched for elective treatment to better quality providers, lower waiting times, better outcomes. This would have increased profits at better quality hospitals, more elective care and volumes. Also Trusts operating in more competitive markets may have had to invest more in quality to attract patients and to be more efficient, but may have had to be operating on lower margins. Hence competitive forces may have been sufficient to sustain quality, but may also have contributed to lower margins.

We have added a sentence to 'Discussion', sub-section 'Limitations', paragraph 5 to acknowledge this point.

The study also needs to consider the effect of the imposed nurse staffing ratios on costs and quality. This was mandated and is imposing additional costs on trusts where shortages existed or where attracting regular nursing staff is problematic (high demand for agency staffing etc). It is important to understand that this was imposed, creating considerable additional costs that would lower margins, but may be improving quality (hence leading to a negative relationship between margins and quality as trust efficiency reduced, but quality improves).

We have modified a sentence in 'Discussion', sub-section 'Conclusions and policy implications', paragraph 1 to acknowledge this point.

The paper does not consider the issues related to foundation trusts. They have stronger incentives to reduce costs as well as increase margins. Likely would also look to have higher revenues by attracting more patients through quality and doing high profit margin procedures. I would consider separating out foundation and non-foundation trusts in the analysis to see if margins are different and relationship between quality and lower margins also differ.

We elected not to separately assess Foundation and non-Foundation trusts but would be happy to add this as a sensitivity analysis if the editors so wish.

Finally the study needs to mention the other issues in health and social care impacting hospital quality and financial performance, namely primary and community care as well as social care. Namely investment in primary care has stagnated over the period of the study with number of GPs not keeping pace with population needs. Evidence that access and quality of GP care reduces secondary care costs through reduced emergency admissions (see Asaria and Cookson). Hence poor financial performance and lower quality process measures may be observed in acute trusts due to rising emergency admissions, which may have nothing to do with quality of secondary care. Similarly access to community care, tertiary and social care has been squeezed, this will impact discharge delay, readmissions and hence hospital costs and quality (see Gaugan, Siciliani and Gravelle for papers on discharge delay and availability of social care). However it is totally beyond the control of the NHS Trusts.

We have briefly touched on primary care systems and how these may impact acute hospital Trusts in the second sentence in 'Discussion', sub-section 'Limitations', paragraph 5.

The discussion may want to elaborate on these points. In particular, that to improve quality and financial performance the trusts may need to integrate care and align resources more closely with primary and tertiary care. Allocating more resources to outpatient services in the community, early diagnosis, primary and secondary prevention as well as discharge management and patient follow up as well as rehabilitation. Current reimbursement and funding structures with payment on for inpatient and hospital based outpatient activity does not encourage this and if the integration and quality of care in the ambulatory sector is not improved hospital financial positions may not get any better. Hence new funding and payment models may be needed to improve financial performance and quality (for instance bundled payments for care across care sectors).

We agree with the reviewer that new funding and payment models may be necessary. However, after careful discussion, we elected not to over reach with implications from our study that provides evidence only of associations rather than causal links. As a result, it might be prudent to first more closely investigate some of the hypothesized causal pathways and intercede with interventional studies such as cluster randomised trials of policy impacts before asserting definitively that new payment models would be beneficial. We touch on this requirement for interventional evidence in 'Discussion', sub-section 'Conclusions and policy implications', paragraph 4.

These points above should be considered in the motivation of the paper and specification of the analysis. For instance, in explaining the specification of your longitudinal regression model and selection of control variables. In my opinion, if the paper wants to make a more significant contribution you should identify a set of quality indicators more directly under the control of the hospital (inpatient waiting times, standardised hospital and 30 day mortality rates, hospital acquired infection rates, adverse hospital events, staff satisfaction and recommendations). I would then look to exploit the longitudinal data and timing of external changes in the reimbursement and health care environment that has put trusts under financial pressure, and look to see how quality had responded to these financial pressures. Hence you could condition on baseline quality, observe trusts who experienced significant changes in their financial situation due to the changes in reimbursement reforms or external situations (such as increased A&E consultations, Reduced supply of primary and tertiary care, reductions in reimbursement rates etc. or increased ISTC competition, imposition of minimum nursing and other staff ratios plus the costs of hiring extra agency staff). These would be exogenous external shocks impacting operating margins and then one could look to see over time following this externally imposed financial pressures, how quality has responded.

Part of the rationale for our quality indicators was logistical constraints (i.e. data that was publicly available and could be easily parsed at Trust level) as discussed in 'Discussion', sub-section 'Study limitations', paragraph 3. We elected not to longitudinally assess the impact of exogenous shocks given the limited granularity of the data ('Discussion', sub-section 'Study limitations', paragraph 1). Further, there are several competing external impacts over the study period, some general (implementation of Health and Social Care act) and some local (factors described by the reviewer above such as increased competition from ISTCs). Parsing out these impacts was not something we felt would be reliably or robustly achieved with the data we had available to us.

Specific comments:

The panel data model is very simplistic and full results are not provided. Did you estimate a GLS random effects model? Did you consider a log-linear specification, also was it possible to condition on baseline outcomes? May want to look at analysis that stratifies by foundation or non-foundation trust status. Did you include a time trend in the analysis?

Our model was an intentionally very simple linear regression with the outcome as the dependant variable and the following independent variables (operating margin, number of beds available and year). Each Trust in each year was treated as a separate observation with standard errors clustered by Trust to account for the non-independence of Trust-level data. We did not make use of random effects.

It is not clear what extra results are requested here. We would be happy to add the output from the statistical software to the supplementary appendix if the editors so wish. We intentionally kept the linear regression analysis as simple as possible, hence very few independent variables. Time trend analyses were not performed for reasons detailed in our response points above.

I would try to look to exploit a before and after comparison of quality in response to specific exogenous shocks that impacted hospital margins exogenously (and possible differentially by hospital).

We have addressed this point in an earlier response above.

Paper needs to provide more background and detail on reimbursement system, competitive environment and incentives to provide high quality care and efficient care: For instance the reduced payments for readmissions, emergency treatment marginal payment reductions, no payment for never events, PbR efficiency saving tariff reductions. CEQUIN and Best Practice Tariffs. Introduction of minimum nursing staff ratios.

We have added brief details on some of these points to the discussion section as detailed in our earlier responses above.

Reviewer: 2

Reviewer Name: Adrián Villaseñor

Institution and Country: Centre for Health Economics, University of York, United Kingdom Please state any competing interests or state 'None declared': I have attempted to give my unbiased and honest review of this paper, however, I would like to state that:

I am currently working on a very similar research article. The outcome variables, interest variables, and sources of data are very closely related. This could have influenced the way in which I frame my comments.

We appreciate the candour of the reviewer and are thankful for the time taken to provide comments on our manuscript.

Please leave your comments for the authors below

Thanks to the authors for addressing this very important issue with clarity and rigour. I would like, however, raise a few points of concern that are likely to improve your paper.

1. I think the paper could benefit from a theoretical or conceptual discussion, with appropriate references, of why these concepts (financial performance and their outcomes) might be correlated. Without any explanation of the mechanisms behind the correlation between two variables, any statistical analysis can be seen as spurious.

We have briefly touched on the rationale behind a correlation towards the end of the 2nd paragraph of the introduction and later in the discussion section (Conclusions and policy implications). We would be happy to add more if the editors feel it would improve the manuscript but are conscious of the burgeoning word count.

2. The claim on page 4, lines 46-51, should be backed up with appropriate references.

We have watered down the wording of the link to quality and added a reference.

3. I would suggest the authors state the financial periods clearly, and in full (e.g. 2010-2011) as it can be confusing for the reader (see page 2, line 12 vs page 5, lines 13-17).

We have adjusted the text on page 2 (abstract) to clarify that these are financial years. We would be happy to adjust all manuscript text to state the full financial year periods if the editors wish us to.

4. I think the paper would benefit from further justifying your measure of financial performance. For Foundation trusts I think it is fine to use "Table ID 1, Subcode 100" as a measure of Turnover, but I would ask you to justify why you are using "Table ID1, Subcode 160" as a measure of Surplus/Deficit. Same point applies to Surplus/Deficit measures for Non-Foundation trusts. These measures of surpluses/deficits include financial items that trusts can use to 'adjust' their surplus/deficit. The use of these financial devices might also be related to trusts outcomes, which will make your sample biased.

We acknowledge that this choice introduces the potential for bias via gamification of financial figures by Trusts as we describe in 'Discussion', sub-section 'Limitations', paragraph 2. It was made on logistical constraints; we would have preferred to use breakeven performance figures had the data been easily available.

5. Moreover, since you are not doing the analysis separately by foundation / non foundation trusts, I think there is a discussion missing of whether the two concepts, "SURPLUS/ (DEFICIT) FOR THE YEAR" for foundation trusts, and "Adjusted retained surplus/(deficit)" for Non-Foundation trust are comparable.

We elected not to analyse the Foundation and non-Foundation Trusts separately but would be happy to add this to the sensitivity analyses if the editors so wish. Within the limitations already discussed, we felt that the measures were comparable between Foundation and non-Foundation Trusts.

6. Your main measure of Financial Performance is "operating margin" as a percentage of turnover. Could you please provide a further explanation of why this might be an appropriate measure that takes into account differences in trusts size?

As different trusts have widely varying revenues based on their size and activity, we chose to assess operating surplus/deficit as a proportion of their turnover. This measure is not perfect as it does not take into account the economies of scale that will apply to larger trusts, i.e. the financial relationship between smaller and larger trusts may not be linear. However, we felt it was a reasonable practical compromise that has previously been used in the literature (see references 17 and 19).

7. A point related to point 4 and 6 then would be, is it appropriate to call your measure "operating margin" (page 2, line 6). Please justify further.

We have used the term in line with other studies in the literature (see references 17 and 19). We would be happy to rephrase if the editors feel it would improve clarity and a superior term can be selected.

8. Tables should be able to be understood even in the absence of the article's text. Please consider 1) Formatting your tables in a way that are visually appealing, and 2) add any table footnotes that allow the reader to understand the table.

We have made some changes to the Table titles and legends to make them clearer. We would be happy to further edit the formatting as per specific reviewer or editor feedback.

9. Table 4 uses the word "impact", which you can't infer from your statistical analysis. Please avoid any causal language throughout your paper.

We have edited this to remove the word 'impact'.

10. I find Figure 1 uninformative. Please consider doing histograms per time period.

We included this figure as a snapshot rather than histograms per time period given the description in the text of the proportion of trusts in deficit. We would be happy to change the figure as suggested, or perhaps remove it altogether, if the editors so wish.

11. Please justify further why you did not add year dummies into your regression analysis and add a footnote to Table 4 stating what control variables you used for each regression.

We did add year dummies into our regression analysis ('Methods', sub-section on 'Statistical analysis', paragraph 2). We have added to the legend of Table 4 to clarify the control variables used.

12. I do not see how the analysis presented in page 11, lines 1-11 with its accompanied figures, add anything to the analysis. It seems to come out from nowhere and makes the analysis hard to follow.

This analysis was included to reinforce the interconnected nature of many of the factors being investigated and provide a visual way of displaying how the associations have changed between the early and late periods in the study. We would be happy to remove it, or move it to the supplementary appendix, if the editors so wish.

VERSION 2 – REVIEW

REVIEWER	Mark Dusheiko Institute of Social and Preventive Medicine, University Hospital of Lausanne and Faculty of Business, Economics, Commerce University of Lausanne. Honorary Research Fellow Centre for Health Economics University of Lausanne
REVIEW RETURNED	21-Aug-2018

GENERAL COMMENTS	<p>The revised version of the paper reads well and is clearer in its presentation of the main findings. The limitations and interpretation as well as policy relevance is better defined. Also the clarity and conceptualisation of mechanisms better addressed.</p> <p>However I think the paper needs some very minor modifications. I think you should put foundation trust status in Table 1. Also you should stratify the analysis by foundation trust status in the supplementary appendix and test whether deficit distributions and the relationship between relative deficits and quality indicators are significantly different between the two sets of hospitals. This is</p>
-------------------------	--

	<p>important given the differing financial pressures and regulatory structures between the two sets of providers.</p> <p>Also in the text you downplay the cancer treatment waiting time targets (e.g. much smaller effects). I think you should not downplay the effects to this extent and it should be given more prominence (magnitude of association may be smaller but one could argue significance and importance in terms of patient quality of a relatively small increase in cancer treatment waiting times is of greater concern than longer A&E waiting times (which can often be for minor injuries or illnesses).</p> <p>Finally, you should read and refer to the attached Nuffield Trust paper in your paper which details a lot of the key underlying financial forces impacting trusts revenues and costs and hence the distribution of trust deficits over time (Gainsbury, The Bottom Line, Nuffield Trust)</p>
--	---

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Mark Dusheiko

Institution and Country: Institute of Social and Preventive Medicine, University Hospital of Lausanne and Faculty of Business, Economics, Commerce University of Lausanne. Honorary Research Fellow Centre for Health Economics University of Lausanne

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below Dear Authors,

Thanks you for your considered response to my comments on your first draft.

The revised version of the paper reads well and is clearer in its presentation of the main findings. The limitations and interpretation as well as policy relevance is better defined. Also the clarity and conceptualisation of mechanisms better addressed.

However I think the paper needs some very minor modifications. I think you should put foundation trust status in Table 1.

We have made this change to table 1 as requested.

Also you should stratify the analysis by foundation trust status in the supplementary appendix and test whether deficit distributions and the relationship between relative deficits and quality indicators are significantly different between the two sets of hospitals. This is important given the differing financial pressures and regulatory structures between the two sets of providers.

We have made this change by adding tables D, E1 and E2 to the supplementary appendix.

Also in the text you downplay the cancer treatment waiting time targets (e.g. much smaller effects). I think you should not downplay the effects to this extent and it should be given more prominence

(magnitude of association may be smaller but one could argue significance and importance in terms of patient quality of a relatively small increase in cancer treatment waiting times is of greater concern than longer A&E waiting times (which can often be for minor injuries or illnesses).

We have made this change as requested by adding to paragraph 2 of ‘Conclusions and policy implications’ section in the discussion. We have also removed the words ‘considerably’ and ‘much’ when referring to the difference in the abstract and discussion.

Finally, you should read and refer to the attached Nuffield Trust paper in your paper which details a lot of the key underlying financial forces impacting trusts revenues and costs and hence the distribution of trust deficits over time (Gainsbury, The Bottom Line, Nuffield Trust)

We have added this reference (23) and referred to it in the text in the discussion (second paragraph of ‘Limitations’ and first paragraph of ‘Conclusions and policy implications’).

VERSION 3 – REVIEW

REVIEWER	Mark Dusheiko Institute for Social and Community Medicine and Department of Business and Economics, University of Lausanne Switzerland
REVIEW RETURNED	08-Oct-2018

GENERAL COMMENTS	<p>Thank you for providing your updated responses to my previous queries and subsequent modifications to the paper. I am happy with the changes, however, I just have a couple of minor comments:</p> <p>1) You do not refer to the new Foundation Trust results in the appendix. I thought they were interesting a required comment - namely that a) Foundation Trusts (or Trusts that became Foundation Trusts) were generally more concentrated in the upper part of the distribution of operating margins (although were still relatively high proportion in the lowest 10th percentile). They also are associated with better performance on quality indicators.</p> <p>But more interestingly b) the association between quality indicators and operating margins were not as strong for foundation trusts compared to non-foundation trusts. Note that for non-foundation trusts association between inpatient satisfaction and operating margins were significant and positive (while foundation trust positive, but smaller and less significant i.e. worse operating margins more strongly associated with lower inpatient quality for non-foundation trusts), and similarly for cancer targets where larger and more significant positive association found between operating margins and quality for non-foundation trusts (Tables E1 and E2). These findings merit a mention in the paper and probably discussion. Foundation trusts moved earlier to new reimbursement and funding models, assuming greater financial responsibility and subject to greater financial stress. Hence may have already been in a stronger pressure to adapt to worsening financial situation or may simply be that financial pressures are not as acute. I think this is an important finding and deserves a brief mention (especially as could also hypothesise that greater financial pressures and autonomy for Foundation Trusts may lead them to prioritise financial performance over quality).</p>
-------------------------	---

	<p>Finally, I found the first sentence on the last paragraph of the discussion a bit tangential and could be dropped or made more relevant.</p> <p>"Finally, our inability to demonstrate causal links on the available macro level public data re-emphasises the need for higher quality interventional studies such as cluster randomised trials specifically assessing policy impacts before implementation en masse. "</p> <p>Your study is looking at the association of financial pressure on quality, and hence you can not randomise trusts to financial difficulties so this is not a relevant option. There are a multitude of policy initiatives and funding decisions, likely to have contributed to current growth of trust deficits so even careful piloting of individual policies may not have identified the risk to trusts of financial difficulties. Hence, I would drop this sentence.</p>
--	--

VERSION 3 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Mark Dusheiko

Dear Authors,

Thank you for providing your updated responses to my previous queries and subsequent modifications to the paper. I am happy with the changes, however, I just have a couple of minor comments:

1) You do not refer to the new Foundation Trust results in the appendix. I thought they were interesting a required comment - namely that a) Foundation Trusts (or Trusts that became Foundation Trusts) were generally more concentrated in the upper part of the distribution of operating margins (although were still relatively high proportion in the lowest 10th percentile).

We have added a comment regarding the distribution of Foundation Trusts among operating margin deciles to the results section (paragraph 4) though our reading is that the proportion looks reasonably similar among decile groups.

They also are associated with better performance on quality indicators.

We have added a comment regarding this to the results section (paragraph 8).

But more interestingly b) the association between quality indicators and operating margins were not as strong for foundation trusts compared to non-foundation trusts. Note that for non-foundation trusts association between inpatient satisfaction and operating margins were significant and positive (while foundation trust positive, but smaller and less significant i.e. worse operating margins more strongly associated with lower inpatient quality for non- foundation trusts), and similarly for cancer targets where larger and more significant positive association found between operating margins and quality for non-foundation trusts (Tables E1 and E2).

We have added a comment regarding this to the results section (paragraph 8). These findings merit a mention in the paper and probably discussion. Foundation trusts moved earlier to new reimbursement and funding models, assuming greater financial responsibility and subject to greater

financial stress. Hence may have already been in a stronger pressure to adapt to worsening financial situation or may simply be that financial pressures are not as acute. I think this is an important finding and deserves a brief mention (especially as could also hypothesise that greater financial pressures and autonomy for Foundation Trusts may lead them to prioritise financial performance over quality).

We have added a brief segment discussing this to the 'Conclusions and policy implications' section of the discussion.

Finally, I found the first sentence on the last paragraph of the discussion a bit tangential and could be dropped or made more relevant.

"Finally, our inability to demonstrate causal links on the available macro level public data re-emphasises the need for higher quality interventional studies such as cluster randomised trials specifically assessing policy impacts before implementation en masse. "

Your study is looking at the association of financial pressure on quality, and hence you can not randomise trusts to financial difficulties so this is not a relevant option. There are a multitude of policy initiatives and funding decisions, likely to have contributed to current growth of trust deficits so even careful piloting of individual policies may not have identified the risk to trusts of financial difficulties. Hence, I would drop this sentence.

We agree with the points raised by the reviewer and have removed the sentence in question.