Work in progress: do not circulate without the authors' discretion

Do financial incentives to improve quality of care lead to better patient outcomes?

Yiu-Shing Lau, Matt Sutton

The University of Manchester

Email for correspondence: yiu-shing.lau@postgrad.manchester.ac.uk

Allocated paper number: A049

Acknowledgements

This project was partially funded by the National Institute for Health Research Health Services and Delivery Research (NIHR HS&DR) programme (project number 08/1809/250). Visit the HS&DR website for more information. The views and opinions expressed are those of the authors and do not necessarily reflect those of the HS&DR programme, NIHR, NHS or the Department of Health. We are grateful to Nancy Prospero of the NHS Greater Manchester Business Intelligence Service and Lesley Kitchen of the Advancing Quality Alliance (http://www.advancingqualityalliance.nhs.uk/) for previous discussions of this work and providing the data.

Abstract

Background: Pay-for-Performance (P4P) schemes typically reward improvements in specific process measures of the quality of care and result in small improvements in these measures. We aim to examine whether the improvements in health outcomes associated with introduction of P4P can be linked to improvements in the quality of care received at patient and/or organisational level for patients with Pneumonia.

Data: Two datasets linked at individual level containing a rich set of provider, area and patient characteristics derived from hospital care records, five process measures of quality, and patient mortality. Our final sample consists of over 98,000 individuals spanning 18 quarters from 2008-2012.

Methods: Cross section and panel data models.

Results: Some of the process quality measures are significantly associated with better health outcomes at a trust level but the magnitudes of the estimated coefficients are too large to represent clinically plausible direct consequences of these process measures.

Conclusion: Our findings suggest that these financial incentives to improve quality weakly lead to improved patient outcomes through their direct effects on the process measures that were incentivised. This P4P scheme appears to have also led to improved patient outcomes by inducing positive spill overs in terms of wider improvements in care quality across unmeasured dimensions and improvements in care for all patients.

Introduction

Pay-for-Performance (P4P) schemes have been introduced as a method to improve health outcomes by incentivising healthcare providers to provide a level of care to earn a financial bonus. The quality of care is measured by delivery of process measures, usually for high-volume health conditions (Herck et al. 2010). The incentivisation of these process measures aims to motivate healthcare professionals to treat patients more consistently within a provider and also between healthcare providers.

P4P schemes use process measures of care as an indicator of quality rather than health outcomes due to the more attributable nature of process measures. A hospital has full control on the delivery of process measures (Ryan et al. 2009), whereas health outcomes may be due to factors outside of the hospitals control.

We consider a P4P scheme introduced in England, the Advancing Quality (AQ) initiative, which is a hospital scheme introduced in October 2008. Under this scheme, hospitals in the North West (NW) region of the England were financially incentivised to provide process measures of care to patients. Delivery of these process measures of care was meant to improve health outcomes and ensure equitable treatment (Ledward, Horne, and Butterworth 2008). Performance is assessed by the providers' achievement rates of the selected process measures.

AQ initially incentivised five clinical areas; acute myocardial infarction, heart failure, community acquired pneumonia, coronary artery bypass graft and hip and knee replacements. For this analysis we examine pneumonia as the sole clinical condition. This has the largest number of patients admitted out of the conditions incentivised under the AQ scheme.

A recent study by Sutton et al. (2012) tested the impact of the Advancing Quality scheme on patient mortality for pneumonia, heart failure and acute myocardial infarction. The authors adopted a triple-difference design with other regions of England and six non-incentivised conditions using patient level data from Hospital Episodes Statistics, and found that AQ had resulted in a significantly lower 30-day within-hospital mortality rate among patients admitted for pneumonia.

We analyse whether this reduction in patient mortality can be attributed to the increase in delivery of process measures of care using a unique patient-level linked dataset. Specifically, we examine whether the outcome gains are attributable directly to the improvements in the quality of care delivered at individual patient level and at the organisational level.

Process measures as predictors of health outcomes

Although process measures of care are selected based on clinical research ("Process of Care Measures - Centers for Medicare & Medicaid Services" 2013) and are a proxy for the quality of care (Lee et al.

2011), they may not result in significant improvements in patient health outcomes such as patient mortality. A paper by Ryan et al. (2009) outlined five reasons why process measures of care may not result in any effect on health outcomes: 1) process measures are ineffective at reducing health outcomes; 2) not all providers with the P4P scheme will implement the process measures correctly; 3) process measures of care may become obsolete over time; 4) attention may be diverted away from processes that are not incentivised and 5) measurement error and gaming by providers.

Why is this important?

Finding the relationship between process measures of care and health outcomes is important as improving health outcomes is the goal of the policy maker (Health 2000). As process measures are incentivised due to them being proxies for quality of care and aim to lead to better health outcomes, it is important that process measures of care have a causal effect on patient health outcomes. If process measures of care are found not to be causal, or even correlational, to improvements in patient health outcomes, new measures should be incentivised that have such direct effects.

What is known?

There have been many studies on the correlations between process measures of care and health outcomes. A recent study by Lee et al. (2011) obtained one year of US data, 2001, from Medicare and Medicaid, and focused solely on pneumonia. Their two outcomes of interest were the 30-day mortality rate and the 30-day readmission rate measured at patient level and the number of process measures a patient had received. Using multilevel logistic regressions, they found no association between both of the health outcomes and the number of process measures of care received. Due to data limitations, these authors were not able to test whether a specific process measure of care impacted on the health outcome, only the number received. These will differ between patients due to severity and need, which were not controlled for.

Four papers have studied the relationship between patient mortality and process measures of care for AMI, HF and PN (Jha et al. 2007; Werner and Bradlow 2006; Werner, Bradlow, and Asch 2008; Ryan et al. 2009). Jha et al. (2007) and Ryan et al (2009) used multivariate logistic regression techniques on patient-level mortality and hospital mean performance on process measures of care. Werner and Bradlow (2006) and Werner et al. (2008) used hospital-level patient mortality and applied Bayesian analysis. All studies used data from the US and found that higher levels of achievement on process measures were associated with a lower mortality rate. Ryan et al (2009) adopted a fixed effects approach to test for causal effects between process measures and outcomes. The authors found that the process measures did not have a causal effect on outcomes. The other three papers all lacked the ability to find any causal effects due to data limitations which meant that panel data analysis could not be

conducted. All four papers also did not have individual level data on both outcomes and process measures which meant that interpretation of findings are restricted due to potential problems with the ecological fallacy (Finney et al. 2011).

Four studies only focused on either HF, AMI or both using data from the US; (Bradley et al. 2006; Peterson et al. 2006) use mortality as the outcome whereas (Luthi et al. 2003; Fonarow et al. 2007) use readmissions as the health outcomes. Each paper used a variety of methods from Poisson regression and hierarchical GLM. One paper did not define the analysis method. The authors found that process measures of care were weakly related to health outcomes, with higher achievements on process measures of care associated with lower patient mortality and readmissions. Each of these papers suffered from poor study designs varying from low number of observations (Luthi et al. 2003), only older populations (Bradley et al. 2006) and non-transparent methods (Peterson et al. 2006).

Three papers used data from outside of the US. All three used multivariate logistic regressions, regressing patient-level mortality as the health outcome. Granger et al. (2005) studied coronary syndromes across 14 countries using process measures of care aggregated to the hospital level. Luthi et al. (2004) used one year of data from 1999 from three Swiss health centres and focused on heart failure patients. Bray et al. (2013) used data from the UK, linking Hospital Episode Statistics to the Stroke Registry. All three studies found that higher achievement in process measures were weakly associated with lower mortality rates. All three papers adopted cross sectional designs which meant that only associations could be found. Luthi et al. (2004) used high quality patient level data, but suffered from a small study sample.

Most previous studies have found weak associations between process measures of care and health outcomes across a variety of clinical conditions. One paper found no association and one paper found no causal effect. These studies have two main limitations; the first is the use of aggregated data where the ecological fallacy may limit interpretation of the relationship between process measures of care and health outcomes. The second is short study periods and cross-sectional designs meaning that correlations cannot be interpreted as causal effects.

Data

We obtained data from the AQ programme. These included hospital records from the Secondary Uses Service (SUS) containing patient and Trust-level characteristics at spell level and data from the programme's Quality Measures Reporter (QMR), which records delivery of the process measures of care for each patient. The linked data contained 98,771 patient spells for patients admitted with

pneumonia across 18 quarters, October 2008 until April 2013. This encompasses almost[1] the entire population of pneumonia patients qualifying for inclusion in the AQ programme from the 24 Trusts in the North West of England.

The introduction of AQ meant that Trusts who volunteered to be a part of the scheme in the North West of England[2] had to record process measures of care given to patients for data collection. Being a new system to the NHS, initial data issues, such as data linkages and missing months of data, arose in some Trusts. This resulted in one trust missing one quarter of data, one trust missing two quarters of data and One other trust has missing data for the final year of observation.

We also obtain data from Hospital Episode Statistics (HES) which provided an in-hospital 30-day unadjusted and risk adjusted mortality rate at trust level over 14 quarters of data, October 2008 until April 2012. This data encompasses all patients with pneumonia from the 24 Trusts in the North West of England, and, is the extended outcome variable used by Sutton et al (2012). The risk adjusted mortality rate controls for patient severity, comorbidities, admission source, admission method, age and sex.

Outcome variable

We generated a dichotomous variable for within-spell in-hospital mortality as our outcome variable. This variable was generated from the discharge method in the hospital record and takes a value of one when the patient was discharged dead. To extend our analysis we use two more outcome variables at the trust level, an in-hospital 30 day unadjusted and risk adjusted mortality rate.

Process measures of care

The five process measures of care for pneumonia under AQ are: oxygenation assessment; initial antibiotic selection; blood cultures before antibiotics; initial antibiotics received within 6 hours of hospital arrival; and smoking cessation advice. For each process measure of care, we know whether the patient was given the measure, failed to be given the measure or was excluded from the measure. Exclusions are made when hospitals are validly able to remove patients from receiving process measures of care; a full list of exclusion criteria is shown on table 1

For our spell level analyses, we generated a set of dichotomous variables taking values of one if process measures were given, and another set of dichotomous variables which took values of one if the patient was excluded from the measure.

[1] The population of AQ qualifying Pneumonia patients is 104,435 over the 18 quarters. However, 5,664 patients did not have all covariates and therefore were dropped from the sample due to missing data. The sample of patients who were removed are slightly older by two years (73), and as a percentage, have a 1% higher in-hospital death rate of 25%

[2] All 24 trusts in the NW of England agreed to participate in the incentive scheme.

Methods

We model the relationship between health outcome and quality of care, where the quality of care is given by the delivery of process measures of care which are incentivised under the AQ scheme. We perform both cross sectional and panel data econometric techniques to extract both causal and correlational effects of process measures on health outcomes. We expect that the process measures of care are inversely related to health outcomes as the process measures. This relationship on an individual level may be hard to find due to selection of patients from hospitals.

Hospital level analysis

For our hospital level analysis, we aggregated all spell-level variables by Trust and quarter. We then estimated both fixed effects and random effects models.

The general model is given by:

$$y_{jt} = \alpha + \boldsymbol{\beta X_{jt}} + \boldsymbol{\gamma p_{jt}} + \theta t_t + a_j + u_{jt}$$

Where $\boldsymbol{X_{jt}}$ is a vector of care process measures, both achievement and exclusions; $\boldsymbol{\beta}$ are the coefficients of interest; $\boldsymbol{p_{jt}}$ is a vector of mean characteristics of patients (average age, percentage male, ethnic group proportions and average area income deprivation score[3]); t_t are time fixed effects. The error component is in two parts where; a_j is a random variable[4] with a constant mean, and u_{jt} is assumed to be independent and identically distributed (*iid*) with a mean of zero and a constant variance. The *j* denotes trust, and *t* denotes time.

The construction of each process measures of care variable is given by:

$$x_{jt} = \frac{1}{N}\sum(A) \; where \; N = A + F + E$$

N is the total number of spells within a trust for a quarter, *A* the number of patients given a process measure of care, *F* the number of patients not given the process measure and *E* the number of patients excluded from the process measure. We also constructed mean rates of exclusions with *E* placed in the numerator.

Spell level analysis

[3] Income deprivation score is from the 2010 Index of Multiple Deprivation. This variable is on an area level, linking a patient to their area deprivation score based on the proportion of the population in the area on income support (IMD, 2013). This area is defined as a Lower Super Output Area (LSOA) which is an area which contains around 1,500 patients.

[4] This error component in the fixed effects estimator are assumed to be time constant, and therefore swept out under the fixed effects estimation.

The spell level analysis was conducted using a OLS and probit regression on the spell level dataset. We estimate this general equation:

$$y_{ijt} = \alpha + \boldsymbol{\beta X_{ijt}} + \boldsymbol{\gamma p_{ijt}} + \delta h_j + \theta t_t + e_{ijt}$$

Where $\boldsymbol{X_{ijt}}$ is a vector of care process measures, both achievement and exclusions; $\boldsymbol{\beta}$ are the coefficients of interest; $\boldsymbol{p_{ijt}}$ is a vector or patient characteristics; age, sex, ethnicity and income deprivation score; h_j are hospital fixed effects; t_t are time fixed effects and e_{ijt} is the error term which is assumed to be *iid* with a mean of zero and a constant variance. *i* denotes patient spell, *j* denotes trust, whereas *t* denotes time.

However, this model has a potential endogeneity problem. The achievement of the patient may be determined by the expected health gain a trust thinks a patient will receive from a process measure of care; therefore a trust will select patients based on the level of health of a patient implying a revere causal effect of health outcomes on process measures of care. Therefore we use trust-quarter mean levels of performance on patient spell level outcomes to remove the endogeneity.

$$y_{ijt} = \alpha + \boldsymbol{\beta X_{ijt}} + e_{ijt} \ (1)$$

$$y_{ijt} = \alpha + \boldsymbol{\beta \bar{X}_{jt}} + \ \rho(\boldsymbol{X_{ijt}} - \boldsymbol{\bar{X}_{jt}}) + e_{ijt} \ \ (2)$$

We get equation two by adding and subtracting the trust mean performance ($\boldsymbol{\bar{X}_{jt}}$) into the right hand side of equation one. We therefore model a trusts average performance by quarter on patients' health outcome. The term ($\boldsymbol{X_{ijt}} - \boldsymbol{\bar{X}_{jt}}$) captures a patient's individual contribution to the trusts mean performance, therefore capturing the reverse causal relationship between y_{ijt} and $\boldsymbol{\bar{X}_{jt}}$**.** We are then left with an unbiased estimate of $\boldsymbol{\beta}$. We therefore estimate the following equation:

$$y_{ijt} = \alpha + \boldsymbol{\beta \bar{X}_{jt}} + \ \rho(\boldsymbol{X_{ijt}} - \boldsymbol{\bar{X}_{jt}}) + \boldsymbol{\gamma p_{ijt}} + \delta h_j + \theta t_t + e_{ijt}$$

Results

Descriptive statistics

Descriptive statistics are shown in Table 2. We find that 24% of patients who were admitted with pneumonia died in hospital. The average age of the admitted population is 73 years.

The achievement rates are much lower than the scores used by the AQ programme as we include patients who were excluded in the denominator. Much higher proportions of patients are excluded from the process measures of care than receive the care measures.

Oxygenation assessment has the highest level of achievement. 60% of patients were given this treatment. Initial antibiotic selection and antibiotics received in a timely fashion have achievement rates of around 30%. Blood cultures performed before initial antibiotic and smoking cessation advice are the two quality measures with the highest exclusion rates of around 80%. Smoking cessation has high as non-smokers are excluded from this measure.

Figure 1 shows the mean achievement rates of process measures and mortality rate over time. We find that mortality rate has steadily fallen by 0.5 percentage points from the first quarter to the last. The trends rates for all process measures, shows that achievement rates have increased steadily with the exception of

Panel data analysis

For our panel data analyses we ran both the random and fixed effects models; the results are shown in table 3. The Hausman test indicated that the random effects estimator is consistent which means the fixed effects estimate is also consistent but not as efficient as the random effects estimator.

We find that a one percentage point change in the achievement rate on the blood cultures measure and smoking cessation advice lowers the mortality rate by around 0.2 percentage points and 0.19 percentage points respectively. This effect is statistically significant at 5%. The timely delivery of antibiotics is correlated positively with the mortality rate across where a one percentage point increase in achievement results in an increase in mortality rate by 0.67 percentage points, although these results are not statistically significant.

Increasing the rate at which patients were excluded from blood cultures and smoking cessation advice by one percentage point lowers the probability of patients being discharged as dead by around 0.2 percentage points when statistical significance is measured at 5%. Timely delivery of antibiotics and antibiotic selection has statistically insignificant low positive effect sizes.

Spell level analysis

Table 3 shows the effects of the Trust-quarter mean achievement and exclusion rates on spell level mortality.

From the results of the probit regression, we find that a one percentage point increase in the provision of oxygenation assessment and timely delivery of antibiotics increases the probability of the patient dying, 0.34 percentage points for oxygenation assessment and 0.09 percentage points for timely delivery of antibiotics, when statistical significance is measured at the 5% level. We find that a one percentage increase in performing blood cultures and offering smoking cessation advice have lowers

the probability of a patient dying by 0.09 percentage points, however, these measures are not statistically significant.

Increasing the exclusion rate by one percentage point for oxygenation assessment, smoking cessation advice and the timely provision of antibiotics, increases the probability of death by 0.38, 0.46 and 0.08 percent points respectively.

The results from OLS are of similar size, sign and significance to the probit regressions. We find that a one percentage point increase in the provision of smoking cessation advice leads to a 0.12 percentage point decline in the mortality rate. Exclusions from smoking cessation advice is not statistically significant with a much smaller coefficient.

Alternative outcomes – panel data analysis

Table 4 shows results from panel data analysis to find the relationship between process measures of care and three mortality variables. In-hospital mortality is the same variable used in the previous analysis; however, to accommodate the two new outcome variables, we limited our sample to 14 quarters of data. Running panel data analysis on in-hospital mortality over 14 quarters of data has not resulted in any statistically significant results.

Using a 30 day in-hospital mortality measure, we find that a one percentage point increase in the provision of smoking cessation advice will decrease mortality rate by 0.18 percentage points. Other process measures of care did not result in any statistically significant results. Exclusion from smoking cessation advice is the only statistically significant process measure from the excluded category where a one percentage point increase in exclusion rate will lead to around 0.2 percentage point decrease in mortality.

Using a risk adjusted 30 day in-hospital mortality measure; we find a one percentage point increase in performing blood cultures before initial antibiotics and giving smoking cessation advice results in a 0.2 and 0.18 percentage point decrease in mortality rate, respectively. A one percentage point increase in the exclusion rate of these process measures of care is also associated with a reduction in the mortality rate by 0.16 and 0.18 percentage points.

Discussion

Findings

In the panel data analysis, we find a weak relationship between the process measures of care and health outcomes. When hospitals perform more blood cultures, lowers the probability of the patient mortality by around 20%; however, this effect may be driven by a small number of patients due to the high exclusion rates for this process measure.

We also find statistical significance for the rate at which trusts exclude patients from smoking cessation advice (21%); this result may be explained by the exclusion criteria of that process measure. As people who do not smoke are excluded from this indicator, our finding suggests that trusts treating more non-smokers have a higher survival rate.

We also did not find much relationship between hospital level achievement on process measures of care and health outcomes at the individual level. The significant results from the achievement of process measures of care is attributed to the fact that hospitals who generally provide the better quality of care, are treating patients who are harder to treat and therefore, that is related to a higher portability of patient mortality.

The results for being excluded from process measures of care from the probit regressions shows different relationships with outcomes when compared to the trust level analysis. Patients' likelihood of survival is lower at trusts, who on average exclude more patients from smoking advice. This result is indicative of the ecological fallacy where results from a trust level outcome variable do not yield the same correlation at patient level. The coefficient on the association of trusts exclusion rate on patient mortality is far higher than what seems plausible with a process measure which is not directly related to providing in hospital treatment. However, this relationship is not found when using the OLS regression method.

The results from the alternative measures of mortality show the importance of risk adjustment of the mortality variable. We find that by risk adjusting mortality, performing blood cultures and smoking cessation advice lowers the probability of mortality. For exclusion measures for blood culture and smoking cessation is more indicative of case mix measures as these two process measures have the highest rates of exclusions.

Looking at the achievement rates over time on figure 1, the lack of data recording for blood cultures may have driven some of the results on the trust level analysis when looking at in-hospital mortality rate. The results from table 4 do not include the final four quarters of data and therefore, is not subjected to the lack of data recording for that process measure of care.

How this relates to literature

Our finding adds to the study conducted by Lee et al. (2011) who also used PN as their sole study condition and aimed to find the relationship between patient mortality with the number of process measures of care a patient receives by introducing the process measures of care into the analysis. Compliment the study by Lee et al. (2011) as we find weak significance between mortality and process measures of care. However we do not find that all process measures of care are associated with lower mortality.

Other literature which also looked at the impact of process measures of care for PN (Jha et al. 2007; Werner and Bradlow 2006; Werner, Bradlow, and Asch 2008; Ryan et al. 2009), all found weak associations between the process measures of care and health outcomes. Like all previous literature we conclude that there are other factors which we do not explicitly capture, which is the cause of patient mortality.

This research also compliments the study by Sutton et al. (2012), the authors found that AQ reduced patient mortality. We used the same mortality variables which the authors used, however, we extended the time periods the variable covers. We found that some of the reduction in mortality is due to process measures of care.

Policy Implications

The selection of process measures for PN was evidence based, which showed that these measures are effective in improving health outcomes. However, this research has found that process measures of care have a weakly causal relationship for an improvement in health outcomes, where weak correlational relationships also seem to exist. More research is needed on the spill over effects of P4P to find where improvements in health outcomes (Sutton et al. 2012) is derived from, from a P4P scheme. Only then, will a policy maker be able to select the best approach to improve the quality of care given to a patient, which will lead to improvements in health outcomes. Even though our study has shown weak association between process measures of care on a patient's mortality for PN; the spill over effects of health care providers providing these process measures of care may be beneficial to the patient.

Other policy implications are; policy makers should weight and wait. Under AQ, all process measures of care have an equal weight when calculating overall trust performance. Our results suggest that not all process measures lead to lower patient mortality, therefore, this research adds support that different process measures should have different weights when rewarding trusts. Another implication of this research is that policy makers should wait for empirical research to be conducted before making decisions to remove process measures of care. Under the AQ scheme, process measures of care have been removed without any evidence base for doing so. This is a concern as we found that on a trust level, performing blood cultures before initial antibiotics is associated with a lower mortality rate. However, this measure has now been removed as a quality measure for AQ as not all patients with pneumonia qualify to receive this measure ("Advancing Quality: Measures for Pneumonia" 2013).

Strengths

To our knowledge, this is the first study which uses an individual level dataset. We have been able to link patients' outcomes to the process measures of care at patient level; received, failed to receive or excluded from, over a four year time period. This has enabled us to use a variety of panel data and

cross sectional econometric methods. We were able to adjust for the endogeneity between process measures of care and health outcomes where previous studies were unable to adjust for.

The flexibility from the data also enables us to deal with other studies limitations such as check the validity of the ecological fallacy by running our analysis on patient and trust level, where the data is based on the same patient sample; this provides a more comprehensive study of the links between the process measures of care and health outcomes.

This study also has the advantage of using data from an incentive scheme where all trusts within the region agreed to participate. This universal uptake of the incentive scheme removes the trust level self-selection bias to join the incentive scheme. This makes results more generalisable unlike the studies from the US where hospitals self-select into P4P programmes.

Limitations

Our research has a few limitations. Our outcome variable is within-spell in-hospital mortality. This is an exclusion criterion for some of the process measures, such as smoking cessation advice. A variable measuring mortality within 30-days of admission on a patient level would be preferable as it will enable us to limit our sample to patients who were discharged as alive. This may lower some endogeneity of the process measures and health outcomes on the individual level as patients who were excluded due to comfort measures will not be in our sample of patients.

We had few measures of patient case-mix, just gender, age, ethnicity and area of residence level deprivation. This issue however cannot be fully addressed with our dataset as such variables are not available.

Due to data recording issues, we do not have a balanced panel for our panel data analysis due to the missing trust data for certain quarters. This problem is encountered during the data extraction process where variables needed for the linkage of all datasets were not available; this issue was not resolved when a fuzzy match to extract the missing data was attempted.

In this study, we only focused on PN as the sole health condition which makes these results not generalizable for the AQ programme. However, this study does show the association of the process measures of care on health outcomes for PN.

Further research

One further change to this current research will be to limit our sample to patients who have been admitted for over two days. This is to reduce the sample to where the hospitals actions may influence the patient outcome. When patients who are admitted but have a high probability of dying, the delivery of process measures of care may not change the outcome due to the patient's poor initial health state.

Another possibility will be to calculate a patient level appropriate care score based on the proportion of process measures a patient received with the number of process measures the patient should have received. This analysis can be conducted on a patient level, and test whether the proportion of correct process measures of care will affect patient outcomes.

We aim to extend this study to look into more clinical conditions under AQ. Like Jha et al. (2007); Werner and Bradlow (2006); Werner, Bradlow, and Asch (2008); Ryan et al. (2009)Werner & Bradlow (2006), Werner et al (2008) and Ryan et al (2009), we will use AMI, HF and PN as three emergency conditions with mortality as the outcome variable. We will also extend our analysis to an elective condition, HIPKNEE, where we will have two outcome variables; readmissions and utility score from the EQ5D where we can control for pre and post-operative health states.

References

"Advancing Quality: Measures for Pneumonia." 2013. Accessed November 29. http://www.advancingqualitynw.nhs.uk/measures.php?measure=Pneumonia.

Bradley, Elizabeth H, Jeph Herrin, Brian Elbel, Robert L McNamara, David J Magid, Brahmajee K Nallamothu, Yongfei Wang, Sharon-Lise T Normand, John A Spertus, and Harlan M Krumholz. 2006. "Hospital Quality for Acute Myocardial Infarction: Correlation among Process Measures and Relationship with Short-Term Mortality." *JAMA: The Journal of the American Medical Association* 296 (1) (July 5): 72–78. doi:10.1001/jama.296.1.72.

Bray, Benjamin D, Salma Ayis, James Campbell, Alex Hoffman, Michael Roughton, Pippa J Tyrrell, Charles D A Wolfe, and Anthony G Rudd. 2013. "Associations between the Organisation of Stroke Services, Process of Care, and Mortality in England: Prospective Cohort Study." *BMJ (Clinical Research Ed.)* 346: f2827.

"English Indices of Deprivation 2010 - Publications - GOV.UK." 2013. Accessed November 12. https://www.gov.uk/government/publications/english-indices-of-deprivation-2010.

Finney, John W., Keith Humphreys, Daniel R. Kivlahan, and Alex H. S. Harris. 2011. "Why Health Care Process Performance Measures Can Have Different Relationships to Outcomes for Patients and Hospitals: Understanding the Ecological Fallacy." *American Journal of Public Health* 101 (9) (September): 1635–1642. doi:10.2105/AJPH.2011.300153.

Fonarow, Gregg C, William T Abraham, Nancy M Albert, Wendy Gattis Stough, Mihai Gheorghiade, Barry H Greenberg, Christopher M O'Connor, et al. 2007. "Association between Performance Measures and Clinical Outcomes for Patients Hospitalized with Heart Failure." *JAMA: The*

Journal of the American Medical Association 297 (1) (January 3): 61–70. doi:10.1001/jama.297.1.61.

Granger, Christopher B, Philippe Gabriel Steg, Eric Peterson, José López-Sendón, Frans Van de Werf, Eva Kline-Rogers, Jeanna Allegrone, et al. 2005. "Medication Performance Measures and Mortality Following Acute Coronary Syndromes." *The American Journal of Medicine* 118 (8) (August): 858–865. doi:10.1016/j.amjmed.2005.01.070.

Health, Department of. 2000. "The NHS Plan: A Plan for Investment, a Plan for Reform". Publication. July 1. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/dh_4002960.

Herck, Pieter Van, Delphine De Smedt, Lieven Annemans, Roy Remmen, Meredith B. Rosenthal, and Walter Sermeus. 2010. "Systematic Review: Effects, Design Choices, and Context of Pay-for-Performance in Health Care." *BMC Health Services Research* 10 (1) (August 23): 247. doi:10.1186/1472-6963-10-247.

Jha, Ashish K, E John Orav, Zhonghe Li, and Arnold M Epstein. 2007. "The Inverse Relationship between Mortality Rates and Performance in the Hospital Quality Alliance Measures." *Health Affairs (Project Hope)* 26 (4) (August): 1104–1110. doi:10.1377/hlthaff.26.4.1104.

Ledward, Jan, Alan Horne, and Jen Butterworth. 2008. "A North West Health System Approach to Advancing Quality". NHS North West.

Lee, Jonathan S, Brian A Primack, Maria K Mor, Roslyn A Stone, D Scott Obrosky, Donald M Yealy, and Michael J Fine. 2011. "Processes of Care and Outcomes for Community-Acquired Pneumonia." *The American Journal of Medicine* 124 (12) (December): 1175.e9–17. doi:10.1016/j.amjmed.2011.05.029.

Luthi, Jean-Christophe, W Dana Flanders, Stephen R Pitts, Bernard Burnand, and William M McClellan. 2004. "Outcomes and the Quality of Care for Patients Hospitalized with Heart Failure." *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care / ISQua* 16 (3) (June): 201–210. doi:10.1093/intqhc/mzh040.

Luthi, Jean-Christophe, Mary Jo Lund, Laura Sampietro-Colom, David G. Kleinbaum, David J. Ballard, and William M. McClellan. 2003. "Readmissions and the Quality of Care in Patients Hospitalized with Heart Failure." *International Journal for Quality in Health Care* 15 (5) (October 1): 413–421. doi:10.1093/intqhc/mzg055.

Peterson, Eric D, Matthew T Roe, Jyotsna Mulgund, Elizabeth R DeLong, Barbara L Lytle, Ralph G Brindis, Sidney C Smith Jr, et al. 2006. "Association between Hospital Process Performance and Outcomes among Patients with Acute Coronary Syndromes." *JAMA: The Journal of the American Medical Association* 295 (16) (April 26): 1912–1920. doi:10.1001/jama.295.16.1912.

"Process of Care Measures - Centers for Medicare & Medicaid Services." 2013. Accessed November 21. http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalProcessOfCareMeasures.html.

Ryan, Andrew M, James F Burgess Jr, Christopher P Tompkins, and Stanley S Wallack. 2009. "The Relationship between Medicare's Process of Care Quality Measures and Mortality." *Inquiry: A Journal of Medical Care Organization, Provision and Financing* 46 (3): 274–290.

Sutton, Matt, Silviya Nikolova, Ruth Boaden, Helen Lester, Ruth McDonald, and Martin Roland. 2012. "Reduced Mortality with Hospital Pay for Performance in England." *New England Journal of Medicine* 367 (19): 1821–1828. doi:10.1056/NEJMsa1114951.

Werner, Rachel M, and Eric T Bradlow. 2006. "Relationship between Medicare's Hospital Compare Performance Measures and Mortality Rates." *JAMA: The Journal of the American Medical Association* 296 (22) (December 13): 2694–2702. doi:10.1001/jama.296.22.2694.

Werner, Rachel M, Eric T Bradlow, and David A Asch. 2008. "Does Hospital Performance on Process Measures Directly Measure High Quality Care or Is It a Marker of Unmeasured Care?" *Health Services Research* 43 (5 Pt 1) (October): 1464–1484. doi:10.1111/j.1475-6773.2007.00817.x.

Table 1: Exclusion Rules

List of exclusion rules	Oxygenation	Antibiotic selection	Blood cultures	Antibiotics received	Smoking Cessation
Patients who were transferred between or within hospitals	x	x	x	x	
Patients receiving C*omfort Measures Only*	x	x	x	x	x
Patients less than 18 years of age	x	x	x	x	x
Patients who had no chest x-ray or CT scan that indicated abnormal findings within 24 hours prior to hospital arrival or anytime during this hospitalization	x	x	x	x	x
Patients with Cystic Fibrosis	x	x	x	x	x
Patients discharged on day of arrival	x	x	x	x	
Patients who died or were a still-birth on the day of, or the day after, arrival	x	x	x	x	x
Patients who discharged him/herself, or were discharged by a relative or advocate, on the day of, or the day after, arrival	x	x	x	x	x
Patients involved in clinical trials	x	x	x	x	x
Patients who did not receive antibiotics or a blood culture			x		
Patients who have received antibiotics within 24 hours prior to hospital arrival				x	
Patients who do not smoke					x

X signifies the exclusion criterion applies to the process measure of care.

Figure 1: Achievement rates and mortality over time

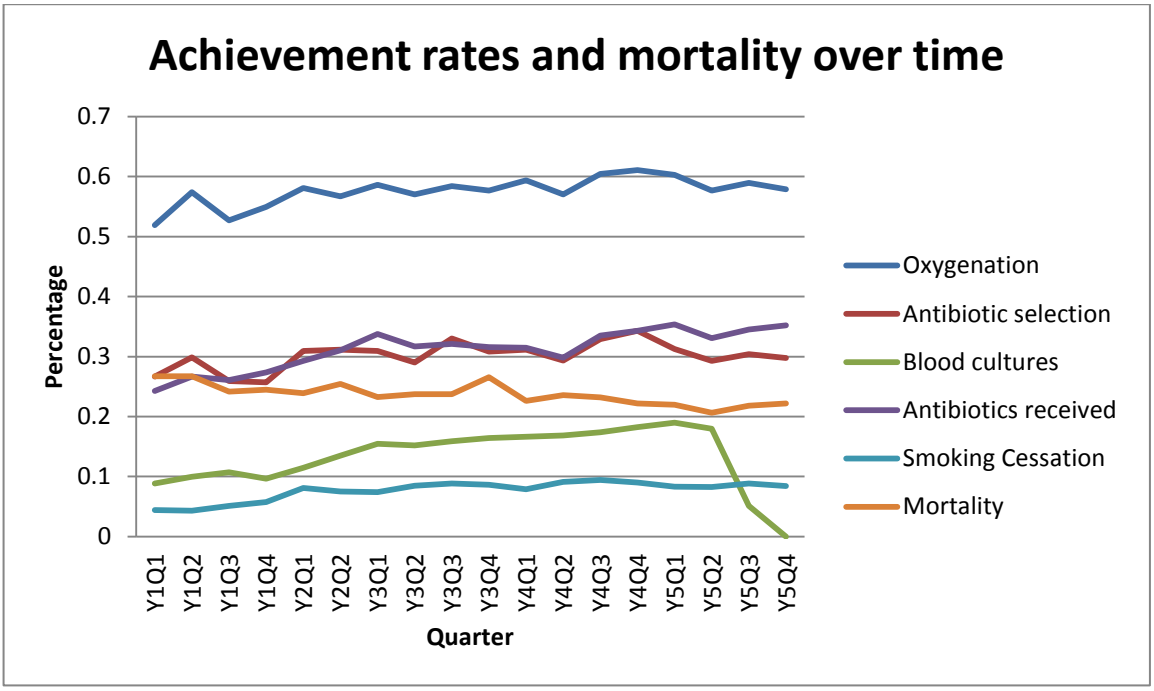

Table 2: Descriptive Statistics

	Individual Level		Trust Level			
	Mean	SD	Mean	SD		
				Overall	Between	Within
Died	0.238	0.426	0.237	0.066	0.047	0.049
Age	72.76	16.481	72.30	3.352	2.63	2.234
Male	0.498	0.500	0.503	0.072	0.022	0.069
Income Deprivation	0.201	0.138	0.190	0.054	0.053	0.015
Ethnicity						
Mixed	0.014	0.116	0.013	0.015	0.013	0.009
Asian	0.018	0.131	0.014	0.024	0.023	0.009
Black	0.004	0.067	0.005	0.011	0.010	0.005
Other	0.007	0.084	0.009	0.025	0.016	0.020
Missing	0.059	0.236	0.081	0.125	0.058	0.112
Achievement						
Oxygenation Assessment	0.602	0.489	0.576	0.173	0.154	0.092
Initial Antibiotic Selection	0.313	0.464	0.301	0.133	0.106	0.084
Blood Cultures	0.135	0.341	0.133	0.103	0.072	0.075
Antibiotics Received < 6 Hours	0.325	0.468	0.312	0.121	0.096	0.077
Smoking Cessation Advice	0.087	0.281	0.077	0.047	0.035	0.031
Exclusions						
Oxygenation Assessment	0.392	0.488	0.418	0.174	0.156	0.090
Initial Antibiotic Selection	0.647	0.478	0.658	0.133	0.113	0.077
Blood Cultures	0.826	0.379	0.827	0.116	0.082	0.085
Antibiotics Received < 6 Hours	0.565	0.496	0.576	0.155	0.133	0.087
Smoking Cessation Advice	0.855	0.352	0.864	0.052	0.044	0.030
Observations	**98771**		**425**			

Table 3: Regression Results

	Trust Level				Individual level			
	Random Effects		Fixed Effects		Probit		OLS	
Age	0.008***	(0.002)	0.007***	(0.002)	0.006***	(0.000)	0.006***	(0.000)
Age Squared	0.000	(0.000)	0.000	(0.000)	-0.000	(0.000)	0.000***	(0.000)
Male	-0.051	(0.034)	-0.035	(0.034)	0.018***	(0.002)	0.018***	(0.002)
Income Deprivation	0.112	(0.100)	0.515**	(0.158)	0.044***	(0.010)	0.040***	(0.010)
Ethnicity								
Mixed	0.166	(0.254)	-0.058	(0.274)	-0.011	(0.011)	-0.005	(0.010)
Asian	-0.109	(0.204)	-0.572*	(0.261)	-0.065***	(0.009)	-0.065***	(0.008)
Black	0.307	(0.408)	0.005	(0.474)	-0.082***	(0.018)	-0.067***	(0.014)
Other	-0.282*	(0.116)	-0.241*	(0.119)	-0.006	(0.016)	-0.008	(0.013)
Missing	0.051*	(0.023)	0.034	(0.024)	0.019***	(0.006)	0.018***	(0.005)
Achievement								
Oxygenation Assessment	-0.271	(0.219)	-0.033	(0.223)	0.343*	(0.141)	0.324*	(0.149)
Antibiotic Selection	0.051	(0.071)	-0.036	(0.075)	0.005	(0.041)	0.005	(0.040)
Blood Cultures	-0.195*	(0.091)	-0.213*	(0.091)	-0.085	(0.055)	-0.077	(0.058)
Antibiotics Received	0.076	(0.062)	0.060	(0.062)	0.087*	(0.039)	0.079*	(0.040)
Smoking Cessation	-0.189*	(0.096)	-0.106	(0.099)	-0.094	(0.069)	-0.123*	(0.062)
Exclusions								
Oxygenation Assessment	-0.166	(0.217)	0.086	(0.220)	0.377**	(0.140)	0.380*	(0.148)
Antibiotic Selection	0.013	(0.083)	-0.081	(0.088)	-0.015	(0.049)	-0.018	(0.049)
Blood Cultures	-0.190*	(0.079)	-0.213**	(0.080)	-0.072	(0.048)	-0.055	(0.051)
Antibiotics Received	0.019	(0.054)	0.043	(0.056)	0.075*	(0.036)	0.066	(0.037)
Smoking Cessation	-0.218*	(0.097)	-0.087	(0.101)	0.458***	(0.066)	0.081	(0.063)
Achievement: trust de-meaned individual level								
Oxygenation Assessment					-0.088***	(0.019)	-0.062***	(0.018)
Antibiotic Selection					0.002	(0.009)	-0.010	(0.005)
Blood Cultures					-0.046***	(0.008)	-0.034***	(0.007)
Antibiotics Received					0.020***	(0.005)	0.017***	(0.004)
Smoking Cessation					0.035	(0.030)	0.023***	(0.003)
Exclusions: trust de-meaned individual level								
Oxygenation Assessment					-0.020**	(0.008)	0.154***	(0.019)
Antibiotic Selection					0.024***	(0.005)	0.030***	(0.006)
Blood Cultures					0.510***	(0.024)	-0.017**	(0.006)
Antibiotics Received					0.006***	(0.000)	0.022***	(0.004)
Smoking Cessation					-0.000	(0.000)	0.165***	(0.003)
Constant	0.189	(0.220)	-0.043	(0.223)			-0.196	(0.138)
Observations	425		425		98771		98771	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects are displayed for probit regressions using margins estimated at means. Models also include quarter dummies and trust fixed effects.

Table 4: Regression Results – Alternative outcome

	In-Hospital Mortality				30 day in-hospital mortality				Risk adjusted 30 day in-hospital mortality			
	Random Effects		Fixed Effects		Random Effects		Fixed Effects		Random Effects		Fixed Effects	
Age	0.011***	(0.002)	0.010***	(0.002)	0.007***	(0.001)	0.004**	(0.001)	-0.001	(0.001)	-0.002	(0.001)
Age Squared	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000*	(0.000)	-0.000	(0.000)	0.000	(0.000)
Sex	-0.063	(0.046)	-0.056	(0.046)	-0.002	(0.035)	-0.004	(0.033)	-0.060	(0.033)	-0.063	(0.033)
Income Deprivation	0.019	(0.116)	0.253	(0.193)	-0.022	(0.098)	0.105	(0.137)	0.074	(0.086)	0.313*	(0.137)
Ethnicity												
Mixed	0.168	(0.281)	-0.060	(0.316)	0.345	(0.219)	0.101	(0.225)	0.176	(0.204)	0.038	(0.224)
Asian	-0.154	(0.229)	-0.631*	(0.293)	-0.385*	(0.183)	-0.667**	(0.208)	-0.387*	(0.167)	-0.511*	(0.207)
Black	0.334	(0.455)	0.001	(0.556)	0.540	(0.360)	0.578	(0.395)	0.343	(0.331)	0.351	(0.394)
Other	-0.220	(0.173)	-0.206	(0.183)	-0.085	(0.132)	0.086	(0.130)	-0.200	(0.125)	-0.140	(0.130)
Missing	0.039	(0.024)	0.025	(0.026)	0.050**	(0.018)	0.039*	(0.018)	0.035*	(0.018)	0.032	(0.018)
Achievement												
Oxygenation Assessment	-0.301	(0.234)	-0.118	(0.246)	-0.113	(0.178)	0.026	(0.174)	-0.065	(0.169)	0.070	(0.174)
Antibiotic Selection	0.050	(0.084)	-0.051	(0.092)	0.010	(0.065)	-0.028	(0.065)	0.018	(0.061)	-0.005	(0.065)
Blood Cultures	-0.191	(0.099)	-0.198	(0.103)	-0.127	(0.075)	-0.113	(0.073)	-0.195**	(0.071)	-0.227**	(0.073)
Antibiotics Received	0.110	(0.070)	0.094	(0.073)	0.101	(0.053)	0.090	(0.052)	0.064	(0.051)	0.058	(0.051)
Smoking Cessation	-0.095	(0.106)	-0.008	(0.113)	-0.179*	(0.081)	-0.155	(0.080)	-0.179*	(0.077)	-0.179*	(0.080)
Exclusions												
Oxygenation Assessment	-0.197	(0.231)	-0.013	(0.242)	-0.071	(0.176)	0.083	(0.172)	-0.087	(0.166)	0.054	(0.171)
Antibiotic Selection	0.033	(0.095)	-0.067	(0.104)	-0.003	(0.073)	-0.040	(0.074)	0.055	(0.069)	0.017	(0.074)
Blood Cultures	-0.117	(0.089)	-0.132	(0.094)	-0.101	(0.068)	-0.086	(0.067)	-0.154*	(0.064)	-0.170*	(0.066)
Antibiotics Received	0.019	(0.059)	0.035	(0.062)	0.039	(0.045)	0.052	(0.044)	0.023	(0.042)	0.043	(0.044)
Smoking Cessation	-0.201	(0.107)	-0.111	(0.116)	-0.267**	(0.083)	-0.165*	(0.082)	-0.185*	(0.078)	-0.181*	(0.082)
Constant	-0.041	(0.243)	-0.180	(0.251)	0.126	(0.184)	0.055	(0.179)	0.625***	(0.175)	0.540**	(0.178)
Observations	333											

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses.