Variations in hospital quality and outcomes under a financial incentive scheme

A thesis submitted to The University of Manchester for the degree of PhD in Medicine in the Faculty of Medical and Human Sciences

2015

Yiu-Shing Lau

The School of Medicine

Institute of Population Health

Manchester Centre for Health Economics

Table of contents

Word count: 71,308

List of tables

List of figures

List of abbreviations

ACEI – Angiotensin Converting Enzyme Inhibitor
ACS – Appropriate Care Score
ARB – Angiotensin Receptive Blocker
CABG – Coronary Artery Bypass Graft
CAP – Community Acquired Pneumonia
CMS – Centres for Medicare and Medicaid Services
CQC – Care Quality Commission
CQUIN - Commissioning for Quality and Innovation
FT – Foundation Trust
HES – Hospital Episode Statistics
HM – Hausman and McFadden test
ICD – International Classification of Diseases
IIA – Independence of Irrelevant Alternatives
IMD – Index of Multiple Deprivation
LSOA – Lower Super Output Area
LVS - Left Ventricle Systolic function
LVSD – Left Ventricular Systolic Dysfunction
NCHOD - National Centre for Health Outcomes Development
NHS – National Health Service
OLS – Ordinary Least Squares
P4P – Pay-for-Performance
PCI - Percutaneous Coronary Intervention
PHQID - Premier Hospital Quality Incentive Demonstration
QMR – Quality Measures Reporter
QOF - Quality and Outcomes Framework
SHA – Strategic Health Authority
SUS – Secondary Uses service
UK – United Kingdom
US – United States of America
VIF – Variance Inflation Factor
VTE - Venous Thromboembolism

Glossary

Advancing Quality	Advancing Quality is a financial incentive scheme covering all NHS hospital Trusts in the North West of England
Area level	Lower-level Super Output Areas, geographical areas containing 1500 people on average.
AQ team	The AQ team is the team of individuals who currently managing the Advancing Quality incentive scheme.
Equity	Equal quality of health care for people with equal needs
Financial incentive	Remuneration awarded to health care providers conditional on meeting criteria set by the policy maker
Process measure	Components of the care provided to patients by providers. This may include investigations or the provision of advice.
Outcome measure	Indicator of key consequences for patients, such as mortality or repeat admission to hospital
Quality Measures Reporter	The Quality Measures Reporter is a dataset which records the process measures of care given to each patient
Spell	The time from when a patient is admitted to a hospital until the patient is discharged

Abstract

High and equitable quality of care are core goals of the English National Health Service. Policy makers have experimented with various ways to improve quality, including use of financial incentives. The effects of these incentives on health outcomes and the distribution of care are not known. The aim of this study was to examine variations in hospital quality and outcomes at patient level under a financial incentive scheme in England.

In October 2008 a financial incentive scheme under which quality of care was measured by process measures was introduced for 24 hospital Trusts in the North West of England. The process measures of care from this Advancing Quality initiative were linked at spell level to health outcomes and administrative hospital records. The data consisted of 252,284 spells between October 2008 and March 2013.

First, I examined whether financially incentivised improvements in quality of care were associated with better patient outcomes. I examined how mortality and readmission were related to process measures using bivariate probit, probit, random effects and fixed effects estimations. I found that several of the incentivised process measures of care are associated with improved patient outcomes. I estimated that Advancing Quality saved 129 lives and avoided 121 readmissions over a four-and-a-half year period.

Second I examined whether quality of care from a hospital incentive scheme is distributed equitably at a patient level. Multinomial and sequential logistic regressions were used to show that process measures of care overall were distributed in favour of patients from lower income score areas. Process measures of care delivered during an emergency admission were distributed in favour of patients from higher income score areas but this was driven by patient severity. Process measures based on advice appeared to be driven by capacity to benefit and were distributed in favour of patients from lower income score areas. Process measures of care for elective admissions regarding delivery of drugs were distributed equitably.

Third, I examined if the quality of care was lower at the weekend. The in-hospital mortality rate is known to be higher for weekend admissions than for weekday admissions but it is not known whether this was due to lower quality of care. Using logistic regressions, incentivised quality of care was found to be consistent throughout the week. The weekend mortality effect can be explained by patient volume, which suggested that patient case mix may be different between weekdays and weekends.

Overall, quality of care under an incentive scheme was found to positively impact on health outcomes, be distributed equitably, and be the same at weekends as weekdays. Further research is needed using quality of care indicators from all Trusts in the English National Health Service. Furthermore further research examining how trusts exclude patients from financial incentive schemes is also needed.

Declaration

No portion of this work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

All chapters were created in collaboration with Matt Sutton who jointly conceived research ideas, recommended econometric methodology and literature, and approved the final version of the thesis. Earlier versions of Chapters 3 and 4 will be incorporated into an article "*Evaluation of the Advancing Quality Pay for Performance Programme in the NHS North West*." This has now been published.

Copyright Statement

Acknowledgements

I would like to thank my supervisors, Matt Sutton, Tim Doran and Stephen Campbell, for their guidance, assistance and ideas throughout the whole process of my PhD. In particular Matt Sutton for his drastically increased: work load; headaches; dare I say it hair loss and grey hairs (the hair that survived); hospital appointments; and alcohol consumption this work has caused him. But especially and most importantly, thank you for sticking by me.

A special thank you goes out to Cheryl Jones who took care of me during the final year of my PhD and especially the write up stage, from: cooking me lunch and dinner; washing my clothes; and putting them on for me.

I would like to thank Søren Rud Kristensen, Alexander Turner and Hannah Forbes for allowing me to stop them working when I wanted so that I could hijack their computers to run my estimation models for this thesis leaving their PC's to operate slower than Windows Internet Explorer. But also most frequently, allowing me to have a not so quick chat.

I like to thank Philip Britteon for recommending me songs albums to listen to whilst I was in the final two months of writing up the thesis. Jack Higgins for his occasional shoulder massages whilst I was slouched over my keyboard staring to the abyss of an empty word document. Luke Munford for his light hearted humour and his kind lunch invites.

To all of the above and all people of Manchester Centre for Health Economics, including Mark Harrison, I would like to thank all of you for your: knowledge; support; advice: food; alcohol; social events and good humour which has made my time doing my PhD seem much shorter and more enjoyable than what it could have been.

I would also like to thank my family: Ching Lau (mum); Yuk-Chow Lau (dad); and Yiu-Kwan Lau (brother).

Dissemination of Work

An earlier version of chapter 3 was presented at: the Zurich European Conference of Health Economics (2012); and the Sheffield Health Economists' Study Group (2014). Both were oral presentations

An early version of chapter 4 was presented at the Health Economists' Study Group in Oxford (2012).

Earlier versions of chapter 5 have been presented at: the Joint conference of the International Health Economics Association and European Conference of Health Economics in Dublin (2014); and the Glasgow Health Economists' Study Group (2014). A recent version of chapter 5 has been presented at the Health Economics Research Centre external seminar series at the University of Oxford (2014). All were oral presentations.

1. Introduction

Quality of care is an important issue for both health care providers and patients. The quality of care is a main determining factor for patients' health outcomes, health management and satisfaction. The English National Health Service (NHS) has made the provision of high quality of care a core mission for current and future generations (Darzi 2008). The ambitious aim for the NHS is to offer world class quality of care, which can compete with other countries around the world (DOH 2010) during times of austerity (Roberts, Marshall, and Charlesworth 2012). The quality of care in the NHS must increase whilst experiencing real term reductions in budget (Roberts, Marshall, and Charlesworth 2012).

The definition of health quality of care is complex and has many aspects. Throughout this thesis I referred to quality of health care as quality of care. With the complex nature of health care, aspects such as provision of care can be assessed through both the perspective of the provider and the recipient of health care (Beattie et al. 2014). The difficulty of defining quality of care is the subjective nature of quality of care (Elwyn et al. 2007). From the patients' perspective, the same quality of care and service can result in different patient perceptions of quality of care (Elwyn et al. 2007).

Currently, there are no universally accepted definitions for quality of care (Campbell, Roland, and Buetow 2000). Campbell, Roland, and Buetow (2000) defined quality as access and effectiveness of care. Access focuses on a patient having a quality of care available which fits a patient's needs. Effectiveness focusses on the positive impact on a patient's health once health care has been accessed (Campbell, Roland, and Buetow 2000).

Maxwell (1984) define quality of care as having six attributes: access to service; relevance; effectiveness; equity; social acceptability; and efficiency and economy. Access to service is having services available to patients need. Relevance focusses on the services of health care which should be suitable based on population need. Effectiveness is defined as care having a direct impact on improved health of patients. Equity is the fairness in the distribution of care. Social acceptability relates to privacy between the healthcare provider and the patient. Efficiency and economy relates to the productivity and costs of care (Maxwell 1984).

Quality of care in England is defined by three main attributes: patient safety; patient experience; and effectiveness of care (Darzi 2008). Patient safety focuses on minimising the risk of patients having adverse health effects when using care. One safety issue is the management of drug prescriptions to reduce mistakes whilst prescribing or administering drugs. Another safety issue is to maintain cleanliness in all healthcare provider premises which minimises the risk of patients acquiring infections within hospital (Darzi 2008). Patient safety was overlooked between 2005 to 2008 in Mid Staffordshire NHS Foundation Trust which resulted in high mortality rates for patients admitted to Accident and Emergency (Francis 2013).

Patient experience measures the quality of care from the patients' perspective. This aspect of quality focuses on whether patients are treated with dignity and respect (Darzi 2008). Patient experience is an aspect of quality of care as it informs healthcare providers on the healthcare services that patients received. This includes: decisions regarding healthcare; support during hospital stay; organisation of care; and continuity of care (Beattie et al. 2014). Improved patient experience has been found to be positively associated with improved patient safety and effectiveness of care (Doyle, Lennox, and Bell 2013).

Effectiveness of care focusses on clinical processes of care having a positive impact on patient health outcomes (Darzi 2008). The care provided should be evidence based and clinical processes should have direct effects of health outcomes (Kohn et al. 2001; Campbell, Roland, and Buetow 2000). Effectiveness of quality of care consists of three areas: avoid premature mortality; providing a better quality of life for people with long term conditions; and helping patients recover from a condition (NHS England 2015). In this thesis, I defined quality of care as the clinical effectiveness of care, where process measures of care should avoid premature mortality or emergency readmissions.

Although quality of care is an important mission of the NHS, little is known about the quality of care that an individual receives. This is because data on the quality of care each patient receives is not routinely collected by all Trusts (Dawson et al. 2005). Instead previous studies on quality has focused mainly on health outcomes such as mortality, which affects 3% of NHS patients after treatment (Dawson et al. 2005). Mortality is therefore poor measure of quality of care due to other factors which may affect mortality which are outside of hospital control, and the low proportion of the population with mortality as the health outcome.

Lord Darzi published a review, entitled High Quality of Care for All (Darzi 2008), which led the movement to focus on quality of care. For up to ten years previous to this publication, the NHS had focussed on increasing hospital capacity, such as hospital beds (DOH 2009). This increase in capacity enabled the focus to now switch towards quality of care (Darzi 2008). Darzi (2008) lead a research team of 2000 clinicians and health care professionals from all over England. This review combined the knowledge from health care professionals and suggested how the quality of care would be improved. The review found that quality of care was poorly measured in the NHS. The authors proposed that having more knowledge on the effects of the care currently provided was essential to improve the quality of care.

Darzi (2008) recommended that the quality of care should be published, so both providers and recipients of healthcare would see the relative quality of care each healthcare provider was delivering. Furthermore, it was recommended a system of incentives should be put in place to incentivise the delivery of high quality of care. This system of incentives would reward providers that provide high quality of care and would be used to support the continual improvement in quality.

Hawkes (2009) commented on the Darzi report by stating the difficulty in measuring the quality of care. If the quality of care should be documented and incentivised, a measurement of quality should be created (Hawkes 2009). The Department of Health aimed to capture quality of care by requesting a group of healthcare professionals to create a list of quality indicators (Hawkes 2009). The author suggested that the NHS would be able to capture process measures of care as a direct indicator for quality of care effectiveness (Hawkes 2009).

Process measures of care as a quality of care measures are increasingly being used in the United Kingdom and countries throughout the world (Bennett 2012; Travaglia and Debono 2009). Process measures of care are indicators of clinical effectiveness of care such as aspirin at arrival or smoking cessation advice. These process measures, which are based on evidence, can be easily documented and reported as a sign of quality (Bennett 2012). Process measures of care may not be used as a sign of quality of care if there is no association between process measures of care and health outcomes (Palmer and Reilly 1979; Bennett 2012).

After the Darzi report, quality incentive schemes aimed at hospitals were introduced in the England. The Advancing Quality scheme was the first hospital quality incentive

scheme introduced in England in October 2008. The incentive scheme incentivised 24 Trusts in the North West of England to provide process measures of care to patients. Under this scheme, Trusts were required to record all process measures of care each patient had received and link these records to spell records, such as the Secondary Uses Service data. Financial incentives were used to incentivise Trusts to provide high quality of care to patients. I explained the Advancing Quality incentive scheme in more detail in Chapter two.

Throughout this thesis I used the words Trust and hospitals interchangeably. A Trust is either a hospital or a collection of hospitals serving a specified geographical area or offering specialised care in the NHS (NHS Choices 2013). I used the term Trust only whenever it is technically correct. Otherwise I used the more popular word, hospitals.

Financial incentives for hospital providers differs from scheme to scheme (Meacock, Kristensen, and Sutton 2014b). Key aspects which are accounted for in the design of incentive schemes are, the size of the financial rewards, who the financial rewards are given to, the patient groups targeted, financial bonus or penalties, how the financial rewards are structured, how the financial rewards are decided and how frequent the financial rewards are offered (Meacock, Kristensen, and Sutton 2014b).

Size of financial rewards under incentive schemes is a consideration in the design of financial incentive schemes. Low financial incentives may not be high enough to act as an incentive. A high financial incentive may result in lower motivation once a target level of reward has been reached. Premier Hospital Quality Incentive Demonstration incentivised Hospitals to earn up to an extra 2% of hospital payments (CMS 2013). The Quality and Outcomes Framework financially incentivised General Practitioners an increase of up to 25% General Practitioner income (Kontopantelis et al. 2015).

Financial incentives also differ in who the financial rewards are given to. Quality and Outcomes framework gave rewards to General Practitioner practices (Kontopantelis et al. 2015), where, the General Practitioners may extract the financial incentive as personal income. Some financial incentive schemes paid financial incentives to Hospitals; this money could not be used as a bonus for individual staff at the hospitals (Ledward, Horne, and Butterworth 2008).

Financial incentives may be given as a bonus or award addition funds. An alternative approach is to pay a financial bonus first, and then request the trusts to pay back the

reward if performance benchmarks were not met. The financial incentive of potentially losing money may have a stronger impact them being awarded a financial incentive. This stems from prospect theory in economics (Kahneman and Tversky 1979). Other types of financial incentives are punishments where bonuses are never issued but fines are issued when performance criteria are not met.

How financial rewards are decided and structured are also considerations for financial incentive schemes. Some structures of financial incentives are performance benchmarks, performance requirements or competition (Meacock, Kristensen, and Sutton 2014b). Performance benchmarks or requirements are set by the policy maker, or agreed between a policy maker and target for the financial incentive such as a hospital. Benchmarks are set to ensure that a certain level of performance must be achieved before the financial incentive would be awarded. Competition is another way to administer financial incentives in the form of a tournament where the financial incentive is awarded only to the best performers. The benefit of a benchmarking system is that if benchmarks have been tailored and set at different levels, poor performers or high performers still have an incentive to increase performance to obtain financial reward. Under a competitive system poor performers who know they would not win may no longer compete to win the financial incentive. Examples of different types of financial incentives aimed at Hospitals in England are given below.

In April 2009, the Commissioning for Quality and Innovation (NHS 2015) framework was introduced. CQUIN was a compulsory scheme where performance targets were agreed between Trusts and healthcare commissioners. A financial penalty was given if Trusts did not agree to meet these targets (Meacock, Kristensen, and Sutton 2014b). As of April 2009, Advancing Quality became a part of CQUIN. Advancing Quality remained a reward based financial incentive scheme.

Best Practice Tariffs were introduced in April 2010 to reduce the variation in the quality of healthcare (DOH 2015). The scheme offered a financial incentive for healthcare providers to provide the most effective healthcare (Meacock, Kristensen, and Sutton 2014b).

In April 2009, Patient Reported Outcome Measures were introduced to capture the effectiveness of care from the patient's perspective (Square 2015). Questionnaires were completed by patients before and after pre-specified types of elective surgery. The four elective surgeries included in the scheme are hip replacements, knee replacements, groin

hernia operations and varicose vein operations (Square 2015). The Patient Reported Outcome Measures questionnaires delivered before and after surgery capture the change in patient perceived health.

This thesis examines the quality of health care, referred to as quality of care, which is financially incentivised through the Advancing Quality initiative aimed at Trusts in the North West region of England. I did not assess the overall quality of care from health care providers, or how the quality of care changed before and after the introduction of the quality incentive scheme. I examined the actual quality of care which is incentivised through the scheme at an individual level.

The aim of this thesis was to examine three aspects of variation in processes and outcomes under a hospital financial incentive scheme. The thesis addressed the following three questions:

1. *Are financially-incentivised improvements in quality of care associated with better patient outcomes?* I aimed to address whether there is a direct effect of the process measures of care from the quality incentive scheme on health outcomes including mortality and readmissions.
2. *"Rich or poor, who gets more"? The distribution of care under a quality improvement program*. I aimed to examine whether the care under a quality incentive scheme was equitable, which is a goal of the NHS. When Trusts improve quality of care to meet targets, does this impact on health care inequity through increased patient selection?
3. *Is the weekend effect on hospital mortality attributable to lower quality of care?* I aimed to test whether the quality of care is driving the weekend effect on mortality, and if not, what are the drivers behind the observed increase in mortality. This aim assessed how the quality of care from a quality incentive scheme is distributed through the days of the week.

In Chapter two I explained the Advancing Quality initiative and the quality of care which was included in the quality incentive scheme. I also explained the data from Advancing Quality and discussed research which has been published on the quality incentive scheme. Chapter three is the first empirical chapter, where I assessed whether process measures of care impact on health outcomes such as mortality and readmissions. Chapter four examined how quality of care is distributed between patients from different income backgrounds. Chapter five examined the weekend effect on the quality of care and aimed

to find the driving factors behind increased weekend mortality. Concluding remarks are included in Chapter six.

2. The Advancing Quality Initiative

The aim of this chapter is to provide information on the Advancing Quality Initiative. Advancing Quality initiative is a hospital level financial incentive scheme which was introduced in October 2008. Under this scheme, hospitals in the North West Strategic Health Authority of the UK were financially incentivised to provide quality of care to patients. Strategic Health Authorities (SHA) were regional organisations within the National Health Service. There were 10 SHAs covering large English regions. An SHA was responsible for managing the performance of healthcare providers, developing health services and improving services within the region. In 2013 SHAs were abolished and replaced by the Trust Development Authority (NHS Choices 2013).

Advancing Quality aimed to provide patients with process measures of care which were meant to improve health outcomes and ensure care was equitable (Ledward, Horne, and Butterworth 2008), where equity is defined as equal level of quality of care for patients with equal need. A Trust's performance within the scheme was measured by the proportion of patients given process measures of care within each year of Advancing Quality. The quality of the recording of process measures of care were audited by the Audit Commission (AQ 2015b).

The Advancing Quality scheme was based on another hospital incentive scheme from the United States (US) called the Premier Hospital Quality Incentive Demonstration (PHQID). This incentive scheme was introduced in 2003, and was an optional scheme for hospitals which operated in the US. Participation in the incentive scheme was voluntary. Not all US-hospitals volunteered to join the incentive programme. Approximately 270 US hospitals adopted the scheme (CMS 2013) representing 5% of all hospitals in the US (Sutton et al. 2012; Lindenauer et al. 2007). Financial incentives were paid to the hospital departments and not to clinicians.

The policy makers who introduced Advancing Quality believed that the quality of healthcare in the North West of England was consistently lower than other regions in England, with large discrepancies in the quality of care between providers within the

North West region (Ledward, Horne, and Butterworth 2008). Advancing Quality was adopted in the North West region to increase and equalise the quality of care in the region, and between the care from North West region with other regions in England, by incentivising the provision of process measures of care and encouraging collaboration between the providers. All 24 hospital Trusts in the North West of England participated in the scheme.

Advancing Quality utilised a 'tournament style' incentive structure (Falk, Fehr, and Huffman 2008). The performance needed to gain a financial reward would be solely based on a competition based on the hospitals relative performance within each clinical area, where hospitals in each clinical area are competing with the same clinical area within another hospital. The incentive structure rewards the hospitals in the top quartile of each clinical area an additional four percent of the healthcare resource group payment to the hospital department. The second quartile was rewarded two percent extra healthcare resource group payment payments. Advancing Quality was a positive financial incentive scheme where no punishments were utilised. The Advancing Quality incentive structure changed since first being introduced in October 2008. From October 2008 until September 2009 the payment structure was purely based on a tournament.

Between October 2009 and March 2010, the scheme also introduced benchmarking as well as a tournament system. Trusts with the largest improvements in the achievement rates of process measures of care were also financially rewarded. Since April 2010 the Advancing Quality scheme became a part of the Commissioning Quality and Innovation payment framework (CQUIN). This changed AQ from a positive tournament based incentive scheme to a locally agreed targets, negative incentive scheme, under which a financial bonus was calculated based on the Trust's expected annual budget, and then deducted if targets were not met.

Advancing Quality initially incentivised five clinical areas: Acute Myocardial Infarction (AMI); heart failure; community acquired pneumonia; coronary artery bypass graft (CABG) and hip and knee replacements (Ledward, Horne, and Butterworth 2008). These clinical conditions were chosen for the following postulated reasons: clinical areas with high volume of patients; high potential for improvements in health; and to reduce costs whilst improving health outcomes (Ledward, Horne, and Butterworth 2008). The Advancing Quality initiative added more conditions to be included in the financial

incentive scheme since the introduction of the scheme in 2008: Dementia; Psychosis; and Stroke (AQ 2015a).

2.1. Process measures of care

Process measures of care are quality of care indicators which are recommended from national care guidelines. All Trusts in the North West of England were incentivised to provide process measures of care to patients. In this section, I described each of the 18 process measure of care, which patient population are incentivised and how hospitals can achieve the process measure. The information used for this subsection was provided by the Advancing Quality incentive scheme in a document which was disseminated along with the data.

A full list of process measures of care are listed in Table 1:

<Insert Table 1>

All process measures of care were observed in my dataset from October 2008 until April 2013, with the exception of Primary PCI which was introduced in October 2009 and Blood cultures which was discontinued from the Advancing Quality scheme in October 2012.

2.1.1. Patient exclusions

In specific circumstances, Trusts in the Advancing Quality program were allowed to exclude patients from each of the process measures. This can benefit both the Trust and patients. The patients will benefit from being excluded from process measures of care if they have an allergy where the process measures of care are not appropriate for the patient (Campbell, Hannon, and Lester 2011). Trusts will benefit from exception reporting as patient factors may determine whether care is appropriate. For example, if a patient does not smoke, then a Trust does not need to provide smoking cessation advice and can exclude the patient. If the patient is not excluded, then Trusts with a higher proportion of smokers will score higher in average achievement rates. Furthermore, if a patient dies within hospital stay, then discharge process measures of care are not required.

Patient safety is the main reason for healthcare providers to exclude patients. This ability for health care providers to exclude patients may also create an opportunity for health care providers to exclude patients who should receive process measures of care. This is known as 'gaming' (Doran, Fullwood, Reeves, et al. 2008; Doran et al. 2012). A study by Doran, Fullwood, Reeves, et al. (2008) examined whether general practitioners were exception reporting patients from a financial incentive scheme. The authors found "little evidence" of gaming (Doran, Fullwood, Reeves, et al. 2008). A study by Gravelle, Sutton, and Ma (2010) also examined whether general practitioners were exception reporting patients from a financial incentive scheme. The authors found that exception reporting varied with supply side health care characteristics, number of general practitioners in practice per patient population and potential competition, which provides evidence for gaming.

2.2. Data

The main data source for all of my studies was the Advancing Quality initiative. From personal communication with Lesley Kitchen of Advancing Quality I obtained data on process measures of care at spell level from 24 Trusts in the North West of England over a four and a half year period starting from the 1st of October 2008 until the 31st March 2013. My estimation datasets were created from the two following datasets, Quality Measures Reporter (AQ 2015) and Hospital Activity dataset (AQ 2015):

2.2.1. Quality Measures Reporter

After the introduction of Advancing Quality, 24 Trusts in the North West of England recorded the process measures of care from the incentive scheme. Data on process measures of care were captured at spell level. There are four to seven process measures of care for each of the five clinical conditions. Each patient spell in the Quality Measures Reporter is repeated by the number of process measures for that clinical condition.

Information on whether or not an individual has received a process measure of care was recorded from hand written medical records from clinicians. More information on how

the data is captured for Advancing Quality is in the appendix. From the October 2009 of Advancing Quality, Trusts were also required to record all exclusions from each process measures of care. Information on the accuracy of the data input and collection of process measures of care for the Quality Measures Reporter is given in Appendix 1.

In the first eight quarters of data, all exclusions were omitted from the Quality Measures Reporter (QMR) dataset. I created observations for all excluded process measures of care to match the data of the final ten quarters of data. Observations for exclusions could be created as each spell should repeat by the number of process measures of care for each clinical condition. If patients do not have any process measures of care data, I assumed that these patients were not included in Advancing Quality.

For each spell, a unique identifier was created by Advancing Quality which is used to link patients' process measures of care data with patients' spell data. The Quality Measures Reporter dataset consists of 252,284 spells for the patients admitted for: AMI (45,380), CABG (10,000), heart failure (34,172); hip and knee replacements (74,455) and pneumonia (88,277) clinical conditions.

2.2.2. Hospital Activity Dataset

The Advancing Quality team, led by Lesley Kitchen in collaboration with the North West Commissioning Support Unit, used Secondary Uses Service (SUS) (HSCIC 2015) data, which is a large administrative dataset at a spell-level. SUS contains patient records of all patients who are admitted to Trusts in England. Using the SUS dataset, patients whose primary clinical conditions was either, AMI, CABG, heart failure; hip and knee replacements and pneumonia, were extracted to create a patient list for each of the 24 trusts within Advancing Quality.

This patient list was circulated by Lesley Kitchen from Advancing Quality and the North West Commissioning Support Unit to each Trust to confirm all patients on the list. Trusts were allowed to omit patients if patients were given an incorrect primary diagnosis in hospital records. By removing a patient from the circulated list, Trusts were removing patients from the Advancing Quality scheme. From quarter nine onwards, Trusts were also allowed to add patients into the Advancing Quality population. Trusts were not

required to provide details on why patients were omitted from the Advancing Quality population.

Advancing Quality then compiled all updated patient lists and created a unique spell level identifier variable to match both the QMR and hospital activity dataset. The Hospital Activity dataset consists of 278,641 spells.

2.2.2.1. International Classification of Diseases (ICD)

International Classification of Disease (ICD) codes version 10 is used by Advancing Quality to identify patient clinical conditions. The ICD-10 codes are used by health care providers and researchers to record and identify clinical conditions (WHO 2015). The ICD-10 codes associated with AMI are I21, I22 and I251. The ICD-10 code associated with CABG is I251. The procedure codes associated with CABG are K402, K403, K451, K453, K454 and K633. The ICD-10 codes associated with heart failure are I500, I501 and I509. The ICD-10 codes associated with hip and knee replacements are M15, M16, M17 and S72. The ICD-10 codes associated with pneumonia are A40, A41, A48, Z16, R65, J96, J13, J14, J15, J16, J17, J18 and J85.

2.2.3. Linking Hospital Activity with Quality Measures Reporter

Patients within the QMR dataset are a subset of patients from the Hospital Activity dataset. Once I linked the two datasets, my dataset contained 252,284 spells. A total of 26,357 spells from the Hospital Activity dataset were not matched to QMR and were subsequently removed. I removed spells without process measures of care as these were used as the main explanatory variables or as the dependent variables in all of my empirical analysis.

From the first eight quarters of data, 23,513 spells were not matched with QMR. 2,844 spells were not matched in the final ten quarters of data. This may be the result of two reasons. The first is that Advancing Quality changed the data management provider from Premier in the first eight quarters, to Clarity Informatics in the final ten quarters of data.

The second reason is that process measures of care have not been previously captured, and therefore there may be more issues with data linkage in the first two years of the scheme.

2.2.3.1. Descriptive statistics

The percentages of unmatched patient hospital records with QMR are, in decreasing order: 15.59% for pneumonia; 11.41% for heart failure; 5.88% for AMI; 3.09% for hip and knee replacements; and 1.13% for CABG. 10,799 spells that did not have matched QMR data ended in the death of the patient. This accounts for 29% of all spells that ended in death. Of the spells that did not end with death, average length of stay for each spell was 2 days longer for unmatched spells. The increased length of stay and mortality rate may be due to a higher average age of 75 compared to 72 for matched spells.

2.3 Limitations of data

The main limitation of the dataset is that the impact of the introduction of Advancing Quality on the quality of care cannot be assessed. All data was from Trusts in the quality incentive scheme only and covers only a time period when the quality incentive scheme has been introduced. The impact of the scheme cannot be assessed as there ass no counterfactual/control group in the dataset.

The impact of the scheme may affect quality of care directly and indirectly. The direct effect will be how the incentive scheme has improved health outcomes for Advancing Quality conditions. The indirect effect will be the spill over effects which the quality incentive scheme would have on clinical conditions not incentivised by Advancing Quality. Another indirect effect will be the effect of Advancing Quality on patients who are not included in the financial incentive scheme, but are admitted with clinical conditions incentivised by Advancing Quality. With the data available, it is only possible to examine direct effects.

2.4 Previous Studies of Advancing Quality

There have been a total of three studies on Advancing Quality. Two studies focussed on the effect of the scheme on mortality (Sutton et al. 2012; Kristensen et al. 2014) and one study focussed on the cost effectiveness of the incentive scheme (Meacock, Kristensen, and Sutton 2014a).

Sutton et al. (2012) assessed the impact of Advancing Quality on 30 day in-hospital mortality. The authors obtained three years of data from Hospital Episodes Statistics from 1st April 2007 until 31st March 2010 which covered eighteen months before and after the introduction of the incentive scheme. The data contained 728,583 patients from 154 Trusts in England. The Advancing Quality clinical conditions the authors used were: AMI; heart failure and pneumonia. The six clinical conditions used as control conditions were: acute renal failure; alcoholic liver disease; intracranial injury; paralytic ileus and intestinal obstruction without hernia; pulmonary embolism; and duodenal ulcer. The authors used a triple difference regression method. The authors found that Advancing Quality resulted in a reduction in mortality by 1.3 percentage points which equates to 890 lives saved. However, this reduction in mortality was only associated with pneumonia and not with AMI and heart failure when statistical significance was measured at a 5% level.

Kristensen et al. (2014) assessed the long term impact of Advancing Quality on 30 day in-hospital mortality. The authors obtained five years of data from Hospital Episode Statistics from 1st April 2007 until 31st March 2012 which covers 18 months pre introduction and 42 months post introduction of the incentive scheme. The data contained 1,825,518 patients from 154 Trusts in England. The Advancing Quality clinical conditions the authors used were: AMI; heart failure and pneumonia. The five clinical conditions used as control conditions were: acute renal failure; alcoholic liver disease; intracranial injury; paralytic ileus and intestinal obstruction without hernia and duodenal ulcer. The authors used a triple difference regression method. The authors found that there were no long term effects of Advancing Quality on 30 day in-hospital mortality when statistical significance was measured at a 5% level.

Meacock, Kristensen, and Sutton (2014) assessed the cost effectiveness of Advancing Quality. The authors obtained three years of data from Hospital Episodes Statistics from 1st April 2007 until 30th April 2010 which covers eighteen months pre and post

introduction of the incentive scheme. An extra month of data was used to create the 30 day readmissions measure. The data contained 856,715 patients from 154 Trusts in England. The patient outcomes were 30 day in-hospital mortality and 30 day emergency readmission. The costs of the quality incentive scheme were obtained from Advancing Quality initiative team, and unit costs were obtained from the Department of Health. The authors found that 649 lives were saved from Advancing Quality and the scheme resulted in 22,802 fewer days in hospital. Calculating the value of the lives saved and the costs of the hospital days avoided, the authors concluded that Advancing Quality was cost effective.

2.4.1. Critique of literature

These three studies of Advancing Quality (Matt Sutton et al. 2012; Kristensen et al. 2014; Meacock, Kristensen, and Sutton 2014) looked at the direct effect and indirect effect on Advancing Quality. However, the reduction in mortality found in pneumonia may be due to either the quality of care which is incentivised or the knock on effects of the quality incentive scheme. This difference cannot be assessed using data without data on process measures of care.

The reason why a direct effect of process measures on mortality cannot be found using HES may be removal of patients which creates problems in identifying which patients admitted with AMI received process measures of care. Trusts were allowed to remove patients from the initial list of patients extracted from SUS. This removal of patients has created a large discrepancy in the number of patients within the HES and linked hospital activity and QMR dataset. Sutton et al. (2012) used a sample of 70,644 patients from the North West of England admitted for AMI, pneumonia and heart failure conditions in the first 18 months of Advancing Quality. The obtained data from Advancing Quality directly contains 32,324 spells from the same clinical conditions, Trusts and time period. Therefore Trusts in the North West excluded 54% of the patient population identified by Sutton et al. (2012). Therefore Advancing Quality only directly affected 46% of the patient sample from Sutton et al. (2012).

The discrepancy in patient population is lower between the Advancing Quality dataset and the dataset used by Kristensen et al. (2014). Kristensen et al. (2014) used a sample of

167,542 patients from Advancing Quality Trusts since the introduction of the quality incentive scheme. Data from Advancing Quality contains 117,656 spells for the same clinical conditions, Trusts and time period. The relative discrepancy may have reduced but, Trusts in Advancing Quality after the ninth quarter were able to introduce patients into the sample. Therefore the patient population of Advancing Quality may not be a subset of the patients from the data used by Kristensen et al. (2014).

2.4.2. Gaps in literature

There are no studies which have examined the direct effect of process measures of care on health outcomes on a patient level. Two goals of introducing Advancing Quality were to improve quality of care and be equitable (Ledward, Horne, and Butterworth 2008). Previous studies have addressed the cost effectiveness of the programme, and the wider impact of Advancing Quality initiative on health outcomes, but no studies have addressed whether the process measures of care have an impact on health outcomes. This issue is important as the care being incentivised should have a direct effect on health outcomes, otherwise, new process measures of care should be considered.

No previous studies have examined how the quality of care is distributed at a patient level. As equity is a goal of Advancing Quality, and the NHS, finding how care is distributed under a hospital quality incentive scheme will contribute to assessing the success of the quality incentive scheme and also decisions of other policy makers and healthcare providers in whether to adopt a similar quality incentive scheme. The following chapter analysed whether process measures of care are associated with improved patient health outcomes.

3. Are financially-incentivised improvements in quality of care associated with better patient outcomes?

Pay-for-Performance (P4P) schemes have been introduced as a means of improving patient health outcomes by incentivising healthcare providers to increase the quality of their care (Christianson, Knutson, and Mazze 2006). The quality of health care is measured typically 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 hospital's control.

I considered a P4P scheme introduced in England, the Advancing Quality initiative, which is a hospital scheme introduced in October 2008 (Ledward, Horne, and Butterworth 2008). Under this scheme, hospitals in the North West 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.

Advancing Quality initially incentivised five clinical areas: acute myocardial infarction (AMI); heart failure (HF); community acquired pneumonia; coronary artery bypass graft (CABG); and hip and knee replacements (hip and knee). For this analysis I examined four of these clinical conditions: AMI; CABG; hip and knee replacement; and pneumonia. I did not analyse heart failure as all process measures of care are related to discharge and therefore, patients who have died in hospital stay will be excluded from all process measures of care associated with heart failure.

A 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. They found that Advancing Quality had resulted in a significantly lower 30-day within-hospital mortality rate among patients admitted for pneumonia. Epstein et al (2014) tested whether care from a similar pay for performance scheme to Advancing Quality in the United States, the Premier Hospital Quality Incentive Demonstration, resulted in worse health outcomes due to deleterious effects of pay for performance. The authors found that pay for performance did not have a negative impact on health outcomes through care substitution.

A follow up study to Sutton et al. (2012) by Kristensen et al. (2014) extended the analysis of Sutton et al. (2012) and tested the long term effects of Advancing Quality using updated data and also using a triple difference design. Long term was defined by a study using three and a half years post introduction of Advancing Quality. The authors found that there was no long term Advancing Quality effect on mortality and the lower mortality rates for the Advancing Quality patients were no longer statistically significantly different to the other hospitals used as the control hospitals.

Although process measures of care are selected based on clinical research (Medicare, 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. Ryan et al. (2009) outlined five reasons why process measures of care may not result in any effect on health outcomes: 1) process measures may be 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) there may be measurement error and gaming by providers.

Finding the relationship between process measures of care and health outcomes is important as improving health outcomes is the goal of the policy maker (Department of Health 2000). As process measures are the focus of the incentives, 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 to, or even correlated with, improvements in patient health outcomes, new measures should be incentivised that have such direct effects.

3.1. Literature review

To find research papers on the effect of quality of care on health outcomes, I used two methods to search for initial papers: Google Scholar and Web of Science. I searched for terms: "mortality"; "readmissions"; "health outcomes"; "quality of care"; "P4P" and "process measures" in various combinations. After finding relevant research papers, I further identified papers by conducting a forward and backward search on citations and the citing papers. I found 19 empirical studies examining a variety of quality of care measures and health outcomes.

I found four papers which used data from outside of the US, and 14 papers using data from the US. I will group the papers from within the US as these studies will be more generalisable to each other due to the same health care system.

3.1.1. Papers from outside the US

Three of the studies examined secondary care (Luthi et al. 2004; Granger et al. 2005; Bray et al. 2013) and one study focused on primary care (Ryan and Doran 2012). I reviewed the secondary care studies first and then reviewed the study on primary care.

All three secondary care studies used patient-level mortality as the health outcome and used multivariate logistic regressions. Luthi et al. (2004) also used the 30 day readmission rate. Granger et al. (2005) studied coronary syndromes across 14 countries using process measures of care aggregated to the hospital level where the patient population was aged 65 and over. Luthi et al. (2004) used one year of data from 1999 from three Swiss health centres and focused on patients with heart failure. Bray et al. (2013) used data from the UK, linking Hospital Episode Statistics to the Stroke Registry.

All three studies found that higher achievements on the process measures were weakly associated with lower mortality rates. Bray et al. (2013) also found that hospitals which provided a higher quality of care were more likely to provide all process measures of care. Luthi et al. (2004) found a weak association between process measures of care with mortality but, did not find an association between readmissions and process measures of care. All three papers adopted cross sectional designs which meant that only associations

rather than causality could be found. Luthi et al. (2004) used patient level data, but suffered from a small study sample of 1,634 patients. Bray et al. (2013) also used patient level data but did not use appropriate panel data methods to analyse the data.

Ryan and Doran (2012) studied the effects of process measures of care on intermediate health outcomes. They used data from 7228 practices from the Quality Management and Analysis dataset from 2004 to 2008 and linked these with the General Medical Statistics dataset data from 2006. The conditions studied were: diabetes; coronary heart disease; stroke; hypertension; and epilepsy. The process measures and intermediate outcomes were aggregated to create composite variables for each of the health conditions. The observation unit was the general practice and the authors used fixed effects due to the longitudinal nature of the data. The authors found that an increase in process measure performance was associated with modest improvements in intermediate outcomes for all conditions studied. A 10 percent increase in the composite process score led to 3.16, 4.32, 7.60, 7.24 and 7.16 percentage point increases in the composite outcome scores for diabetes, coronary heart disease, stroke, epilepsy and hypertension respectively. All of these effects were statistically significant at the 1% level.

A limitation of the study by Ryan and Doran (2012), which was mentioned by the authors, was that the study is only at practice level. Due to the ecological fallacy (Finney et al. 2011), this relationship between process measures and intermediate outcomes may not be translated to a patient level. Furthermore, these findings on a primary care setting may not be generalisable to a secondary care setting. Process measures in primary care are more long term due to the focus on preventative care, in contrast to the process measures of care in secondary care, which focus on shorter-term management of conditions.

3.1.2. Papers from the US

Fourteen papers on the associations between process measures of care and health outcomes are from the US (Kontos et al. 2014; Bradley et al. 2006; Wu et al. 2014; Fonarow et al. 2007; Chung et al. 2008; Luthi et al. 2003; Peterson et al. 2006; Lee et al. 2011; Lee et al. 2014; Sucov, Valente, and Reinert 2013; Jha et al. 2007; Werner and Bradlow 2006; Werner, Bradlow, and Asch 2008; Ryan et al. 2009). Together the studies

examined three health conditions: AMI; HF and PN. I review the findings on each of the health conditions separately.

3.1.2.1. AMI

Two studies from the US focused solely on AMI (Bradley et al. 2006; Kontos et al. 2014). The papers analysed the association of process measures on health outcomes in two different ways. Bradley et al. (2006) used one year of hospital level data from the National Registry of Myocardial Infarction from 2002 to 2003. The authors used a 30 day mortality rate as the outcome variable and used a hierarchical generalised linear model design. The authors found that process measures were correlated with mortality, however the process measures accounted for a small amount of the variation in mortality. Kontos et al. (2014) used patient level data from Medicare and Medicaid services from 2007 to 2011. The authors used in-hospital mortality as the outcome and hierarchical logistic regression. In contrast to Bradley et al (2006), they found that patient mortality was not correlated with process measures of care.

Both of these papers suffer from limitations. The hospital level dataset available to Bradley et al. (2006) restricted the study to using cross sectional methods so only associations could be found. The hospital level data means results may not be generalisable to individual level due to the ecological fallacy. Kontos et al. (2014) did not have direct process measures of care linked with mortality on an individual level. Therefore, there may not have been sufficient variation in the hospital-level process measures of care to explain individual-level in-hospital mortality.

3.1.2.2. Heart failure

Four studies focused on heart failure as the only clinical condition (Wu et al. 2014; Fonarow et al. 2007; Chung et al. 2008; Luthi et al. 2003). Wu et al. (2014) used 30 day and one year mortality rates as the outcome variables. Fonarow et al. (2007) and Chung et al. (2008) examined readmission rates, as well as 60-90 day, and short term, mortality respectively. Luthi et al. (2003) used 21-month readmission rates after discharge as the

sole outcome variable. The data was sourced from Medicare and Medicaid services (Luthi et al. 2003; Chung et al. 2008; Wu et al. 2014) and a quality improvement program OPTIMIZE – HF program (Fonarow et al. 2007).

The varieties of methods used by the studies are: Poisson regression (Luthi et al. 2003); multivariable regression (Fonarow et al. 2007); survival analysis (Chung et al. 2008) and logistic regression (Wu et al. 2014).

Two studies found that the prescription of Angiotensin Converting Enzyme Inhibitors (ACEIs) at discharge resulted in lower numbers of readmissions (Luthi et al. 2003) and was associated with lower probability of 30-day and one-year mortality (Wu et al. 2014). The Fonarow et al. (2007) study also found that the prescription of ACEI at discharge was correlated with a reduction in both mortality and readmission rates. However, this relationship was not statistically significant once patient characteristics were adjusted for. Chung et al. (2008) used a composite score to account for all quality measures of care and found that patients who were given all quality metrics were less likely to be readmitted.

These four papers suggested that the relationships between heart failure process measures and health outcomes were weak. Only the prescription of ACEI looked likely to be linked with lower readmissions or mortality, and this association may not be strong. These heart failure studies had key limitations. The studies had small sample sizes: 2943 patients in the (Luthi et al. 2003) study and 400 patients in the (Chung et al. 2008) study. This may limit the generalisability of the results. In addition, the short study period of 18 months meant that only cross sectional analysis could be conducted resulting in correlations in one specific year (Fonarow et al. 2007).

3.1.2.3. AMI & Heart Failure

Peterson et al. (2006) studied the relationship between following clinical guidelines and mortality. Using data from a national quality improvement initiative from 2001 to 2003, the authors used Pearson correlation coefficients to assess the correlation between adherence to clinical guidelines (which include process measures of care) and mortality. The authors found that adherence to clinical guidelines were associated with a reduction in mortality. The main limitation with the methods that the authors used is the possible confounders which may affect the correlation. These explain factors may influence who

one variable is affected with another, if ignored, this leads to an omitted variable bias (Wooldridge 2013)

3.1.2.4. Pneumonia

Three studies have looked at pneumonia as the sole clinical condition (Lee et al. 2011; Lee et al. 2014; Sucov, Valente, and Reinert 2013). All three studies used data from Medicare and Medicaid services. Lee et al. (2011) obtained one year of data from 2001, (Lee et al. 2014) obtained data from 2006-2010, and Sucov, Valente, and Reinert (2013) obtained data from one emergency department where no date was stated. Lee et al. (2014) and Lee et al. (2011) used all-cause 30-day mortality and readmission within 30-days as outcome variables and used multilevel logistic regressions to test for associations between process measures and health outcomes. Sucov, Valente, and Reinert (2013) used in-hospital mortality and used Wilcoxon and Chi-squared tests to measures the associations between process measures and the health outcome. Lee et al. (2011) and Sucov, Valente, and Reinert (2013) both found no associations between process measures of care and health outcomes. (Lee et al. 2014) found that provision of process measures for pneumonia reduced the probability of both mortality and readmissions.

Each of these three studies had limitations regarding study design, sample selection and weak study methods. (Lee et al. 2011) were not able to test whether a specific process measure of care impacted on the health outcome, only the number of process measures of care received. These will differ between patients due to severity and need, which were not controlled for. Sucov, Valente, and Reinert (2013) had a selective study sample of patients from one hospital emergency department which will lack generalisability. Lee et al. (2014) selected a study sample which was patients who were not excluded from any of the process measures of care. This patient selection would be restrictive and may introduce bias to the results as the selected patients would all have required similar care.

3.1.2.4. AMI, heart failure & pneumonia

Four papers have studied the relationship between patient mortality and process measures of care for AMI, heart failure, and pneumonia (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 four studies found that higher levels of achievement on process measures were associated with lower mortality rates.

These four studies had the following limitations. Jha et al. (2007) used a provider level mortality rather than individual level mortality, and Ryan et al. (2009) used aggregated process measures of care rather than individual level of care. The problem with using aggregated data is the potential ecological fallacy of results. The two studies by Werner and Bradlow (2006) and Werner, Bradlow, and Asch (2008) used cross sectional methods and therefore were unable to capture an causality.

3.2 Aim

The aim of the study was to analyse whether the short-term 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, I examined 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.

3.2.1. Potential Contribution of this study

The results from the literature suggest that there is either no or, at best, a weak relationship between health outcomes and process measures. However the studies in this review have three main limitations:

1. The use of provider level data. The ecological fallacy may limit interpretation of the relationship between process measures of care and health outcomes at an individual level.
2. Aggregation of process measures of care, so the relationship between each of the process measures of care and health outcomes cannot be found. The weak relationship found may be due to certain process measures driving the results, and the effects are diluted.
3. Short study periods and cross-sectional designs meaning that correlations cannot be interpreted as causal effects.

This study addressed the limitations of the previous literature, by contributing in three ways.

1. I examined the link between process measures and mortality at patient level with unique data from a quality improvement program which incentivised providers to provide process measures of care and record data at patient level. This allowed us to perform both provider level and patient level analysis. Furthermore I had data on each of the process measures of care that each patient has received and therefore I was able to see which process measures are associated with health outcomes. I also created a composite care score called the appropriate care score, which was whether patients received all of the process measures of care that they were not excluded from.
2. Our study sample covers 18 quarters which means I was able to use both cross sectional and panel data methods in separate regressions.
3. Our study will be the first to my knowledge which used clinical conditions under elective surgery; CABG and hip and knee replacements. The Trusts that participated in Advancing Quality were incentivised to provide process measures of care of patients with these elective surgery clinical conditions, yet no research has been conducted on the effect of incentivising these process measures.

3.2.1.1. Endogeneity

When estimating the receipt of process measures of care on health outcomes such as in-hospital mortality, there is a potential endogeneity problem. Previous studies have not used spell level data and, as such, did not need to account for the endogeneity between the process measures and the health outcomes at patient level.

The achievement of process measures for the patient may be determined by the expected health gain that a Trust thinks a patient will receive from a process measure of care. Trusts might select patients based on patient's level of health, which would generate an unobserved health of patient effect of health outcomes on process measures of care. Patients in low health with a high probability of mortality may not receive process measures of care as the expected benefit will be minimal to the patient. On the other hand patients in good health may not receive process measures of care with the same reason being a low potential health benefit from the process measure can be achieved. I could not control for patient severity or expected health gain as these are not measured in administrative datasets. When under an emergency, trusts may not select patients; however, Trusts can choose which patients to be admitted. If Trusts choose to provide process measures to more patients in good health than in bad health, then, the results I found will be biased to having a larger relationship than what is true, if I did not control for this endogeneity.

The probability that a patient would receive process measures of care is directly related to how quickly the Trust to which they were admitted responded to Advancing Quality. A Trust's speed and size of response to Advancing Quality should not directly affect patients' health outcomes other than through the delivery of process measures. I proposed that differences in Trust speed and size of responses to Advancing Quality can be considered to be exogenous to individual patients, and that these provide an exogenous source of variation in process measures that can be used to examine the causal relationship between process measures and health outcomes.

3.3. Methods

I performed both cross sectional and panel data econometric techniques to estimate the association between the process measures and the health outcomes. I expected that the process measures of care are inversely related to the bad health outcomes. All analysis was conducted using Stata 13 version MP4

3.3.1. Trust level analysis

For my Trust level analysis, I aggregated all spell-level variables by combinations of Trust and quarter. I then estimated models which contained either fixed effects or random effects for Trusts (Wooldridge 2013). For the Trust level analysis, I treated each trust as a unit of observation. As each trust was a unit of observation, these results do not directly relate to an individual level analysis due to differences occurring through aggregation.

The general model is given by:

$$\boldsymbol{y_{jt} = \alpha + \beta X_{jt} + \gamma p_{jt} + c_t + a_j + u_{jt}} \qquad \text{Equation (1)}$$

In equation (1): The *j* denotes Trust, and *t* denotes time in quarters. α is the constant term. y_{jt} are the health outcomes; α is the constant term; $\boldsymbol{X_{jt}}$ are 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); $\boldsymbol{\gamma}$ are a vector of coefficients on mean characteristics of patients; and c_t are time fixed effects. The error component is in two parts where a_j is a random variable 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 error component, a_j, in the fixed effects estimator is assumed to be time constant and therefore removed under the fixed effects estimation.

3.3.1.1. Variance Inflation Factor

Due to the low number of observations, ranging from 76 to 426, in the Trust-quarter level analyses, I estimated Variance Inflation Factors (VIF) to test for the extent of multicollinearity between the process measures of care (O'Brien 2007). I removed process measures of care where the VIF was greater than 100. Multicollinearity occurs when variables are highly correlated with each other (Wooldridge 2013). This may result in the inflation of standard errors of the estimated model. I identified which process measures of care were highly correlated, using the VIF, and remove them from the Trust-quarter level estimations.

In total, I removed nine process measures of care. From AMI analysis, I removed the five process measures from seven process measures: aspirin prescribed at discharge; Angiotensin Converting Enzyme Inhibitor (ACEI) or Angiotensin Receptor Blockers (ARB) for Left Ventricular Systolic Dysfunction (LVSD); smoking cessation advice; beta blocker at discharge and primary Percutaneous Coronary Intervention (PCI). From the CABG analysis, I removed two process measures from four: aspirin at discharge; and antibiotic selection. From the pneumonia analysis, I removed two process measures from five: oxygenation assessment; and smoking cessation advice.

3.3.2. Individual level analysis

The spell level analysis was conducted using three sets of estimations. The first set used spell level outcomes and Trust-quarter level hospital quality. I used probit regression for this analysis (Wooldridge 2013). The second set of estimations used spell level outcomes and spell level process measures of care, estimated using a probit regression. The third set of estimations used spell level outcomes and spell level appropriate care score using a bivariate probit regression (Wooldridge 2010).

I estimated three sets of regressions, where each will provide an improved estimate of process measures of care on patient outcomes. Regression set one reproduced what has been achieved in the published literature where process measures of care are aggregated to a trust quarter level where the level of observation is at the trust level. Regression set two improved on the literature by estimating the association between mortality and health outcomes on an individual level. Regression set three addressed the endogeneity problem when estimating regression set two.

3.3.2.1. Regression set one

In regression set one I analysed the following model for each of the clinical conditions:

$$\boldsymbol{y_{ijt} = \alpha + \beta \overline{X}_{jt} + \gamma p_{ijt} + \delta h_j + \theta c_t + e_{ijt}} \qquad \text{Equation (2)}$$

In equation (2): The *i* denotes individual; *j* denotes Trust and *t* denotes time in quarters. y_{ijt} is health outcomes; α is the constant term; $\overline{\boldsymbol{X}}_{\boldsymbol{jt}}$ are mean values of all process measures of care, including both achievement and exclusion rates, at the Trust-quarter level; $\boldsymbol{\beta}$ are the coefficients of interest; $\boldsymbol{p_{ijt}}$ is a vector of patient characteristics (age, sex, ethnicity and income deprivation score); γ is a vector of coefficients on patient characteristics; h_j are hospital fixed effects; δ are coefficients on hospital fixed effects; c_t are time fixed effects; θ are coefficients on 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 also estimate this regression using a Trust-quarter level appropriate care score.

3.3.2.2. Regression set two

In regression set two I analysed the following model for each of the clinical condition:

$$\boldsymbol{y_{ijt} = \alpha + \lambda X_{ijt} + \xi p_{ijt} + \pi h_j + \varpi c_t + e_{ijt}} \qquad \text{Equation (3)}$$

In equation (3): The *i* denotes individual; *j* denotes Trust and *t* denotes time in quarters. y_{ijt} is health outcomes; α is the constant term; $\boldsymbol{X_{ijt}}$ is a vector of individual-level process measures of care or appropriate care score.; λ are the coefficients of interest; $\boldsymbol{p_{ijt}}$ is a vector of patient characteristics (age, sex, ethnicity and income deprivation score); ξ are a vector coefficients for patient characteristics; h_j are hospital fixed effects; π are coefficients on hospital fixed effects; c_t are time fixed effects; ϖ are coefficients on the 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.

3.3.2.3. Regression set three

The number of potentially endogenous variables created a substantial estimation problem. For AMI, for example, I would have 14 endogenous variables. Due to this, I am limited to using an appropriate care score to capture the quality of care patients' received.

In regression set three, I estimated the association between the appropriate care score and health outcomes for each clinical condition. As the receipt of quality may be endogenous, I estimated the following models, using a recursive bivariate probit model (Maddala and Lee 1976):

$$\boldsymbol{y_{ijt} = \beta_1 ACS_{ijt} + \beta p_{ijt} + h_j + c_t + e_{1ijt}} \quad \text{Equation (4)}$$

$$\boldsymbol{ACS_{ijt} = \gamma p_{ijt} + hc_{jt} + h_j + c_t + e_{2ijt}} \quad \text{Equation (5)}$$

In equations (4) and (5): The *i* denotes individual; *j* denotes Trust and *t* denotes time. ACS_{ijt} is the individual appropriate care score. β_1 ss the estimated coefficient on ACS. p_{ijt} is a vector of patient characteristics. β are a vector of coefficients on the patient characteristics. γ are a vector of coefficients on patient characteristics. h_j are hospital fixed effects; c_t are time fixed effects. The error terms, e_{1ijt} and e_{2ijt} are identically and independently distributed with a bivariate normal distribution with mean of zero, variance of one and a correlation of ρ. In equation (4) y_{ijt} are the outcome variables, morality or readmission; and β_1 is the coefficient of interest.

A bivariate probit regression model estimates equation (4) and equation (5) simultaneously which allows for the errors of both models to be correlated. This means that the endogeneity though simultaneity will be adjusted for (Heckman 1978) as ACS_{ijt} will no longer be correlated with e_{1ijt}. If I assumed that a common variable such as patients health status, which is unobserved, is correlated with both health outcomes and ACS_{ijt}, then I would have an endogeneity problem. Modelling equation (4) and equation (5) will mean that health status will be present in the error terms, e_{1ijt} and e_{2ijt} respectively. Correlation between e_{1ijt} and e_{2ijt} will be adjusted for in the bivariate probit estimation, hence adjusting for the common unexplained variables such as health status. I included a restriction criterion which are the Trust-quarter interaction terms into equation (2) to improve identification of the model (Jones 2007) as including hc_{jt} into equation (2) will explain ACS_{ijt} reducing e_{2ijt}. As the bivariate probit model implies that the error terms are distributed bivariate normal, controlling for more unobserved heterogeneity within the model strengthens the assumption on the error terms (Jones 2007).

This model can be estimated without an exclusion restriction. However, I expected that a Trust's decision to provide patients with all appropriate process measures of care will be determined by how quickly a Trust responds to the financial incentive scheme. How quickly a Trust responds to Advancing Quality should affect a patient's health outcome only through the care which is provided.

3.3.2.4. Marginal effects

After running the regressions using probit and logistic methods, I estimated average marginal effects to aid interpretability of the results (Bartus 2005). To obtain average marginal effects, marginal effects will first be calculated for each covariate for each patient. The marginal effects calculated for each patient are then averaged for each variable.

3.3.3. Calculation of the impact of Advancing Quality on health outcomes

I calculated the potential impact of Advancing Quality on health outcomes by estimating the number of lives saved from AMI, CABG and pneumonia, and the number of readmissions avoided for hip and knee patients. I calculate this using equation (6):

$$\sum_{t=2}^{T}\sum_{j=1}^{J}\left(\boldsymbol{ACS_{jt}} - \boldsymbol{ACS_{j1}}\right) \times \boldsymbol{\beta_{ACS}} \times \boldsymbol{N_{jt}} \qquad \text{Equation (6)}$$

In equation (6), *t* denotes time period in quarters; *j* denotes Trust; *ACS* denotes the Trust average appropriate care score; β_{ACS} is the coefficient from either the fixed effects or random effects estimation for the appropriate care score and N_{jt} is the number of patients within each Trust and time period.

In this equation I first calculated the additional proportion of patients receiving appropriate care over the base value of this proportion in this Trust in the first quarter. I then multiplied this proportion by the number of patients affected and then estimated effect of the appropriate care score on the outcome.

When estimating equation (6) I assumed that all of the increase in the appropriate care score over the value recorded in the first quarter is attributable to the existence of the Advancing Quality scheme. This assumption includes holding the quality of care constant over the observation period for Trusts not included in Advancing Quality, which includes not technological improvement or advancements in health care over time. The effect of Advancing Quality without this assumption was not possible. This is because we did not have data on process measures of care before the introduction of Advancing Quality, we also did not have process measures of care data on Trusts not within Advancing Quality.

3.3.4. Data

I obtained data from the Advancing Quality programme (AQ 2015). These included hospital records from the Secondary Uses Service (SUS) (HSCIC 2015) which contained patient and Trust-level characteristics at spell level and data from the programme's Quality Measures Reporter (QMR) (AQ 2015), which records delivery of the process measures of care for each patient. The linked data contained around 252,284 patient spells for patients admitted with AMI, CABG, hip and knee and pneumonia across 18 quarters, October 2008 until April 2013. This encompassed almost the entire population of AMI, CABG, hip and knee and pneumonia patients qualifying for inclusion in the Advancing Quality programme from the 24 Trusts in the North West of England, more information on data is found in chapter 2, section 2.2.

The introduction of Advancing Quality meant that Trusts who were a part of the scheme in the North West of England had to record process measures of care given to patients. 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, October 2008 until January 2009, and one Trust missing two quarters of data, October 2008 until April 2009. One other Trust had missing data for the final year of observation, April 2012 until April 2013.

A list of exclusion reasons for each process measure of care is shown in Table 2. Patients are excluded from discharge measures if they are discharged dead. I did not assessed the impact of process measures on care on health outcomes on heart failure as the process measures under Advancing Quality scheme are all discharge measures. This meant that

for the heart failure condition, all patients who have died in-hospital were excluded from all process measures of care.

<Insert Table 2>

3.3.4.1. Outcome variables

I generated a dichotomous variable for in-hospital mortality for each spell as the outcome variable for AMI, CABG and pneumonia patients. For hip and knee patients, I used a 30 day readmission rate as only 2.11% of patients undergoing hip and knee replacement died in hospital. The 30 day readmission dichotomous variable was generated by the data team at Advancing Quality; it captured readmissions for any cause up to 30 days after surgery. The in-hospital mortality variable was generated using the discharge method field in the hospital record and takes a value of one when the patient was discharged dead and zero otherwise. If the discharge method was missing, I did not use that data.

The in-hospital mortality variable is limited in that a patient's health state prior to hospital admission will affect mortality regardless of what process measures of care are given to patients. In principle, mortality either in or out of hospital would be a preferred measure. However, linked information on out of hospital deaths is not available; furthermore 80% of deaths for AMI and pneumonia occur in hospital 30 days after hospital arrival (Kristensen et al. 2014). For the Trust-quarter level analysis, I aggregated the dichotomous variables and readmission variables to Trust-quarter level.

To extend my analysis for the emergency care conditions, AMI and pneumonia, I used two more outcome variables at the Trust level. These were an unadjusted in-hospital 30-day mortality rate and a risk-adjusted in-hospital 30-day mortality rate. These were obtained from Kristensen et al (2014) and were originally sourced from Hospital Episode Statistics (HES) (HSCIC 2012). They were provided at Trust-quarter level over 14 quarters, October 2008 until April 2012. These data encompassed all patients admitted with pneumonia and AMI from the 24 Trusts in the North West of England. The risk adjusted mortality rate controlled for primary diagnosis using International Classification of Disease version 10 codes (ICD10), comorbidities using Elixhauser conditions (Southern 2004), admission source, admission method, age and sex.

3.3.4.2. Process measures of care

A full list of process measures of care are shown in Table 1. For each of the process measures of care the datasets included information on whether a patient had: received a process measure; failed to be given a process measure; or been excluded from the measure. As a result of having three independent categories, each of these process measures were accounted for through two dichotomous variables: a dichotomous variable which takes a value of one if a patient has been given the process measure and zero otherwise; and a second dichotomous variable which takes the value of one if the patient was excluded and zero otherwise. By using two variables for each process measure of care, each pair will be compared with patients who failed to be given a process measure of care. For analysis on the Trust level, I aggregated the dichotomous variables for each of the conditions for both achievement and exclusion, by each Trust and observational quarter. For spell level analysis I used both individual and Trust-quarter level process measures of care.

Using the set of dichotomous individual level variables for each process measure, I was able to generate an appropriate care score. The appropriate care score was a dichotomous variable that is equal to one if a patient has been given all of the process measures of care from which they have not been excluded, and zero otherwise. This dichotomous variable indicated whether a patient has received all the quality of care which has been incentivised under the Advancing Quality scheme.

3.3.4.3. Other covariates

I obtained data from two other sources for the other covariates. I accounted for patients' area-level income deprivation scores from the Index of Multiple Deprivation (IMD. 2013) and time-varying Foundation Trust status sourced from the Monitor website (Monitor. 2014).

The income deprivation score was measured at a lower super output 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.

3.4. Results

3.4.1. Descriptive statistics

Descriptive statistics are shown in Table 3. I found that 9%, 2% and 24% of patients who were admitted with AMI, CABG and pneumonia died in hospital, respectively. For patients who were admitted with hip and knee, 16% were readmitted within 30 days. The mortality rate for CABG is lower than the mortality for hip and knee replacements. I do not have readmission rates for patients admitted with CABG.

<Insert Table 3>

The average age of the admitted population was 72 years and there were marginally more males than females at 50.3% of the sample.

The achievement rates shown in Table 3 were much lower than the scores used by the Advancing Quality programme website, http://www.advancingqualitynw.nhs.uk/results-by-hospital/, as I included process measures of care that are excluded in the denominator. I found that more patients on average are given a process measure than are excluded from them. On average across all process measures: 55% of total process measures are achieved; 3% of total process measures are failed; and 42% of all process measures were recorded as exclusions.

3.4.1.1. AMI

Aspirin given at arrival was the process measure with the highest achievement percentage for AMI patients at 69%. Aspirin given at discharge and beta blockers prescribed at discharge have achievement percentage of around 55%. Angiotensin Converting Enzyme Inhibitor (ACEI) or Angiotensin Receptor Blockers (ARB) for Left Ventricular Systolic Dysfunction (LVSD), smoking cessation advice, fibrinolytic therapy and primary Percutaneous Coronary Intervention (PCI) have high levels of exclusions of over 80%, with the latter two indicators having over 90% exclusions.

3.4.1.2. CABG

Achievement rates for all process measures of care for CABG were over 90% with the exception of discontinuation of antibiotics which has an achievement rate of over 80%.

3.4.1.3. Hip and knee replacement

Achievement rates for all process measures for hip and knee replacement were all over 80%. Exclusion rates for all indicators for hip and knee were less than 10%.

3.4.1.4. Pneumonia

Oxygenation assessment has the highest level of achievement, with 60% of patients 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 were the two quality measures with the highest exclusion rates of around 80%. Smoking cessation has a high exclusion rate as non-smokers were excluded from this measure.

3.4.2. Achievements over time

Figure 1 shows the mean achievement percentage of process measures for AMI over time from October 2008 until and including March 2013 by each quarter. I found that achievement rates for aspirin at discharge and primary PCI have increased by 8 percentage points between the 1st October 2008 and 31st March 2013. The achievement rates of beta blockers on arrival have increased by 13 percentage points. Achievement rates of aspirin at arrival, ACEI or ARB for LVSD, and fibrinolytic therapy have lowered by 2, 0.7 and 9 percentage points.

<Insert Figure 1>

Figure 2 shows that achievement rates for all process measures of care for CABG have improved over time. Achievement rates of aspirin at discharge exhibited the lowest improvement of 0.2 percentage points. Achievement rates of antibiotics received and antibiotics discontinued increased by 7 percentage points and Achievement rates of antibiotic selection have increased by 3 percentage points.

<Insert Figure 2>

Figure 3 shows that all process measures of care for hip and knee patients follow similar trends. All exhibiting an increasing non-linear concave shape. The variation in achievement rates of the process measures also reduced overtime. Achievement rates of all process measures of care increased over 20 percentage points with the exception of antibiotics discontinued which increased by 17 percentage points.

<Insert Figure 3>

Figure 4 shows the mean achievement rates of process measures over time for patients admitted with pneumonia. The trends showed that achievement rates have increased steadily over time. Achievement rates of antibiotics received exhibited the largest percentage point increase of 11 percentage points. I found that oxygenation assessment, antibiotic selection and smoking advice experienced an increase of: 4.9; 3.6; and 3.9 percentage point increase respectively. Achievement rates of blood cultures reduced in the final two quarters as the provision of blood cultures was no longer financially incentivised from October 2012.

<Insert Figure 4>

Figure 5 shows the mortality rates for AMI, CABG and pneumonia and the readmission rate for hip and knee patients over time. I found that the mortality rate for patients admitted with CABG has increased by 0.5 percentage points. Mortality rates for AMI and pneumonia have decreased by 2.9 and 4.6 percentage points respectively. Readmission rates for hip and knee patients have fallen by 11.1 percentage points. The rate of readmissions is lower during year two quarter two and year five quarter four. The rate of readmission after 30 days will be lower for year five quarter four as no data was available after that quarter. The rate of readmission after 30 days may be lower for year two quarter

two as the change in data provider may have resulted in the loss of readmission data. I included quarter variables to address this issue.

<Insert Figure 5>

3.4.3. Trust level analysis

3.4.3.1. AMI

Results from Trust level analysis for patients admitted for AMI are shown in Table 4. I found that both regression methods generated similar results. The results of the Hausman tests suggest that the random effects estimator cannot be rejected; therefore I used the random effects estimator. I found that a one percentage point increase in the provision of aspirin at arrival lowers mortality by around 0.35 percentage points. This effect is statistically significant at the 0.1% level

<Insert Table 4>

Higher percentage of patients excluded from aspirin at arrival and primary PCI were associated with lower mortality. A one percent increase in patients excluded from aspirin at arrival lowers mortality by 0.3 percentage points. This is statistically significant at the 0.1% level. Exclusions from primary PCI lowers mortality by 0.06 percentage points. This is statistically significant at the 5% level.

The Trust-quarter level results for the appropriate care score for patients with AMI are shown in Table 5. I found that a one percentage point increase in the appropriate care score lowers mortality by 0.28 percentage points.

<Insert Table 5>

3.4.3.2. CABG

The results for CABG are shown in Table 6. I found that no process measures of care have a statistically significant association with a change in mortality. Achievement of

antibiotic received has a small positive effect on mortality, where a one percentage point increase increases mortality by 0.0004 percentage points, but not statistically significant.

<Insert Table 6>

The relationship between the appropriate care score and mortality is shown in Table 7. Similar to the results of models containing each process measure of care, Trusts that have a higher rate of providing all appropriate process measures of care do not have a statistically significant effect on mortality. The effect size is also small. A one percentage point increase in appropriate care score, increases mortality rate by 0.04 percentage points.

<Insert Table 7>

3.4.3.3. Hip and knee

The results for hip and knee Trust-quarter level regressions are shown in Table 8. Like the findings from CABG, another elective condition, no process measures of care were statistically significant. All process measures of care with the exception of prophylactics received exhibits the hypothesis that an increase in the provision of process measures of care leads to a decrease in negative health outcomes.

<Insert Table 8>

The relationship between the appropriate care score and mortality is shown in Table 9. A one percentage point increase in the appropriate care score reduces the readmission rate by 0.15% points, where the results were statistically significant at the 0.1% level.

<Insert Table 9>

3.4.3.4. Pneumonia

The results for patients admitted for pneumonia are shown in Table 10. The results of the Hausman test suggest that the random effects estimator is the preferred model. I found that a one percentage point increase in the achievement rate on the blood cultures measure lowers the mortality rate by 0.29 percentage points. This effect is statistically significant

at the 0.1% level. I do not found that any other process measures of care were statistically significant. The effect sizes of antibiotic selection and antibiotics received were an order of 10 smaller than the effect size of blood cultures.

<Insert Table 10>

Increasing the proportion at which patients were excluded from blood cultures by one percentage point lowers the probability of patients being discharged as dead by 0.26 percentage points when statistical significance is measured at 0.1%. A one percentage point increase in exclusions from oxygenation assessment increases patient mortality by 0.09 percentage points, statistically significant at the 5% level.

The relationship between the appropriate care score and mortality is shown in Table 11. A one percentage point increase in the appropriate care score reduces the mortality rate by 0.11 percentage points. This result is statistically significant at the 0.1% level.

<Insert Table 11>

3.4.4. Spell level analysis

3.4.4.1. AMI

Table 4 shows the effects of Trust and individual level achievement on all process measures of care on in-hospital patient mortality. Patients attending hospitals which perform one percent higher on aspirin at arrival have a decreased mortality probability of 0.19 percentage points, which is statistically significant at the 5% level. No other Trust-quarter level process measures were statistically significant. Patients attending Trusts which exclude one percentage point more patients have a decrease in mortality percentage by 0.19 percentage points. Results for the individual spell level process measures complement the findings from the Trust-quarter level analysis. Patients who were given aspirin at arrival have a 7.4 percentage point lower probability of mortality. Patients who were excluded from aspirin at arrival have a 4.7 percentage point decrease in mortality probability.

Table 5 shows the estimated effects of the appropriate care score for AMI patients on in-hospital mortality. I found that a one percentage point increase in the appropriate care score decreases patient mortality by 0.12 percentage points statistically significant at the 5% level. Patients who have been given all process measures of care were 14.1 percentage points less likely to during their hospital stay compared to patients who were not given all of the process measures of care from which they were not excluded from. Removing the endogeneity, I found that patients who were given all appropriate process measures of care were 2.7 percent less likely to die during their hospital stay. This is statistically significant at the 1% level.

3.4.4.2. CABG

Table 6 shows the effects of the performance on the process measures of care for CABG on spell level in-patient mortality. As with the Trust-quarter level analyses, I did not find any statistical significance, effects at a 5% for any of the process measures of care.

Table 7 shows the effects of the appropriate care score for CABG patients on in-hospital mortality. I found that there is not a statistically significant relationship between Trust performance on the appropriate care score and patient mortality at the 5% significance level. Patients who have been given all appropriate measures of care have a 1.6 percentage point lower probability of mortality, which is statistically significant at the 0.1% level. From the bivariate probit results, at a 5% statistically significance level, patients who have been given all appropriate care measures have a 1.5 percent higher in-hospital mortality risk.

3.4.4.3. Hip and knee

Table 8 shows the effect of performance on process measures of care for hip and knee on readmissions. I found that a one percentage point increase in average Trust performance on antibiotics selected decreases the probability of patient's readmission rates by 0.05 percentage points significant to 1% level. The effect sizes of average Trust achievement rates of hip and knee were low. The effect of selected antibiotics is the largest. Rates of

antibiotics within one hour of surgery and receipt of prophylactic antibiotics cause a 1.5 and 1.6 percentage point decreases in the probability of being readmitted. They were statistically significant at the 1% and 5% levels respectively. Patients who have prophylactic antibiotics ordered have a 3 percentage point higher probability of being readmitted. This is statistically significant at the 1% level.

The estimated effects of the appropriate care score on 30-day readmission are shown in Table 9. Patients attending Trusts which increased their mean appropriate care score by one percent have a 0.074 percentage point decrease in probability of being readmitted. This is statistically significant at the 1% level. Patients who receive all appropriate process measures of care had a 1.6 percentage point lower probability of being readmitted when compared to patients who did not receive all appropriate process measures of care. This is statistically significant at the 0.1% level. In the bivariate probit regression, I did not find any statistical significance. I also found that the correlation between the two errors of the bivariate probit models is not statistically significant. This implies that the bivariate probit was not necessary when modelling the relationship between appropriate care score and 30 day readmissions.

3.4.4.4. Pneumonia

Table 10 shows the effects of the performance on process measures of care for pneumonia on spell level in-hospital mortality. I found that Trusts that increase achievement rates of initial antibiotics received within six hours of arrival by one percentage point increases the probability of patient mortality by 0.08%. Trusts which had a one percent higher exclusion rate for oxygenation assessments increase the probability of patient mortality by 0.31 percentage points.

I found that Trust performance on appropriate care score scores is not related to a patient's mortality probability when the threshold of statistical significant is at 5% level (Table 11). Patients who were given all appropriate process measures of care had a 16.5 percentage point lower probability of in-hospital mortality. This is statistically significant at the 0.1% level. When I remove the endogeneity from the patient level appropriate care score, I found that patients who were given all appropriate process measures of care had a

1.7 percentage point lower mortality risk than patients who did not receive all appropriate process measures.

3.4.5. Robustness check for AMI and pneumonia

Table 12 and Coefficients signify proportions; a 0.01 coefficient signifies 1%
Table 13 show results from Trust level analysis to examine 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, the crude 30 day mortality rate and the risk adjusted 30 day mortality rate, I limited my sample to 14 quarters of data. I found that limiting the sample to 14 quarters did not change the statistical significance of the findings compared to estimations with 18 quarters of data.

3.4.5.1. AMI

Table 12 shows that one percentage increases in achievement or exclusion of aspirin at arrival were both associated with a 0.3% percentage point decrease in 30 day in-hospital mortality rate. This is statistically significant at the 0.1% level. Trusts that excluded a one percentage point higher proportion of patients from smoking cessation advice had lower mortality rates by 0.09 percentage points. This is statistically significant at 5% level. No other process measures of care were statistically significant at the 5% level

Using a risk adjusted 30 day in-hospital mortality measure, I found that a one percentage increases in achievement or exclusion of aspirin at arrival were associated with 0.28% and 0.25% reduction in mortality respectively. These results were statistically significant at 0.1% and 1% respectively. With risk adjusted in-hospital mortality, the rate of exclusion from smoking cessation advice is no longer statistically significant at the 5% level.

3.4.5.2. Pneumonia

Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 13 shows that a one percentage point increase in the provision of blood cultures decreases the 30 day in-hospital mortality rate by 0.2 percentage points. This is statistically significant at the 1% level. Other process measures of care do not produce any statistically significant results. Exclusion from smoking cessation advice and blood cultures had statistically significant effects. One percentage point increases in the exclusion rates lead to around 0.22 and 0.16 percentage point decreases in mortality respectively.

I found a one percentage point increase in performing blood cultures before initial antibiotics results in a 0.23 percentage point decrease in the risk-adjusted 30 day mortality rate. This is statistically significant at the 0.1% level. A one percentage point increase in the exclusion rate of this process measure is also associated with a reduction in the mortality rate by 0.18 percentage points. This is statistically significant at the 1% level.

3.4.6. Potential Effect of Advancing Quality on health outcomes

Table 14 shows the estimated the potential effects of the Advancing Quality scheme on lives saved and readmissions avoided. I found from the Trust level analysis that Advancing Quality saved 467 and 537 lives respectively for AMI and pneumonia over the four and a half years since the quality incentive scheme was introduced. This is 1% of the AMI patient sample and 0.5% of the pneumonia patient sample. I found that 1122 readmissions were avoided for hip and knee replacement patients. This is 1.6% of the hip and knee patient sample.

From the Trust level analysis using 42 months of data, I found that Advancing Quality directly saved 317 and 238 lives for AMI and pneumonia respectively. When the outcome variable is risk adjusted 30 day in-hospital mortality, I found that Advancing Quality saved 239 and 274 lives for AMI and pneumonia respectively.

From individual level analysis over my study period, I found that Advancing Quality directly saved 45 and 84 lives for AMI and pneumonia respectively. This is 0.1% of the AMI and pneumonia patient samples. 121 readmissions were avoided for patients admitted with hip and knee replacements. This is 0.17% of the hip and knee patient

population. I prefer the results of the individual level analysis as my data links patients quality of care directly to health outcomes.

3.5. Discussion

3.5.1. Summary of key Findings

I found that overall process measures of care were associated with better patient outcomes. However, not all incentivised process measures of care were associated with health outcomes. Process measures of care may not be associated with health outcomes due to the following two reasons. Firstly health outcomes such as mortality and readmissions were affected by many factors which may or may not include process measures of care such as patient condition severity. The ability to find correlation between process measures and health outcomes may therefore be weak. Secondly, all process measures of care for a clinical condition were used together in the analysis. Process measures of care for a clinical condition may be correlated, therefore using all process measures of care together will result in some process measures of care having statistically insignificant associations with health outcomes. Due to correlated process measures of care, I could not test the association of each process measure of care with health outcomes due to omitted variable bias which may influence statistical significance (Wooldridge 2013) of each process measures. Findings of some process measures of care being statistically significant or not may not provide more information than using a composite quality score due to correlated process measures of care.

I found that using an appropriate care score that process measures were associated with health outcomes for AMI, hip and knee replacements and pneumonia. I summarise the result in more detail for all clinical conditions in sections, 3.5.1.1, 3.5.1.2, 3.5.1.3 and 3.5.1.4.

3.5.1.1. AMI

I found consistently that Trusts that had higher achievements of and exclusion for aspirin at arrival had lower mortality rates. The result on exclusions suggests that Trusts may not incorrectly exclude patients as being excluded from aspirin at arrival is associated with lower Trust-quarter level mortality and in-hospital patient level mortality.

These results for AMI suggest that a financial incentive for NHS Trusts in the North West of England to improve patient quality of care has led to improvements in patient outcomes. my findings suggest that providing process measures as a package of care reduces mortality, with aspirin at arrival being the main driver for improvements in reducing mortality on the Trust level. I found that Trusts who had higher appropriate care scores had lower mortality.

However, as achievement and at exclusion percentage were high, people who had failed to receive the process measure were low, just 1.2 percent of patient AMI population. The low number of failed patients may be driving the large consistent effects for the process measure. The provision of providing aspirin at arrival has hit a ceiling effect where achievement rates without exclusions were high so that it is not possible for Trusts to improve achievement rate by one percent. As the effect of aspirin on mortality is so strong, the high performance of Trusts providing aspirin at arrival should be maintained.

3.5.1.2. CABG

I did not find any evidence of an association between process measures of care and health outcomes for patients admitted with CABG at the Trust level. The individual appropriate care score regressions suggest that patients who had received all appropriate process measures of care were more likely to die in-hospital.

In the VIF for CABG, I found that process measures for CABG were highly correlated. This is because I had a low number of observations and low variation of health outcomes due to the low mortality rate. These factors may explain the statistically insignificant appropriate care score.

In individual process measures of care, all process measures were statistically insignificantly related to mortality where the standard errors were also large implying that the probability of the process measures having an effect on patient mortality is low. This

elective condition exhibits lower mortality rates when compared to the emergency conditions of AMI and pneumonia. The low mortality rate may not have enough variation to explain if the process measures have an effect.

3.5.1.3. Hip and knee replacements

I found that Trusts that perform better on the appropriate care score for hip and knee replacement had lower 30 day readmission rates. This is also reflected when I analyse spells. Patients who had received all appropriate process measures of care had a lower probability of readmission.

For the Trust-quarter level panel data results, I found that the process measures of care were not statistically significant individually. However, when combined into a composite quality metric, I found that all process measures of care together were jointly significant.

From the individual level regressions, I found that process measures of care for hip and knee patients were not endogenous. The appropriate care score from the individual level probit and bivariate probit methods has resulted in similar coefficients. The health outcome for hip and knee patients is a 30 day readmission rate, as this condition is elective, the level of a patient's health may not affect the quality of care received.

In individual level process measures from the probit regressions, I found that prophylactic antibiotics ordered is associated with an increased risk of readmissions. This process measure is likely to be picking up patient severity, where more severe patients will be administered the process measure.

3.5.1.4. Pneumonia

I found that Trusts who had higher percentages of appropriate care score achievement for pneumonia had a lower mortality rate. This is also reflected in the individual level analysis.

I did not find a relationship between hospital level achievement on process measures of care and health outcomes at the individual level. The significant results from the

achievement of antibiotics received may be attributed two reasons. The first that hospitals that generally provide the better quality of care were treating patients who were harder to treat and therefore, that is related to a higher probability of patient mortality. The second is that the provision of initial antibiotics after 6 hours of hospital arrival may be too late. Antibiotics delivered after 4 hours were linked to increase mortality (Cassel et al. 2014).

In the panel data analysis, I found a relationship between the process measures of care and health outcomes for pneumonia. When hospitals perform one percentage point more blood cultures for pneumonia patients, it lowers the probability of the patient mortality by around 0.29 percentage points. However, this effect may be driven by a small number of patients due to the high exclusion rates for this process measure.

3.5.1.5. *Removal of endogeneity*

I found that the bivariate probit was the preferred method for AMI, CABG and pneumonia. For hip and knee replacement surgery, I found that the 30 day readmission rate and provision of process measures of care did not exhibit an endogeneity problem which means that the probit regression is the preferred estimator for hip and knee replacement clinical condition.

For AMI, CABG and pneumonia, I found that addressing the endogeneity through simultaneity, has made the estimated coefficients less negative. This means that the endogeneity exhibited a negative bias for process measures of care and health outcomes. This finding is consistent with the intuition that hospitals excluded patients in poor health due to lower probability of health gain. Providing process measures of care to populations in good health will result in a stronger relationship between process measures of care and mortality as patients in good health will not result in death regardless of provision of process measures of care.

3.5.2. How those key findings relate to literature

3.5.2.1. AMI

Our study agrees to previous the studies by (Bradley et al. 2006; Jha et al. 2007; Peterson et al. 2006; Werner and Bradlow 2006; Werner, Bradlow, and Asch 2008). However, I further identified that the administering aspirin at arrival is a driving factor when considering all process measures of care together. This result also agrees to the study by Epstein et al. (2014) which tested whether there were potential negative effects of pay for performance schemes on the quality of care for patients. Epstein et al. (2014) focused on the provisions of healthcare for patients and found no detrimental effects from potential care substitutions.

3.5.2.2. Pneumonia

Our findings adds to the studies conducted by Lee et al. (2011) who used pneumonia 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. Complementing the study by Lee et al. (2011), I found weak significance between mortality and process measures of care. However I did not find that all process measures of care were associated with lower mortality which reflects the study by (Sucov, Valente, and Reinert 2013) which found that time for first antibiotics is not related to mortality.

Other literature which also looked at the impact of process measures of care for pneumonia (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 I conclude that there were other factors which I do not explicitly capture, which is the cause of patient mortality.

This research also complements the study by Sutton et al. (2012). The authors found that Advancing Quality reduced patient mortality. I used the same mortality variables which these authors used. I extended the time periods the variable covers. Sutton et al. (2012) found that the first 18 months of Advancing Quality reduced 890 deaths from patients admitted with AMI, heart failure and pneumonia. Using the same mortality variable over 42 months of Advancing Quality, I found that Advancing Quality directly saved 513 lives

from patients admitted with AMI and pneumonia. my study finds lower lives saved as my study finds the direct effects of Advancing Quality. The study by Sutton et al. (2012) uses all patients admitted to North West hospitals with AMI, heart failure and pneumonia. Therefore the lives saved calculated by Sutton et al. (2012) included direct and indirect effects of Advancing Quality.

The findings of my research is not consistent with Kristensen et al. (2014). Kristensen et al. (2014) did not find that Advancing Quality affected mortality after the first 18 months. The authors concluded that Advancing Quality did not result in long term mortality reductions. I found that between the 43rd and 54th month of Advancing Quality saved a further 449 lives. Kristensen et al. (2014) did not use data up to 54 months after the introduction of Advancing Quality which may explain the difference in results. I found that the final year of my study period resulted in a large effect of Advancing Quality on improved patient outcomes.

Our study further adds to the literature as I found for hip and knee patients that all process measures of care as a care package for patients were linked to reductions in readmissions. This association using elective care indicated that process measures of care to improve health outcomes had been successful, in both emergency and elective conditions.

3.5.3. Strengths of this study

To my knowledge, this is the first study which used an individual level dataset. I had been able to link patients' outcomes to the process measures of care at patient level over a four and a half year time period. This has enabled us to use a variety of panel data and cross sectional econometric methods. I were able to adjust for the endogeneity between appropriate process measures of care composite scores and health outcomes which previous studies had not adjusted for.

The flexibility from the available data also enabled us to deal with other potential limitations, such as checking the concern over of the ecological fallacy by running my analysis on patient and Trust level. This provided a more comprehensive study of the links between the process measures of care and health outcomes.

This study also had 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 possible selection bias to join the incentive scheme. This makes the results more generalisable unlike the studies from the US where hospitals self-select into P4P programmes.

3.5.4. Limitations of this study

Our research has the following limitations.

1. 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 my 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 my sample of patients. I used a Trust-quarter level 30 day mortality rate for AMI and pneumonia to test the robustness of my results, however, a patient level 30 day mortality rate would be ideal for my analysis.
2. I had few measures of patient case-mix (gender, age, ethnicity and area of residence level deprivation). This issue however cannot be fully addressed with my dataset. By including all process measures of care, some of this patient severity will be captured by achievement of the process measures of care I used (Sucov, Valente, and Reinert 2013). However, severity of the care condition which is not captured in administrative datasets would account for condition severity.
3. Due to data recording issues, I did not have a balanced panel for my 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.
4. For the Trust level analysis, I did not weight the observations to account for Trusts treating different number of patients. Having no weights assumes that all Trusts were the same size. As each trust level variable have no aspect of size, the relationship between health outcomes and process measures will not be correctly reflected.

5. Only six NHS Trusts in North West of England provides CABGs. As some hospitals have low numbers of patients for this condition, there were a low number of observations on the Trust-quarter level analysis. This means that I do not have enough observations for the number of parameters in my models. On the other hand, the number of observations for the spell level regressions for CABG is of a suitable sample size.
6. In this Chapter, I assumed that health outcomes were only affected by the Advancing Quality initiative. Patient health outcomes may also be affected by other health care factors during my study period. I did not capture this increase in overall quality of care. I use quarter time dummies to adjust for differences in quality of care from Advancing Quality, but not for how the general quality of care has improved over time. This affects my calculation as I do not know over time how the general quality of care has changed, or, I do not know how the mortality rate has changed over time. Therefore, any associations between process measures of care and health outcomes that I found may not be directly attributed with the direct effect of AQ due to no control groups or prior data from the introduction of AQ.

3.5.5. Policy Implications

The selection of the process measures for Advancing Quality was reported to be evidence based. This research has found that improvements in these process measures of care were associated with health outcomes. However, I have found that not all process measures of care were related to health outcomes. This is especially the case for CABG where I did not find a relationship for any of the measures.

I suggest that policy makers implementing P4P schemes should weight and wait. Under Advancing Quality, all process measures of care have an equal weight when calculating overall Trust performance. my results suggest that not all process measures led to lower patient mortality or readmission rates. This research adds support that different process measures should have different weights when rewarding Trusts as not all process measures of care impact on health outcomes (Cassel et al. 2014). Policy makers should also wait for empirical research to be conducted before making decisions to remove process measures of care (Reeves et al. 2010). Under the Advancing Quality scheme, process measures of care have been removed without any evidence base for doing so.

This is a concern as I found that performing blood cultures before initial antibiotics was associated with a lower mortality rate. However, this measure has now been removed as a quality measure for Advancing Quality as not all patients with pneumonia qualify to receive this measure (Advancing Quality 2013).

3.6. Conclusion

In this chapter, I found that incentivised process measures of care were positively associated with improved patient outcomes. Advancing Quality incentivised effective quality of care which would potentially result in 129 lives saved for AMI and pneumonia combined and avoided 121 readmissions from hip and knee replacement conditions, if I would assume that the provision of process measures of care were solely related to AQ and that health outcomes since October 2008 to April 2013 were fixed. In the next chapter, I will analyse whether incentivised quality of care is distributed equitably to patients from different socioeconomic backgrounds.

4. “Rich or poor, who gets more?” The distribution of quality of care under a hospital pay-for-performance programme

Equitable distribution of health care is a core goal of the NHS (Department of Health 2000). The equity of healthcare delivery can be assessed in numerous ways. Vertical equity is when the distribution of healthcare varies by different need (Goddard and Smith 2001). Horizontal equity states that patients with equal need should be given equal healthcare (Dixon et al. 2003). Equity may also be assessed through supply within both horizontal and vertical equity. For this study I considered supply side horizontal equity (Goddard and Smith 2001). I defined equity as the quality of health care being the same across different patient groups.

The introduction of financial incentives for quality improvement in hospitals may call into question the achievement of the goal of equal distribution of care. A tournament payment structure in particular may increase healthcare inequalities at the patient level. As financial incentives were paid based on achievement, poorer performing hospitals may not increase the provision of process measures of care (Ryan et al. 2012). Although still a part of the incentive scheme, the hospital may choose not to compete. In the tournament however, as a system of benchmarking is introduced, the differences in the quality of healthcare across different patient groups may narrow as all healthcare providers were now incentivised.

I considered a quality incentive program, Advancing Quality, introduced into the North West of England, which financially incentivised Trusts to provide process measures of care to patients. All Trusts in the North West of England were included in this quality incentive scheme, and in the first year, competed against each other through the provision of process measures of care in a tournament style where the top 50% of performing Trusts were awarded a financial bonus. After the first year, a system of benchmarking and year on year quality incentives were introduced (Ledward, Horne, and Butterworth 2008).

The aim of this chapter was to examine whether the achievement of process measures of care incentivised by Advancing Quality has been distributed differently across different patient groups, and whether this has changed over time. I included emergency and elective clinical conditions to test whether there is a difference in the distribution of planned and unplanned care.

4.1 Literature review

For this literature review, I discussed two systematic reviews which have examined the distribution of care within financial incentive systems. I then reviewed papers which focus on the general practice setting and finally on the hospital setting. I found literature using a google scholar and Web of Science, searching for the following terms, "financial incentive, "P4P", "socioeconomic", "deprivation" and "quality of care". Once I found papers, I performed pearl growing to search for literature using referenced and citing articles.

4.1.1. Systematic reviews

Two systematic reviews were found which aimed to find how the quality of care under a financial incentive scheme has been distributed. These studies examine the effect of how quality if care under a financial incentive scheme has been distributed between socioeconomic groups.

Verlinde et al (2012) conducted a systematic review to examine how patient social gradient would affect communication with their doctor, and therefore affect the quality of care received by the patient (Verlinde et al. 2012). The authors searched literature from 1965 to 2011 and found a total of 20 papers, where 5 papers focused on researching doctor-patient communication and social class where social class was measured by income, education or occupation or any combinations of the three.

One reviewed study, examined the patients' income level on the level of advice given by the healthcare physician (Taira et al. 1997). Using data from questionnaires over a sample of 6549 patients, the authors found that patients who have low incomes were more likely to have received smoking cessation advice. Patients who have high incomes were more likely to have discussed diets and exercise with their doctors.

Two reviewed studies, examined the patients' education level on the level of advice given by the healthcare physician (Street 1991) and (Street 1992). Both papers used data from

videotaped consultations with 41 patients. Both reviewed studies found that patients who have a higher educational background receive more information from the doctor due to patients having a more effective communication style and asking more questions.

Two studies reviewed by (Verlinde et al. 2012) examined a patient's education and income level as measures of social class on the level of information given by the doctor (Hall, Roter, and Katz 1988; Maly et al. 2009). Hall, Roter, and Katz (1988) used data from professional health care providers from 157 patients and found that patients from a higher social class received more information than patients from lower social class. Maly et al. (2009) collected 327 patient responses from a survey for breast cancer patients. The authors found that a higher education level was related to patients receiving more information from health care physicians.

Verlinde et al. (2012) found one study which looked into the relationship between education and drug prescriptions (Stewart 1984). The study used data from audio tapes from 140 patient consultations. The study found that patients who have a higher education level such as having a university degree were more trusted with receiving drug prescriptions, whereas patients with lower education attainment received more emotional support.

One limitation of this study is that the review looks at the distribution of care where no quality incentive programs were taking place. Therefore the study does not examine the distribution of care when quality of care in incentivised which is the aim of my study. The other issue is the relevant studies reviewed were old. The relevance of these studies may not be the same in modern day. Within the scope of this study, the authors have not found more recent literature. On the other hand, I can still use the findings from the papers reviewed by Verlinde et al. (2012) as they provide insights in how a healthcare physician will react to patients of different income and educational backgrounds.

4.1.1.2. Alshamsan et al. (2010)

Alshamsan, et al. (2010) reviewed the impact of financial incentives on inequalities in quality of health care. The authors searched for literature from 1st January 1980 until 1st November 2008 for articles published in English. This systematic review consisted of 22

studies. Twenty were based on the Quality and Outcomes Framework. Most studies were observational with a cross-sectional design using data from the Quality Management and Analysis System dataset which meant that data were at practice level. This study selected studies that focused on pay for performance (P4P) for a pre-specified level of quality and healthcare inequalities by: socioeconomic deprivation; age; sex; and ethnicity.

The majority of papers reviewed by the authors found that the quality of care in more deprived areas were lower than the quality of care from less deprived areas with the exception of one paper, Sutton and McLean (2006), which found the opposite effect. The majority of studies found that the effect size was small and did not remain after a few years of the incentive scheme. Furthermore, inequalities between age, gender and ethnicity did not change as there was persistence in the quality differential between the different patient groups. The authors concluded that the current P4P schemes were unsuccessful in reducing healthcare inequalities.

One discussion point which the authors mentioned was the limitation of practice level data when examining healthcare inequalities between patients. Once patient level data is aggregated into practice level, the data on practice level becomes the practice population average of the four inequality stratum. This means that the effects of deprivation will not be based on patients' strata, but a less precise and sensitive aggregated measure. This means difference between age, gender, socioeconomic deprivation and ethnicity will not be identified as the significant factors will be aggregated to more homogeneous measures.

4.1.2. The general practice setting

4.1.2.1. Provider level

In this section I reviewed papers that have examined the achievement rates of the QOF but have stratified the sample based on aggregated deprivation levels. The aggregation is the average deprivation of the catchment areas which the practice operates. Most studies used the Index of Multiple Deprivation.

The overall conclusion from the six research papers (Ashworth et al. 2011; Dixon, Khachatryan, and Tian 2012; Doran et al. 2008; Ashworth et al. 2007; Millett et al. 2007;

Saxena et al. 2007) is that the quality of care does vary depending on the level of deprivation. More deprived areas have worse scores on process measures of quality. Only one study I reviewed showed no effects of deprivation levels on healthcare quality (Sutton and McLean 2006). However, this study used a small sample from a nationally unrepresentative region where deprivation is higher than the national average. The findings from these research papers seem to provide evidence of the inverse care law (Doran, Fullwood, Kontopantelis, et al. 2008) where the patients who were less in need of healthcare will be the patients who will receive the most care (Hart 1971).

Further findings have found that practices who were situated in more deprived areas have been improving the fastest even though overall quality performance is lower (Dixon, Khachatryan, and Tian 2012; Doran, Fullwood, Kontopantelis, et al. 2008; Ashworth, Medina, and Morgan 2008). This finding indicates that poorer performing practices have exerted more effort to provide care for the patients. The reasons for this improvement in quality of care may be down to a two reasons:

1. Good performing practices that were situated in less deprived areas did not need to exert much effort to achieve the highest achievement rate necessary for GPs to gain the financial incentive (Doran, Fullwood, Kontopantelis, et al. 2008). Therefore the incentive to improve performance in terms of achievement rates did not require much effort.
2. Prevalence rates of diseases were higher in areas where deprivation is greater (Doran, Fullwood, Kontopantelis, et al. 2008). As the QOF is based of achievement rates from the eligible practice list which means that these increases in achievement rates require more work for GPs in more deprived areas.

These studies all add to the literature base which examines how the quality of care varies by aggregate level practice deprivation. They also provide insights into how or where the quality of care has improved over time.

However, the main issue with using aggregated indicators of quality is that practices were rarely situated in areas where the catchment areas of the practice have the same levels of deprivation. This means that the practice average of deprivation based on the average deprivation of the catchment area where the patients live will dilute the effects of deprivation measures. Therefore the true effect of deprivation on the quality of care will be harder to find due to lower levels of variation within the deprivation of practices.

Four papers (Ashworth et al. 2011; Doran, Fullwood, Kontopantelis, et al. 2008; Ashworth, Medina, and Morgan 2008; Ashworth et al. 2007) used data on practice area level data attributed to the practice from the UK 2001 census. Although census data were complete, as it is compulsory to be completed by residents in England, they were dated and may not reflect the current area characteristics of the population. Furthermore the methods used for this analysis such as Ordinary Least Squares (OLS) regression (Dixon, Khachatryan, and Tian 2012; Ashworth, Medina, and Morgan 2008; Millett et al. 2007; Saxena et al. 2007; M. Sutton and McLean 2006). These may not be the most appropriate due to the negative skewed nature of the achievement rates.

In the next section, I focused on studies which have looked into deprivation at an individual level to look for the effects of deprivation. Although the previous studies have shown that practices operating in the most deprived areas have lower quality of care, these were the areas where improvement is growing fastest. There were concerns about whether the research conducted answers my research question on the types of individuals who were affected by financial incentives.

4.1.2.2. Patient level

Even though the QOF incentive scheme is aimed that a practice level, individual level analysis will be needed to mitigate the concerns about the 'ecological fallacy' (Freedman 2001) which is the assumption that the relationship between practice level deprivation and quality of care can hold even on an individual level. In other words there may be a correlation between aggregated quality of care and aggregated deprivation strata and causation is not only assumed on an aggregated level, but also assumed for individual level.

Other than the ecological fallacy, the other advantage of using patient level data is the increase in accuracy of the data. Research into quality of care stratified by deprivation either used the deprivation level of the practice, or the deprivation level of the area where the patients live. McLean et al. (2008) found that studies which looked into the area deprivation of the practice capture the same relationship, but underestimated the effect of deprivation when compared to patients' area deprivation.

There is not much evidence on the relationship between deprivation and quality of care on an individual level. However from the research conducted, the effect of patient area deprivation on the quality of care given is mixed. Two studies have found that generally patient area level deprivation is not significant when analysing the trends pre and post QOF on the incentive conditions (McGovern et al. 2008; Millett et al. 2009). Two papers which examined recording levels of practices identified that patients who were from more deprived areas have less chance of getting the quality of care documented, although more deprived areas have a higher prevalence of co-morbidities (Simpson et al. 2007; Simpson et al. 2006).

The sole paper I found which used deprivation at a patient level was conducted by (Crawley et al. 2009). The data the authors used was from The Health Survey for England over two time points, 2003 and 2006. The paper examinined patients under three clinical areas: coronary heart disease; diabetes and hypertension, where the outcome measures were clinical indicators of care. The deprivation measure used was whether a respondent's social class, was manual or non-manual. The authors found that the distribution of clinical indicators incentivised under the QOF was not different between patients from different social classes.

Although Crawley et al. (2009) used patients' deprivation, the measure of deprivation was banded into two groups due to low number of observations within each category (Crawley et al. 2009). This means that the measure of deprivation was a highly aggregated measure in terms of categories and hence may not have been sensitive enough to capture differences in healthcare distibution across deprivation groups.

4.1.3. The hospital setting

Two papers Ryan et al. (2012) and Jha, Orav, and Epstein (2010) were found on the hospital setting. Ryan et al. (2012) tested if the change of incentive structure impacted on the distribution of incentive payments across socioeconomic groups. They used aggregated data from Premier, and Medicare, CMS index and American hospital annual survey data for finance and research all five clinical areas; AMI, heart failure, pneumonia, coronary artery bypass graft and hip and knee replacement. The incentive structure change was going from a tournament system where hospital compete on achievement

rates of process measures of care to a benchmarking system where each trust were set target achievement rates of care (Ryan et al. 2012). The study used hospital level data from the Hospital Quality Incentive Demonstration (HQID) over five years for 229 hospitals. The authors captured patient's socioeconomic status at hospital level by using the Medicare Disproportionate Share Index and created quartiles to group hospitals. The research was econometric analysis using a difference in differences design. The authors found that under the tournament system, there were large disparities in healthcare payments based on quartiles of deprivation. Hospitals had a 32.8 percentage point differential in the likelihood of receiving payment from supplying care based on socioeconomic deprivation. Under a benchmark system, no statistically significant results were found. This study did not focus on healthcare quality, only payments. The socioeconomic deprivation is based on income from aggregated hospital data and therefore analysis is not on individual patients.

Jha, Orav, and Epstein (2010) examined whether pay for performance for hospital quality impacted on hospitals serving more poor patients differently to hospitals with fewer poor patients. The authors created a disproportionate share index to proxy the hospitals share of poor patients using aggregate data from Premier HQID, for three conditions, AMI, HF and PN. The authors used three years of HQID data for 255 hospitals from October 2003 to September 2006. Using bivariate and multivariate regression analysis, Jha, Orav, and Epstein (2010) found that AMI and pneumonia, the quality of care for hospitals serving proportionally more poor patients responded to financial incentives and quality of care caught up with hospitals serving proportionally less poor patients. With aggregated data, the authors cannot look at who benefited or gained in quality from the patients served by the hospital

However, one may not generalise all HQID research and apply to the Advancing Quality initiative due to the differences between the two schemes (Sutton et al. 2012). Two studies by Sutton et al. (2012) and Ryan et al. (2009) aimed to find the effects of a financial incentive scheme, Advancing Quality and HQID respectively, on patient mortality. The two studies used patient level hospital administrative data on emergency conditions of AMI, heart failure and pneumonia. Ryan et al. (2009) also included CABG, however Sutton et al. (2012) omitted patients with CABG due to low number of Trusts in the North West region of England performing that procedure. A triple difference (difference-in-difference-in-differences) design was implemented in both papers. Sutton et al. (2012) found that Advancing Quality incentive scheme has led to a reduction in

mortality whereas Ryan et al. (2009) found no effect on mortality. Reasons for the difference in findings between the two studies may be due to differences in the pay-out size and structure (Ryan and Blustein 2011).

Results from studies in inequities from the US may therefore not translate over to the UK due to the differences in the overall healthcare in distribution of care in both systems. The UK has universal health insurance coverage for all patients where a goal is the distribution of care of healthcare. The US has a more complex structure of private, public and no insurance coverage which results in an inequitable healthcare system (Blumenthal and Dixon 2012).

Another problem with most studies using the PHQID is the aggregated nature of the data. Although, aggregated data will produce the overall effect of an incentive scheme, the results may be misleading due to the ecological fallacy (Schwartz 1994). Results from aggregated analysis may not translate to an individual level. This applies for both patients and incentivised process measures of care.

Studies on the PHQID which look into healthcare inequalities among patients from different socioeconomic backgrounds found that quality of care at the start of the incentive scheme is distributed pro rich, however, over time the difference in care between narrows (Ryan et al. 2012; Werner 2010; Jha, Orav, and Epstein 2010).

4.2. Aim

The aim of this chapter was to examine how financial incentives have been distributed between different patient groups. My study in this chapter was the first paper to analyse the types of patients who receive process of care measures from the Advancing Quality initiative which will highlight the distributions of the quality of care at a patient process measures of care level.

Most research conducted previously uses patients' area income as the main deprivation variable (Alshamsan et al. 2010). Linking a patients' area income level to patients who were emergency cases in need of process measures of care incentivised by Advancing Quality is expected to be low. When admitted, hospitals may not distinguish the financial backgrounds across patients. As Trusts will be financially rewarded to supply process

measures of care, I hypothesised that there should be no variation in the quality of care under a financial incentive scheme under a financial incentive. This is because the health care providers should provide care to everyone to meet the targets of the financial incentive. The theory behind incentives in economics is the Principal Agent theory (Ross 1973). Under the principal agent framework, a principal (in this case Advancing Quality) delegates tasks to be conducted by an agent (Trusts); where an agent may have different interests with the principal. The principal must therefore create an incentive to motivate the agent to conduct the delegated tasks. An incentive is therefore defined as a medium that will induce an agent's action (inaction) to conduct tasks set by the principal. In our case, this is creating financial incentives to incentivise Trusts to provide process measures of care.

As this research will be conducted for each process measure of care, I grouped the process measures of care. These groups are:

1. Process measures which were given in emergencies
2. Process measures based on drugs and tests
3. Process measures based on giving advice.

4.3. Methods

4.3.1. Econometric models

I applied multilevel cross section econometric analysis to address the aim of this chapter. The analysis was conducted at process measure level where a spell will have repeated observations by the number of process measures of care for that condition. The outcome variable was a categorical variable with four outcomes. The regression techniques I employed to measure these four outcomes simultaneously were the multinomial logistic regression (Wooldridge 2013) and the sequential logistic regression (Buis 2011).

4.3.1.1 Multinomial Logit

The multinomial logit regression estimated separate probability models simultaneously. The advantage of regressing separate probability models simultaneously is an improved efficiency of the models which results in different coefficients and lower standard errors.

The probabilities of other outcomes will be relative to the population of patients who have been given the clinical indictor; this is given by equation 8.

$$\mathbf{Pr}(\boldsymbol{y} = \boldsymbol{a}|\boldsymbol{x}) = \frac{\mathbf{exp}(\boldsymbol{x}\boldsymbol{\beta}_{\boldsymbol{a}|\boldsymbol{b}})}{\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{i}}\mathbf{exp}(\boldsymbol{x}\boldsymbol{\beta}_{\boldsymbol{i}|\boldsymbol{b}})} \qquad \text{Equation 7}$$

Where *y* is a categorical outcome variable which has *i* outcome categories such as *a* and *b*; $\boldsymbol{x}$ is a vector of covariates; $\boldsymbol{\beta}$ is a vector of parameters. Outcome *b* is used as the base outcome.

The formal model of the multinomial logit can be written as (Long & Freese, 2006):

$$\mathbf{ln}\boldsymbol{\Omega}_{\boldsymbol{i}|\boldsymbol{b}}(\boldsymbol{x}) = \boldsymbol{ln}\frac{\mathbf{Pr}(\boldsymbol{y}=\boldsymbol{i}|\boldsymbol{x})}{\mathbf{Pr}(\boldsymbol{y}=\boldsymbol{b}|\boldsymbol{x})} = \boldsymbol{x}\boldsymbol{\beta}_{\boldsymbol{i}|\boldsymbol{b}} \ \ \mathbf{for}\ \boldsymbol{i} = \mathbf{1}\ \mathbf{to}\ \mathbf{4} \qquad \text{Equation 8}$$

When using the multinomial logit, I assumed that for each process measure of care, a hospital will decide whether a patient will be given a process measure (pass); fail to be given a process measure (fail); be excluded alive or dead at one time point as each outcome is independent of one another. The independence of the process measures of care is illustrated in Figure 6.

<Insert Figure 6>

The main assumption of multinomial regressions is the independence of irrelevant alternatives (IIA) (Dow and Endersby 2004). This assumption states that if one outcome is more likely than another (say A is preferred to B), and then the appearance of a third option (C) will not change the distribution of probability between the first two choices. For example, assume that there were 100 economic agents, and 60 agent choose A and 40 choose B. If a choice C appears which is a perfect substitute to A, then the 60 agents who chose A will be split between A and C, whereas B is unaffected. This may lead to 30 agents choosing A, 40 agents choosing B and 30 agents choosing C. Here IIA has failed as the relationship between A and B have changed as now more agents choose B over A.

This assumption of IIA will be tested after my estimation by using the Hausman McFadden test for IIA (Cheng and Long 2007).

If the Hausman McFadden test for IIA shows that the IIA assumption does not hold, I will be unable to use the regression results from the multinomial logit. I therefore can only use the results from the sequential logit regressions, whereas if the IIA assumption holds, I can use both results of the sequential logit and multinomial logit to provider a better overview of how the quality of care is distributed between different patient groups.

4.3.1.2. Hausman and McFadden (HM) test for IIA

The Hausman and McFadden test checks for independence among the different outcomes by fitting a model with all outcomes (full model), and then fitting the model with one or more of the outcomes removed (restricted model) (Long and Freese 2006). The HM statistic is calculated using the coefficients and variance of the full and restricted models:

$$\boldsymbol{H = \left(\widehat{\beta}_R - \widehat{\beta}_F^*\right)' \{\widehat{var}\left(\widehat{\beta}_R\right) - \widehat{var}\left(\widehat{\beta}_F^*\right)\}^{-1}\left(\widehat{\beta}_R - \widehat{\beta}_F^*\right)} \qquad \text{Equation 9}$$

The test statistic has a chi squared distribution where the degrees of freedom should equal the rows of the restricted model to pass the IIA assumption (Long and Freese 2006).

4.3.1.3. Sequential Logit

Unlike the multinomial logit, the sequential logit regression does not require the IIA assumption to hold as the model uses conditional probabilities. The sequential logistic regression will be used to test associations between the covariates and a selected outcome. This method requires a user to specify a decision tree using the outcomes on the independent variable. Unlike the multinomial logit, this means that outcomes were not independent and therefore estimated probabilities were conditional probabilities (Amemiya 1975; Kahn and Morimune 1979). Also as a decision tree is specified, this method has more flexibility than the multinomial logit (Kahn and Morimune 1979), where I can specify a decision tree based on hospital choice which can be estimated.

Results for the sequential logit regressions compare patients who receive care with patients who failed and were excluded alive from process measures of care, where the probability if receiving care is conditional on not being excluded due to death. The decision tree is illustrated in Figure 7.

<Insert Figure 7>

Our four outcomes are represented as ovals, whereas each decision is represented by a diamond. Each outcome along the process tree is conditional on the previous decision; this is shown by the following equations (Buis 2011):

$$\boldsymbol{p_1 = \Pr(y_1 = 1|x)} \quad \text{Equation 10}$$

$$\boldsymbol{p_2 = \Pr(y_1 = 2|x, y_1 = 1)} \quad \text{Equation 11}$$

$$\boldsymbol{p_3 = \Pr(y_1 = 3|x, y_1 = 1)} \quad \text{Equation 12}$$

$$\boldsymbol{p_4 = \Pr(y_1 = 4|x, y_1 = 1)} \quad \text{Equation 13}$$

Here outcomes 1, 2, 3 and 4 correspond to the outcomes within the oval nodes in Figure 7. I assumed that the decision made by the hospital will be conditional on whether the patient is alive. The decision to provide process measures of care, failing to do so, and exclusions will be made instantly.

The predicted probabilities of the sequential logit and multinomial logit offer different interpretations on coefficients. The multinomial logit compares between outcomes, and the sequential logit uses conditional probabilities. Both techniques will not be used as substitutes and were not directly comparable; however, they will be used as complementary methods. The sequential logit model will estimate the distribution of receiving care compared to not receiving care. The multinomial model compares receiving care to failing to receive care and being excluded from care.

4.3.1.4. The general estimated model

The general estimated model is the model specification of both the sequential and multinomial regressions. The generalised model is used to illustrate what were the outcome variables, and all of the explanatory variables I use for estimations. I estimated condition specific and process measures of care specific regressions for both the multinomial and sequential logit regressions. The model has the following general form and the analysis was conducted using Stata MP4:

$$y_{ixjt} = \beta_0 + \boldsymbol{\gamma x_{qjt}} + \theta\delta_x + \emptyset\boldsymbol{\delta_x t_t} + \boldsymbol{\tau p h_j} + \beta_1 t_t + \beta_2 r_{xj} + u_{jt} + \varepsilon_{iqjt}$$

i. y_{iqjt} is the outcome variable/ variables.
ii. β_0 is the constant term
iii. $\boldsymbol{x_{qjt}}$ is a vector of explanatory variables at the patient level.
iv. δ_q this will be the variable of interest which is the income score.
v. $\emptyset\boldsymbol{\delta_q t_t}$ Interaction variable between time and the income score. These variables will be of most interest as they will show how the distribution of care between income groups has changed over time.
vi. $\boldsymbol{h}$ Hospital level variables such as Foundation Trust status and size.
vii. t_t Time dummy variables which will capture how quality of healthcare has changed over time.
viii. r_{qj} clinical condition. This categorical variable will not be used when estimating each clinical process measure separately.
ix. u_{jt} error component from random variation between patients who receive treatment from different hospitals.
x. ε_{iqjt} these are idiosyncratic errors.

Where i indicates process measures of care, q indicates the patients, j represents Trust and t indicates time.

The error components ε_{ijt} are independent and identically distributed (iid), where the errors have a log Weibull distribution (Dow and Endersby 2004). This means that the shape and nature of the distribution of the error terms will change depending on the variables in the model. The regressions will also cluster the patients to control for correlation between outcome variables for the same patient.

4.3.1.5. Marginal effects

To understand and interpret the estimation results after the multinomial regressions and sequential logit regressions, I estimated the marginal effects of the coefficients from the above model (Wooldridge 2013). Marginal effects must be calculated to be able to interpret results from multinomial logit regressions and sequential logit regressions so that the regression coefficients are on a percentage scale. I interpreted the multinomial regression results by comparing each of the outcomes to whether a patient has received a process measure. For the sequential logit results, I obtained the probability of a patient receiving care conditional on patients not being excluded due to death. The Stata commands for marginal effects are shown in Appendix 2.

4.3.2 Data

In this section, I described the datasets which I have used to create the estimation sample. In total the estimation dataset was generated by combining six separate datasets from several sources.

The hospital activity dataset (AQ 2015), explained in Chapter 2, was my main dataset which was created by members of the Advancing Quality initiative. This was an individual level dataset which contains spell information and personal characteristics. I extracted patients': age; gender; ethnicity; admitted clinical condition; admission date; LSOA of patient residence and Trust of treatment.

The Secondary Users Service (SUS) (HSCIC 2015) explained in Chapter 2, were a collection of raw hospital record extracts. This was the source file from where hospital activity dataset was created. This dataset contains more comprehensive information about a patient spell when compared to hospital activity dataset. However this dataset does not contain a direct patient identifier for Advancing Quality patients. I used this dataset to extract information which is missing from the hospital activity dataset.

Quality Measure Reporter (QMR) (AQ 2015), explained in Chapter 2, was a process measure of care level dataset which describes the incentivised process measures of care each patient received in a spell. This dataset was provided by Advancing Quality and

described whether a patient has received, failed to receive, or been exception reported from a process measure of care. This dataset was directly linked to hospital activity dataset at a spell level.

The Care Quality Commission (CQC) dataset was created using information from the 2007/2008 annual health check report by the Healthcare Commission (CQC, 2008). This hospital Trust level dataset contained information on how well a NHS Trust had performed under set quality guidelines (CQC 2009), and how well the Trust had performed in managing financial resources. These data were provided for financial year 2007/2008. I chose data for CQC before the implementation of Advancing Quality as the impact of Advancing Quality may affect the quality rating of the Trusts.

Monitor dataset was constructed using information for the Monitor website (Monitor, 2014) which holds information on Trusts which have Foundation Trust (FT) status, as well as the date in which the Trust have achieved such status. This information is used to construct a binary variable for FT status which varies over time.

I obtained Trust type from the National Clinical and Health Outcomes Knowledge Base (NCHOD) (NCHOD 2015). I extracted Trust type in four categories: Large; medium; small; and specialist and teaching.

The Index of Multiple Deprivation (IMD) (IMD 2013) is an area level dataset which contains area level deprivation scores. This contains an overall deprivation score as well as all domain, and subdomain, scores which were used to create the overall score. This will assign each patient in the hospital activity dataset with an area level income score at a Lower Super Output Area (LSOA). An LSOA is an area which contains at least 1,000 and maximum of 3,000 individuals, or 400 to 1,200 households.

Each dataset adds more information with regard to a spell on different levels: individual; Trust; and area.

4.3.2.1. Dependent variable

The dependent variable was a categorical variable with four mutually exclusive and exhaustive categories of whether a patient has received a process measure of care or not. The categories are:

1. Excluded from a process measure of care and the patient have died.
2. Excluded from a process measure of care but the patient is still alive.
3. The Trust failed to provide the process measure to the patient.
4. The Trust provided the process measure to the patient.

I had four categories for the outcome variable as I was interested in how quality of care has been distributed between patients. Therefore each outcome will be from the perspective of the health care provider. A healthcare professional under the Advancing Quality scheme would have decided to give care to patients, not give care to a patient or exclude giving care to a patient. Due to this perspective, a difference can be made between how a doctor has excluded a patient, such as the patient is alive or not. If a patient is alive, a healthcare provider could decide to give care, not give care or exclude a patient. However, if a patient had died, the doctor could automatically exclude patients from process measures of care.

This variable was created using both the hospital activity dataset and the QMR datasets. The QMR dataset lists process measures of care a patient has been given, failed to be given, and after the 8^{th} quarter, whether a patient has been excluded from a process measure. I extracted whether a patient has been excluded from process measures of care from identifying spells in hospital activity dataset which do not match QMR for first 8 quarters. I extracted whether a patient has died by using the spell discharge method from the hospital activity dataset.

Our outcome variable was a categorical variable where the outcomes were mutually exclusive. This was because a patient can only have one result from a single clinical process measure. My outcome variable is also exhaustive as there were no other events which the patient can be in. So for a patient in the Advancing Quality dataset for a given clinical process measure, there are no other group that can be included in the outcome variable.

4.3.2.2. Exclusion of patients

Trusts were able to exclude patients from process measures of care. Once patients were excluded from a measure, these patients will not be included in the calculation of overall achievement scores of that process measure. All process measures of care in Advancing Quality share two common exclusion criteria: patients under the age of 18years and whether the patient is a part of a clinical trial. A full list of exclusion criteria and for each process measure is shown in Table 2.

4.3.2.3. IMD variable

From the IMD dataset, I used the income domain for my analysis. This domain measures the percentage of the working population within an LSOA who were either unemployed or have low earnings which were supplemented through income support. The measure was the proportion of people in an LSOA who received: income support; jobseekers allowance; pension credit; child tax credits and asylum seekers on subsistence or accommodation support (IMD 2013).

I converted this income deprivation percentage into an income percentage, where the higher the value for the income score signifies that a higher proportion of people in the patient's area of residence were not receiving low-income benefits. I used the 2010 IMD income deprivation; therefore, the variable is fixed over time.

4.3.2.4. Interaction between income and year

I also included a set of four continuous interaction variables between income score and year. These variables will capture the change in the effect of the income score over time. I will therefore be able to see how the distribution of care for Advancing Quality had changed over time between patients living in areas with different income scores. With interaction terms being used, the coefficient on the income variable represents the income deprivation on the provision of process measures of care in the first year, 1st October 2008 until 31st March 2009 of Advancing Quality.

4.3.2.5. Patient characteristics

I included patients': age; gender and ethnicity. To capture the potential non-linearity between age and provisions of process measures of care, I included an age squared variable. To create this variable I centred the age variable on its mean value. This is creating using the following equation 7:

$$\boldsymbol{Mean\ centered\ age\ squared} = (\boldsymbol{Age} - \overline{\boldsymbol{Age}})^{\mathbf{2}} \qquad \text{Equation 14}$$

I used a mean-centred age-squared variable, as age and age-squared were highly correlated without mean centering.

To account for ethnicity, I used two dichotomous variables: one variable was whether a patient was not of white ethnicity and the other variable was if a patient's ethnicity was not recorded. The base category was white ethnic background. I have controlled for ethnicity using three broad groups due to limited numbers in several categories.

4.3.2.6 Mean Trust achievement

Trust mean achievement proportion was the average achievement proportion of all process measures of care for each individual Trust over all clinical conditions. This variable does not use the excluded population to create the variable. This variable captures the average response of a Trust to the Advancing Quality incentive scheme. I included mean Trust achievement rates to capture how quickly a Trust has adjusted to the Advancing Quality scheme. This variable captured the variation of performance between Trusts. I was interested in within Trust distribution of quality which may be affected by between Trust variations in Quality.

4.4. Results

This section described the descriptive statistics, multinomial logit and sequential logit results.

4.4.1. Descriptive statistics

Descriptive statistics for five categories are now described

4.4.1.1. Outcomes on clinical conditions level

Descriptive statistics are shown in Table 15. I found that the achievement proporions vary from 27% for pneumonia to 87% for hip and knee replacements patients. I found that the proportion of patients failing to receive process measures of care were the lowest when compared to my other outcomes, with the exception of hip and knee replacements where the probability of being excluded due to death was lowest at 0.5%. For AMI, heart failure and pneumonia, being alive and excluded from process measures were the largest outcome group.

<Insert Table 15>

4.4.1.2. Patient characteristics

I found that on average, patients reside in LSOAs where 80% of people were not either unemployed or receiving some form of income support. There were a higher proportion of males in my sample with AMI and heart failure, 62% and 52.4% respectively. There were a higher proportion of females in my sample with hip and knee replacements and pneumonia, 61.3% and 50.2% respectively. I found that patients with heart failure were on average 78 years old, being the oldest average age out of the four clinical conditions. The average age of patients with AMI, hip and knee replacements and pneumonia were 70, 71 and 73 years respectively. On average, 89% of my patient samples were of a white ethnic background.

4.4.1.3. Trust characteristics

Trusts in my sample on average score excellent or good on the CQC quality score. Just 14% of Trusts scored poor. I found that on average most Trusts were scored excellent or good on the CQC financial management rating. Only 25% scored poorly. Most patients (54%) attended Foundation Trusts, and the mean Trust achievement performance in Advancing Quality is 82%.

Patients attending large Trusts and medium sized Trusts were the majority of the population at 38.7% and 32.3% respectively, followed by specialist or teaching Trusts at 18% and small Trusts at 11% of the patient population.

4.4.1.4. Time characteristics

Over time, the number of patients in each year increased in the sample. The first year contained 18.6% of the total number of patients in my sample and the final year contained 26% of the total number of patients in my sample.

4.4.1.5. Outcomes for each process measure of care

Descriptive statistics on outcomes for each of the process measures of care were shown in Table 16.

<Insert Table 16>

The proportion of the population achieving process measures of care for AMI varied from 3.2% for fibrinolytic therapy to 69% for aspirin at arrival. The process measures of care for pneumonia also exhibit a large differential in achievement rates from 8.6% of patients given smoking cessation advice and 60% given oxygenation assessment. Achievement proportion for heart failure also exhibited a large variation from 6% for smoking cessation advice and 66% for evaluation of left ventricle systolic (LVS) function. Process measures of care for hip and knee have a lower variation of achievement rates when compared to the process measures for emergency conditions. The highest achievement proportion for

hip and knee is 89% for venous thromboembolism (VTE) Prophylaxis ordered with the lowest being antibiotics received at 86%.

4.4.2. Multinomial Logit

The multinomial logit regression models a categorical dependent variable which was analogous to simultaneously modelling logistic regressions. I presented the findings from the condition specific results first, and then present the results from process measure of care specific results. For the condition specific results, I presented patient and Trust characteristics, whereas for process measures level regressions, I presented the patient characteristics only.

4.4.2.1. Condition-specific multinomial logit regressions

Patient characteristics for condition specific regression results were shown in Table 17 and the Trusts from the same regression models were shown in Table 18. In this section I will discuss the results from these tables.

<Insert Table 17>

<Insert Table 18>

4.4.2.1.1. AMI

I found that patients from areas with the highest income score had a 0.6 percentage point higher probability of receiving a process measure of care than failing to receive it, when compared to patients from areas with the lowest income scores. This result was statistically significant at the 1% level. From the results of the interaction terms, this relationship was consistent throughout my observation period.

Patients from the highest income score areas had a 6.9 percentage point higher probability of being excluded from a process measure than receiving a process measure when compared to patients from the lowest income score areas. This finding was statistically

significant at the 1% level. Over time, this relationship remained consistent over the estimation period.

I found that the distribution of care is not statistically significantly different between different ethnic groups. On average, as patient's age by one year, patients are: 0.1 percentage points more likely to fail to receive a process measure; 4.9 percentage points more likely to be excluded and die, and were less 4.2 percentage points less likely to be excluded from a process measure when compared to receiving a process measure of care. All findings on patient's age were statistically significant at the 0.1% level.

Patients attending Trusts which received a poor rating of quality from the CQC ratings when compared to patients attending Trusts with an excellent CQC quality rating were: 0.2 percentage points more likely to receive a process measure when compared to failing to receive a process measure and 5.6 percentage points less likely to be excluded then pass a process measure. These results were significant to a 0.1% level.

4.4.2.1.2. Heart failure

I found that in the first year, there was no statistically significant difference between the patients receiving and failing to receive process measures of care for heart failure between patients from areas with different income. I found that patients from highest income score areas were 5.6 and 5.5 percentage points more likely to receive process measures of care than fail in the third and fourth years of Advancing Quality respectively when compared to patients from the lowest income score areas. These results were statistically significant at the 0.1 and 1% level respectively.

Patients from the highest income score areas were 9.8 percentage points more likely to be excluded from process measures of care then receive process measures of care when compared to patients from the lowest income score areas. This is statistically significant at the 0.1% level. This relationship is consistent throughout time.

Patients not from a white ethnic background were 3 percentage points and 2 percentage points less likely to receive process measures of care than failing to receive care and being excluded from a process measure of care respectively. These results were statistically significant at the 0.1% and 1% level respectively. Patients from a non- white

ethnic background were 3.5 percentage points less likely to be excluded and died than being given a process measure of care.

Similar to result for AMI, patients attending Trusts which have a poor CQC quality of care rating were 1 percentage point and 2 percentage points more likely to receive process measures of care than: fail to receive process measures of care and be excluded from process measures of care respectively, when compared to patients who attended hospitals with an excellent CQC quality rating. Results were 1% and 0.1% level respectively.

4.4.2.1.3. Hip and knee replacements

Patients from the highest income score area were 4.2 percentage points less likely to receive process measures of care than be excluded from process measures of care when compared to patients from the lowest income score areas. This difference was statistically significant at the 0.1% level. Over time this relationship changes. In the fifth year, patients from the highest income score areas were 0.7 percentage points more likely to receive process measures of care than be excluded from process measures of care when compared to patients from the lowest income score areas, statistically significant at the 1% level.

On average, a one year increase in a patient's age increases the probability of: failing to receive process measures of care by 0.4 percentage points; excluded from process measures of care by 1.6 percentage points and excluded and died by 0.8 percentage points when compared to receiving process measures of care. All results were statistically significant at the 0.1% level.

Patients not from a white ethnic background are: 0.7 percentage point more likely to receive process measures of care when compared to failing to receive process measures of care, and, 2.3 percentage points more likely to receive process measures of care than excluded from process measures of care. Results were statistically significant at the 1% and 0.1% level respectively.

Patients who attended Trusts with excellent financial CQC rating were more likely to be given process measures of care than failed to be given care by: 4.3 percentage points when compared to poor performing Trusts and 2.9 percentage points when compared to

good performing Trusts, results were statistically significant at the 0.1% level. Patients from Trusts with excellent financial ratings were also more likely to receive process measures of care than be excluded from process measures of care by: 2.9 percentage points when compared to poor financial performing Trusts and 3.4 percentage points when comparing to Trusts with good financial ratings, results were statistically significant at the 0.1% level.

4.4.2.1.4. Pneumonia

Patients from the highest income score areas were 2.7 percentage points more likely to receive process measures of care than fail to receive care when compared to patients from the lowest income score areas, where results were significant to the 0.1% level. This pro rich distribution of care reduces over time. In the fifth year of Advancing Quality, this relationship becomes 0.4 percentage points more likely for patients from highest income score areas. This result is statistically significant at the 1% level.

Patients from the highest income score areas were 5.4 percentage points more likely to be excluded from process measures of care than receive process measures of care when compared to the patients from the lowest income score areas. This result is statistically significant at the 1% level and the relationship does not change over time.

On average, one year increase in patients' age increases the probability of receiving process measures of care by: 0.8 percentage points than failing to receive care and 3.8 percentage points than being excluded alive. These results were statistically significant at the 0.1% level.

Patients who attended Foundation Trusts have a 1.6 percentage point, 1.5 percentage point and 2.2 percentage point decrease probability in receiving process measures of care when care compared to failing to receive care, being excluded and alive and being excluded due to death, respectively. These results were statistically significant at the 0.1% level.

4.4.2.2. Process measures of care specific regressions

I now show the results from the process measures of care specific regressions.

4.4.2.2.1. AMI

Results from the admission and during hospital stay process measures of care are displayed in Table 19. I found that patients' income score area has no statistically significant effect on whether patients receive admission process measures of care: aspirin at arrival and fibrinolytic therapy, when statistical significance is measured at the 5% level. Patients from the highest income score area were 30 percentage points less likely to receive smoking cessation advice than be excluded alive, results were statistically significant at the 0.1 percent level.

<Insert Table 19>

Results from the discharge process measures of care were displayed in Table 20. I found that patients from highest income score areas were more likely to receive process measures of care by: 2.5 percentage points than failing to receive ACEI/ARB for LVSD and 0.9 percentage points than failing to receive aspirin at discharge when comparing to patients from the lowest income score areas. The results were statistically significant at the 1% and 5% level respectively. These relationships did not change over time when statistical significance is measured at the 5% level. Patients from highest income score areas were more likely to be excluded alive than receive process measures of care by: 12 percentage points for beta blocker at discharge and 8 percentage points for aspirin at discharge when compared to patients from the lowest income score areas. These results were statistically significant at the 0.1% and 5% level respectively.

<Insert Table 20>

4.4.2.2.2. Heart failure

Results from process measures of care for heart failure are presented in Table 21. I found that patients from highest income score areas were more likely to be excluded alive than receiving process measure of care by: 6.9 percentage points for evaluation of LVS

function; 10.2 percentage points for ACEI/ARB for LVSD and 16.6 percentage points for smoking cessation advice when compared to patients from the lowest income score areas, where results were statistically significant at the: 5%; 5% and 0.1% level respectively.

Patients from lowest income score areas were 3.4 percentage points more likely to be classed as, failed to be given smoking cessation advice than be given the advice when compared to patients from the highest income score areas, where the result was statistically significant at the 0.1% level. On the other hand, patients who were from the lowest income score areas were more likely to be given discharge instructions by 13.8 percentage points than fail to be given the instructions when compared to patients from the highest income score areas. The results were statistically significant at the 0.1% level.

<Insert Table 21>

4.4.2.2.3. Hip and knee replacements

The results for process measures of care for hip and knee replacements were presented in Table 22. I found that patients from the lowest income score area were more likely to be given than be excluded alive from process measures of care by: 3.7 percentage points for antibiotic selection; 5.8 percentage points for VTE prophylaxis ordered and 5.7 percentage points for VTE prophylaxis received when compared to patients from the highest income score areas. Results were statistically significant at the: 1%; 0.1% and 0.1% level. Over time, these relationships change. In year five I found that there were no differences in exclusion and provision of VTE prophylaxis ordered over patients from income score areas. For VTE prophylaxis received, the relation becomes less pro poor over time. In the fifth year of Advancing Quality, patients for lowest income score areas were 0.4 percentage points more likely to be given care than be excluded when compared to the patients from high income score areas, where results were statistically significant at the 5% level. On the other hand, for antibiotics selection, the initial pro poor relationship becomes a pro rich relationship by 1.3 percentage points, statistically significant at the 5% level.

<Insert Table 22>

Patients from high income score areas were more likely to be given than being excluded through death from process measures of care by: 0.8 percentage points for antibiotic selection and 1.4 percentage points for antibiotics discontinued when compared to patient from low income score areas. Results were statistically significant at the: 5% and 1% level. The distribution of care does not change over my study period for patients from different income score areas when statistical significance is measured at the 5% level.

Patients from the highest income score area were more likely to receive VTE prophylaxis than fail to receive the process measure by four percentage points when compared to patients from the lowest income score area, results were statistically significant at the 0.1% level.

4.4.2.2.4. Pneumonia

The results for process measures of care for patients with pneumonia are shown in Table 23. I found that patients from highest income score areas were more likely to receive than be classed as 'failed' care by: 2 percentage points for antibiotic selection and 11.7 percentage points for smoking cessation advice when compared to patients from the lowest income score areas, results were statistically significant at the: 5% and 0.1% level respectively. These relationships change over time. In the fifth year of Advancing Quality, patients from lowest income score areas were 1.2 percentage points more likely to have antibiotics selected, and, smoking cessation advice become less pro rich by 4 percentage points. The results were statistically significant at the 5% level.

<Insert Table 23>

Patients from lowest income score areas were more likely to receive than be excluded from care alive by: 8.4 percentage points for antibiotics received; and 24.7 percentage points for smoking cessation advice when compared to patients from highest income score areas.

I found that care became pro rich over time for oxygenation assessment. In the fifth year patients from low income score areas were 6.1 percentage points more likely to be given care than be excluded when alive when compared to patients from the highest income score areas. This pro rich finding is also consistent with blood cultures, where in the fifth

year; patients from low income score areas were more likely to be given than excluded through death from care.

4.4.3. Hausman and McFadden IIA test

The results of the Hausman and McFadden tests for independence of irrelevant alternatives are shown in Table 24. I found that from the condition specific regressions: heart failure and pneumonia have passed the tests for IIA, which means that the multinomial logit estimator cannot be rejected for these two conditions.

<Insert Table 24>

From the process measure specific regressions, I found that: fibrinolytic therapy for AMI; ACEI/ARB for LVSD for heart failure; smoking cessation advice for heart failure; antibiotic selection for pneumonia; blood cultures for pneumonia; antibiotics received for pneumonia and smoking cessation advice for pneumonia did not fail the IIA assumption for the multinomial logit.

In total I found that being excluded alive was not an independent and irrelevant alternative in a total of 6 tests, whereas achieved an indicator has failed 11 times. I found for same process measures across conditions such as smoking cessation advice and ACEI/ARB for LVSD, that the outcomes for the process measures across conditions results in differences in how the outcomes were related.

4.4.4. Sequential Logit

I presented the findings from the condition specific results first, and then presented the results from process measure of care specific results. For the condition specific results, I presented patient and Trust characteristics, whereas for process measures level models, I presented the patients characteristics only.

4.4.4.1. Condition specific regressions

The results from the condition specific sequential logit models were presented in Table 25. I referred to this table for all results in this sub section.

<Insert Table 25>

4.4.4.1.1. AMI

I found that patients from the highest income score area have a 4.1 percentage point increase in receiving process measures of care when compared to patients from the lowest income score area. This result was consistent over time and was statistically significant at the 1% level. In year four, receiving process measures of care for AMI became more pro poor by an additional 6.1 percentage points.

Patients who were older by one year were 0.7 percentage points more likely to receive process measures of care; this relationship was statistically significant at the 0.1 percent level. Patients from not a white ethnic background were 0.7 percentage points more likely to receive process measures of care. This was statistically significant at the 5% level.

Patients attending Foundation Trusts were 2.8 percentage points less likely to receive process measures of care. This was statistically significant at the 0.1 percent level. Patients which attend Trusts with poor CQC rating were 6.7 percentage points more likely to receive care when compared to patients who attended Trusts with an excellent CQC rating. These results were statistically significant at the 0.1 percent level.

4.4.4.1.2. Heart failure

Patients from the lowest income score area had a 10.2 percentage point higher probability of receiving process measures of care when compared to patients from the highest income score area; the result was statistically significant at the 0.1% level. This finding is consistent over time with the exception of year four when this relationship changes to pro

rich where patients from the highest income score area had a 1.7 percentage point increase in probability receiving care when compared to patients from the lowest income score area. This result was statistically significant at the 0.1% level.

Patients' probability of receiving care decreased over time with an average of 3.5 percentage points for an increase of one year of age, from the age squared variable, I found that this decrease in probability increases as patient's age, these results were statistically significant at the 0.1 percent level.

Patients attending Trusts with excellent CQC quality scores were less likely to receive process measures of care by: 2.6 percentage points compared to patients from Trusts rated good and 4.3 percentage points from Trusts rated poor. On the other hand, patients attending Trusts with excellent CQC financial rating have a 1.9 percentage point higher probability of receiving process measures of care from either: patients attending Trusts with good or poor CQC financial ratings. All results from CQC ratings were statistically significant at the 0.1% level.

4.4.4.1.3. Hip and knee replacements

Patients from the lowest income score area were 3.2 percentage points more likely to receive process measures of care in the first year when compared to patients from the highest income score area. This pro poor relationship became a pro rich relationship over time. In the fifth year patients from the highest income score area were 1.7 percentage points more likely to receive process measures of care when compared to patients from the lowest income score areas.

Similar to the results from heart failure, on average, patients probability of receiving care reduced by 3 percentage points as patients age by one year where the probability decrease in getting care increases as patients become older, results were statistically significant at the 0.1% level.

Patients that attended Trusts with poor CQC quality rating had a 2.7 percentage point increase in probability of receiving care when compared to patients that attend Trusts with excellent CQC rating. Patients that attend Trusts with excellent CQC financial ratings had higher probability of receiving process measures of care by: 7.2 percentage points when

compared to poor CQC financial ratings and 6.3 percentage points when compared to Trusts with good CQC financial rating. Results were statistically significant at the 0.1 percent level.

4.4.4.1.4. Pneumonia

Patients from the lowest income score area were 5.3 percentage points more likely to receive process measure of care when compared to patients from the highest income score area; this result was statistically significant at the 0.1 percent level. This result was consistent over time.

Patients that attended Trusts with poor CQC quality rating were more likely to receive process measures of care by 0.7 percentage points when compared to Trusts with an excellent CQC rating where the result is statistically significant at the 5% level. Patients that attend Trusts that perform excellent in the CQC financial ratings have a higher probability of receiving care by: 3.9 percentage points when compared to poor CQC financial rating; and 1.8 percentage points when compared to Trusts with good CQC financial rating. Results were statistically significant at the 0.1 percent level.

4.4.4.2. Process measures of care specific regressions

The following section presents the results from the process measures of care specific sequential logit regressions.

4.4.4.2.1. AMI

Results from the process measures of care specific regression for AMI are shown in Table 26. I found no statistically significant relationships between admission process measures of care aspirin at arrival and fibrinolytic therapy, and patients' area income when statistical significance is measured at the 5% level. These results were also consistent throughout time when statistical significance was measured at the 5% level.

<Insert Table 26>

Patients from the highest income score area had a 26.6 percentage point lower probably to receive smoking cessation advice when compared to patients from areas with the lowest income score area. This is statistically significant at the 0.1% level. This finding is also consistent throughout time when statistical significance is measured at the 5% level.

4.4.4.2.2. Heart failure

Results from the process measures of care specific regression for heart failure are shown in Table 27. I found that patients from lowest income score areas have a higher probability of receiving process measures of care by: 10.5 percentage points for ACEI/ARB for LVSD; 20.1 percentage points for discharge instructions; and 10.6 percentage points for smoking cessation advice, when compared to patients from the highest income score areas. These results were statistically significant at the: 5%; 0.1%; and 0.1% level, respectively.

<Insert Table 27>

For smoking cessation advice, the distribution of care by income score area remains the same when statistical significance is measured at a 5% level. I found that for discharge instructions that the differential in quality of care between patients from lowest income score area 3.1 percentage points more likely receive the discharge instructions when compared to patients from the highest income score areas.

4.4.4.2.3. Hip and knee replacements

Results from the process measures of care specific regression for hip and knee replacements are shown in Table 28. I found that initially patients from the highest income score area have a 6.1 percentage point lower probability of having antibiotics

selected when compared to patients from the lowest income score area. This result however, goes from pro poor to pro rich. In the fifth year, patients from the highest income score area have a 6.6 percentage point higher probability of having antibiotics selected when compared to patients from the lowest income score area. The results were statistically significant at the 0.1 percent level.

<Insert Table 28>

Patients from lowest income score area have a 5 percentage point higher probability of having VTE prophylaxis ordered when compared to patients from highest income score areas. This result was statistically significant at the 1% level.

4.4.4.2.4. Pneumonia

Results from the process measures of care specific regression for pneumonia are shown in Table 29. I found that patients from the lowest income score area have a higher probability of receiving process measures of care by: 6.6 percentage points for oxygenation assessment; 13.2 percentage points for antibiotics received; and 17 percentage points for smoking cessation advice when compared to patients from the highest income score area. Results were statistically significant at the: 5%; 0.1%; and 0.1% level, respectively. The relationship between income score area and the provision of blood cultures remain consistent throughout time when statistical significance was measured at the 5% level. For antibiotics received, the relationship becomes less pro poor from 13.2 percentage points in the first year to 4.8 percentage points in the fifth year. For oxygenation assessment, the initial pro poor relationship changes to become pro rich where patients from the highest income had a 2.1 percentage point increase in probability of receiving the quality metric when compared to patients from the lowest income score area.

<Insert Table 29>

Patients from the highest income score areas have a 7 percentage point increase in probability of having antibiotics selected when compared to patients from the lowest income score areas. Results were statistically significant at the 5% level. This relationship was consistent throughout time when statistical significance was measured at 5% level.

In year one, there was no relationship between a patient's income score area and the provision of blood cultures. However this result became pro rich over time. Patients from the highest income score area had an 11 percentage point increase in the probability of receiving blood cultures when compared to patients from the lowest income score area. Results were statistically significant at the 1% level.

A summary of the results for the effect of patient area income in the first and fifth year of Advancing Quality for all condition specific and processes measures of care specific regression are shown in Table 30.

<Insert Table 29>

4.5. Discussion

For all emergency conditions I found that quality of care is distributed pro poor and this was consistent over time. Patients from lower income score areas were more likely to receive process measures of care than be excluded. For elective conditions, care was initially distributed pro poor but this relationship changed to become pro rich over time. I found that patients from lower income score areas have a higher probability of being excluded alive in the first year, and, had a lower probability of being excluded in the fifth year when compared to patients from higher income score areas.

Quality of care given under emergencies was distributed initially pro poor but became pro rich over time. Patients from higher income score areas were less likely to be excluded through death, more likely to be excluded alive and more likely to receive care. The pro rich distribution of care could be case mix. Patients from lower income score areas had higher levels of co-morbidities and disease severity (Neuburger et al. 2013). Higher disease severity may lead to higher probably to death and higher probability of exclusions.

Process measures of care based on advice were distributed pro poor in both the first and fifth year. The level of inequity falls in the fifth year to become less pro poor. I found that patients from lower income score areas were more likely to receive care than be excluded or fail. The finding for smoking may be due to higher prevalence of smoking for patients in lower income score areas (Haustein 2006; Borland and Rudolph 1975). This result may be explained from the hypothesis that patients from lower income score areas had a higher potential health benefit from advice.

Process measures of care relating to drugs and tests had a weakly pro poor distribution which does not change over time. This result is driven from patients being excluded alive for patients receiving ACEI/ARB for LVSD. Most process measures of care relating to tests and drugs did not exhibit differentials in the quality of care.

Process measures of care related to providing and stopping antibiotics for elective admissions exhibit no effect on equity. Care processes such as ordering had an initial pro poor distribution which changes to become pro rich over time.

4.5.1. How findings relate to the literature

I found opposite effects from literature examining income score area and provision of care from general practitioner practices. My overall pro poor finding was inconsistent with the pro rich finding from primary care (Ashworth et al. 2011; Dixon, Khachatryan, and Tian 2012; Doran et al. 2008; Ashworth et al. 2007; Millett et al. 2007; Saxena et al. 2007).

Crawley et al. (2009) used patient level survey data and found that there were no differences in distribution of care based on socioeconomic deprivation. This agrees with my findings when process measures of care such as drugs and test were administered with no emergency pressure. The findings indicate that healthcare providers may not inequitably distribute quality of care to patients, given patients have equal need.

Our findings were consistent with the findings of Jha, Orav, and Epstein (2010). Jha, Orav, and Epstein (2010) examined the distribution of quality of care between hospitals serving higher or lower proportion of income deprived patients. The authors found that

over time, the quality of care serving proportionally high numbers of income disadvantaged patients increased overtime.

4.5.2. Strengths of this study

The unique dataset which I have used allowed us to research each individual process measure of care at patient level. Due to the variety of process measures incentivised under Advancing Quality, I were able to test various hypotheses which highlight areas where care may be inequitable. In this paper I have been able to highlight which areas within which conditions and process measures were driving variations in the quality of care.

A key strength of this study was having data on which patients were excluded from process measures of care. If I assumed that patients were being excluded from process measures of care for valid reasons, then I would be able to find whether the quality of care was distributed evenly for patients with even need.

I was also able to test between elective and emergency care. My findings from hip and knee process measures found that patients from lower income score areas have received more care than patients from higher income score areas. This finding was not consistent with Ryan et al. (2009) and Jha, Orav, and Epstein (2010) and is opposite from (Verlinde et al. 2012). This finding suggests that even within secondary care, similar drug prescribing process measures can vary in quality across different patient groups.

4.5.3. Limitations of this study

The results of this study may not be generalisable to other healthcare systems due to the universal coverage of healthcare in England. As found by Sutton et al. (2012) results from the US did not translate over to the UK in terms of the effect of the scheme on mortality. As process measures of care have an added complication in that Trusts within Advancing Quality have different policies regarding the implementation of clinical care. I may not

find the same result of distribution of care to be the same within the UK where all citizens have full national health insurance.

Another data issue which limits this study was the unavailability of patient level deprivation within hospital administrative data. Patient level income deprivation was available using survey or census data; however this improvement in data quality for a patient's income was negated by the reduction in data quality regarding a patient being a part of Advancing Quality. If I used census or survey data, I would not have been able to conduct my analysis using my methods and my sample will not be patients who were under Advancing Quality.

4.5.4. Policy implications

Quality incentive schemes may address socioeconomic care differentials by supplying care in a pro poor manner which is driven by severity and an incentive schemes exclusion criterion. I found that this pro poor distribution of care did not come at the expense of process measures of care which directly affect a patient's health outcome. I found that the care was equitable.

Quality incentive schemes may need to be tailored based on the patient population of each Trust. Trusts which provide care for patients from poorer areas may provide more advice than Trusts which serve a higher proportion of patients from richer areas. This creates an imbalance of work for each Trust when achievement benchmarks were used.

4.6. Conclusion

Overall I found that quality of care was distributed pro poor and over time the pro poor distribution reduced, however process measures of care were distributed equitably.

For process measures of care under an emergency, patient condition severity may be driving the differences in quality of care between patients from different income groups as results were driven by patients being excluded. Advice for patients was pro poor and

may be driven by both demand and supply (Goddard and Smith 2001). Patients from poorer areas were more likely to smoke and therefore require more from the hospital. Trusts may increase the supply of advice for patients from poorer areas due to higher expected health gains.

Process measures of care that were not provided in an emergency situation and were directly related to improving health outcomes such as antibiotics, drugs at discharge and tests, were distributed equitably. This implies that when a patient's condition severity no longer impacts on the probability of mortality and patients have an equal probability of health gain, the quality of care provided is equitable.

When I look at the results of the multinomial logit regressions which did not fail the IIA assumption tests, I found overall that income score area has no impact on probability of receiving care with the exception of giving advice.

In the current chapter, I found that the incentivised quality of care was overall distributed pro poor. Incentivised quality of care differentials between patients from different socioeconomic groups were driven by the perceived health gains and patient severity. Quality of care under no emergency with direct impact on health was distributed equitably. This was in line with the A priori hypothesis, where I would expect under a financial incentive scheme that the process measures of care should be distributed equitably. In the next chapter I examined the distribution of incentivised quality of care over days of the week.

5. Is the weekend effect on hospital mortality attributable to lower quality of care?

The patient mortality for weekend admissions has been found to be higher than for weekday admissions for emergency care. This was commonly known as the weekend effect (Cram et al. 2004). This finding was seen worldwide, in countries including Canada (Laupland, Ball, and Kirkpatrick 2009; Saposnik et al. 2007), United States (Kostis et al. 2007; Sharp, Choi, and Hayward 2013), Australia (Coiera et al. 2014; Concha et al. 2014) and the UK. Possible explanations for higher mortality at weekends were reduced senior staffing levels (Dr Foster Intelligence 2011) and reduced services, such as a lower provision of diagnostic tests and invasive procedures (Kostis et al. 2007).

Research into the weekend effect has become of increasing interest in the National Health Service (NHS) in England. Plans to put greater emphasis on providing patients with more consistent levels of care by offering full hospital services 24 hours a day, seven days a week (NHS 2013). These service changes were currently taking place. The current medical director of the NHS, Sir Bruce Keogh, created plans to implement seven day services in hospitals (NHS 2013) after he co-authored a paper finding a weekend effect on mortality (Freemantle et al. 2012).

To date there have been no published studies examining the distribution of process measures of quality of health care throughout the days of the week using data at the patient level. As process measures of care were under the direct control of hospitals, establishing if there is a weekend effect on measurable process measures of care will help explain the weekend effect. Finding if quality of care differs in the weekend could potentially help hospitals and policy makers improve weekend care.

In the next sections I will: present a literature review by examining literature grouped by country of study; introduce the Advancing Quality scheme and the clinical conditions I study; explain the data and how I created my variables; show the estimation methods; and present the results followed by a discussion.

5.1. Literature review

The aim of this literature review was to find the relationship of the weekend effect on both mortality and quality of health care from hospitals. To find research papers on the weekend effect on mortality and care quality, I used two methods to search for initial papers; Google Scholar and Web of Science. I searched for the terms "weekend effect", "hospital", "quality" and "mortality". After finding relevant research papers, I further identified additional papers by conducting a forward and backward search on citations and the citing papers. I found 33 empirical studies on the weekend effect examining a variety of health conditions and outcomes published between 2004 and 2014.

All of the studies that I found used patient level data and used a variety of statistical and econometric methods, including logistic regression, survival analysis, Ordinary Least Squares (OLS), negative binomial, chi squared tests and Cox models. I pay particular attention to: studies which were condition-specific or of combinations of conditions; the outcome variables that were used; and whether AMI or pneumonia was a clinical condition used in the study. I draw particular attention to AMI and pneumonia, as these were the clinical conditions that will be the focus of a subsequent study.

I reviewed the literature outside of England and then by number of studies within each country in descending order. I then reviewed studies from England. I grouped studies by country as the weekend effect may be dependent on the organisation of health system. I also compared studies focusing on one condition with studies grouping multiple conditions.

5.1.1. United States

Thirteen studies on the weekend effect were from the US (Kostis et al. 2007; Becker 2007; Hoh et al. 2010; Dorn et al. 2010; Echaiz et al. 2014; Groves et al. 2014; Horwich et al. 2009; James et al. 2010; Sharp, Choi, and Hayward 2013; Cram et al. 2004; Needleman et al. 2002; Schilling et al. 2010; Goldstein et al. 2014). These ranged from: studies of specific conditions (Kostis et al. 2007; Becker 2007; Hoh et al. 2010; Dorn et al. 2010; Echaiz et al. 2014; Groves et al. 2014; Horwich et al. 2009; James et al. 2010); studies of multiple conditions (Sharp, Choi, and Hayward 2013; Cram et al. 2004; Needleman et al. 2002; Schilling et al. 2010); and specific studies of children (Goldstein et al. 2014).

Within the condition-specific studies, two studies focus on AMI. Kostis et al. (2007) used 15 years of data at patient level from 1987 to 2002. In-hospital mortality was the outcome variable and the estimation method was logistic regression. Becker (2007) also looked at AMI as the sole condition using data from 1989 to 1998. The three cardiac related procedures were: cardiac catherisation; percutaneous transluminal coronary angioplasty, and coronary artery bypass graft surgery. Kostis et al. (2007) and Becker (2007) used linear regression models. Kostis et al. (2007) found that mortality on the weekend was higher and fewer invasive surgeries were performed. This supported the finding by Becker (2007), who found that patients admitted on weekends were less likely to receive immediate cardiac procedures.

The other condition-specific studies focused on: stroke (Hoh et al. 2010); upper gastrointestinal haemorrhage (Dorn et al. 2010); E.coli and urinary tract infections (Echaiz et al. 2014); ruptured aortic aneurysms (Groves et al. 2014); heart failure (Horwich et al. 2009) and acute kidney injury (James et al. 2010). All studies used mortality within the spell as the outcome of interest. Dorn et al. (2010), Groves et al. (2014), Horwich et al. (2009) and James et al. (2010) found that weekend admissions were associated with higher in-hospital mortality than weekday admissions. However, Hoh et al. (2010) and Echaiz et al. (2014) did not find a difference in mortality between weekend and weekday admission. No studies examined process measures of care as a measure of quality to explain the difference in mortality.

Goldstein et al. (2014) analysed the weekend effect on urgent surgery for children. They used in-hospital mortality as the main outcome variable and logistic regressions. They found that mortality during weekend admissions was higher than admission during weekdays. This study however, may not be generalisable to the whole population due to the sample of patients not including adults, where results from studies contain only children may not be applicable to adults.

Three studies looked at six or more conditions. Only one study conducted condition specific regressions (Cram et al. 2004). Sharp, Choi, and Hayward (2013) examined all emergency conditions. Cram et al. (2004) examined 50 emergency conditions including AMI and pneumonia. Schilling et al. (2010) examined six common conditions: AMI; pneumonia; heart failure; hip and knee; stroke; and gastrointestinal bleeding. All papers used in-hospital mortality as the main outcome and all used a logistic regression approach. These three papers, (Cram et al. 2004; Schilling et al. 2010; Sharp, Choi, and

Hayward 2013), found that mortality during weekend admissions was higher than during weekday admission. However, when examining condition specific regressions, Cram et al. (2004) found that patients admitted to hospital on weekends with AMI had a higher probability of mortality with an odds ratio of 1.09.

Needleman et al. (2002) used data on hospital discharges and staffing. The outcome used was in-hospital mortality and the study adopted hospital level analysis using negative binomial analysis and ordinary least squares regressions. The authors found that there was a negative correlation between nursing staff and in-hospital mortality. However, the authors found no difference between weekend and weekday admissions in mortality, or in staffing levels.

5.1.2. Canada

Four studies were from Canada (Suissa et al. 2014; Bell and Redelmeier 2001; Laupland, Ball, and Kirkpatrick 2009; Saposnik et al. 2007). One study used 17 years of data from all hospitals from the province of Quebec in Canada (Suissa et al. 2014). One study used 10 years of data from all emergency departments from Ontario (Bell and Redelmeier 2001). One study used four years of data on four hospitals (Laupland, Ball, and Kirkpatrick 2009). One study used one year of data from all stroke centres in Canada (Saposnik et al. 2007).

Three studies look at the mortality on day of admission. (Laupland, Ball, and Kirkpatrick 2009; Bell and Redelmeier 2001; and Saposnik et al. 2007). These three studies all used logistic regression. Saposnik et al. (2007) found no differences in complications from operations, but found a higher mortality for weekend admissions. Bell and Redelmeier (2001) found that there was a weekend effect on mortality for 23 out of 100 conditions including AMI. However Laupland, Ball and Kirkpatrick (2009) found that there was no difference in patient mortality. Laupland, Ball and Kirkpatrick (2009) also extended the analysis by having the time of the day as well as the days of the week. When testing all conditions together, the authors did not find differences in mortality for patients admitted on different times of the days and days of the week.

Suissa et al. (2014) is the only study found which looks at pneumonia. However, the authors focussed on hospital stays over the weekend rather than examining admissions over the weekend. The authors found that patients staying over weekends and Fridays had a higher mortality.

5.1.3. Australia

Three studies were from Australia (Coiera et al. 2014; Concha et al. 2014; Asha, Titmuss, and Black 2011). Two studies use seven years of data from 501 hospitals in Australia from 2000 to 2007 (Coiera et al. 2014; Concha et al. 2014). One study used data from one Australian hospital from 2006 to 2008 (Asha, Titmuss, and Black 2011).

Coiera et al. (2014) used data from all admissions and conducted survival analysis. Concha et al. (2014) used data from all emergency admissions. Coiera et al. (2014) and Concha et al. (2014) used in and out of hospital mortality for weekend and weekday admissions. Coiera et al. (2014) and Concha et al. (2014) found that mortality for weekend admissions were higher than admissions during weekdays. Coiera et al. (2014) and Concha et al. (2014) did not control for Trust level characteristics or capture quality of care in the study.

Asha, Titmuss, and Black (2011) looked at the differences in mortality and length of stay for all emergency admissions by the time of the day, focussing on out of hours care and non-out of hours care. The findings were similar to the study conducted by Laupland, Ball, and Kirkpatrick (2009) in Canada, that there was no difference in mortality or length of stay by whether patients were admitted out of hours.

5.1.4. Other Countries

Three other studies were identified using: one year of data from 2010 from Iraq (Al-Asadi and Kadhim 2014); four years of data from one hospital in Spain (Barba et al. 2006); and 10 years of data from all hospitals in Italy (Gallerani et al. 2012). All studies used logistic regressions. One study also used Kaplan-Meier survival analysis (Gallerani et al. 2012)

Al-Asadi and Kadhim (2014) focused on AMI as the sole condition and in-hospital mortality as the outcome variable. The authors found that the mortality during weekend admissions was not higher than admissions on weekdays. However, unlike the studies by Laupland, Ball, and Kirkpatrick (2009) and Asha, Titmuss, and Black (2011), the authors found that patients admitted out of hours have a higher mortality than patients admitted during usual operating hours.

Barba et al. (2006) analysed all emergency conditions and used two-day mortality and a 'global' mortality, which was not defined. Gallerani et al. (2012) analysed acute aneurysm rupture or dissection with in-hospital mortality. Both studies found that patients admitted on weekends have a higher mortality than weekday admissions. Barba et al. (2006) only found the weekend effect on mortality when taking 2 day mortality into account. The weekend effect on mortality is not found when using a global mortality measure.

Findings from Al-Asadi and Kadhim (2014), Barba et al. (2006) and Gallerani et al. (2012) found that there is no definitive weekend effect on mortality. All three papers do not capture quality and therefore, do not provide any explanation to why there is or is not a weekend effect on mortality.

5.1.5. Scotland

Three studies were from Scotland (Schmulewitz, Proudfoot, and Bell 2005; Smith et al. 2014; Handel et al. 2012). One study used one year of data from one hospital (Schmulewitz, Proudfoot, and Bell 2005). One study used two years of data from one hospital (Smith et al. 2014). One study used ten years of data from all hospitals (Handel et al. 2012). All three studies used logistic regressions.

Schmulewitz, Proudfoot, and Bell (2005) analysed the weekend effects for six emergency conditions, including pneumonia, cerebrovascular accidents, chronic pulmonary disease, pulmonary embolism, upper gastrointestinal bleeds, and syncope and collapse on in-hospital mortality, length of stay and readmissions. The authors found that there was no weekend effect on mortality, length of stay or readmissions after six months from operation. The authors found that mortality for pneumonia is lower during weekend

admission when compared to admissions during weekday. This finding however, is specific to the hospital of study, Edinburgh Royal Infirmary during 2001.

Two more recent studies by Handel et al. (2012) and Smith et al. (2014) analysed the weekend effect on in-hospital and out of hospital mortality for all emergency conditions combined. They found that mortality was higher on the weekend admissions compared to the weekday admissions. The 11 year sample period from 1999 to 2009 of Handel et al. (2012) found that the differential between mortality of weekend and weekday admissions was falling over time. The study by Handel et al. (2012) controlled for a variety of patient factors. However, it did not control for any hospital level variables such as size and volume.

Smith et al. (2014) assessed the effect of admissions during public holidays on mortality as well as admissions during weekends. The authors found that there were higher mortality for patients admitted on public holidays when compared to non-public holidays. Smith et al. (2014) did not combine admissions on weekend with admissions on public holidays. Combining weekend and public holiday admissions is reasonable as hospital services on public holidays were expected to be similar weekends (Phillips et al. 2004)

5.1.6. England

Seven studies were conducted using English data (Aylin et al. 2010; Freemantle et al. 2012; Mohammed et al. 2012; Aylin et al. 2013; Bray et al. 2013; Ferguson et al. 2014; Jairath et al. 2011). Four studies used data from Hospital Episode Statistics (HES): for the years 2005 to 2006 (Aylin et al. 2010); for the years 2009 to 2010 (Freemantle et al. 2012); for the years 2008 to 2009 (Mohammed et al. 2012); and three years of data from 2008 to 2011 (Aylin et al. 2013).

Three studies analysed condition specific datasets: stroke (Bray et al. 2013); appendicectomy (Ferguson et al. 2014); and acute upper gastrointestinal bleeds (Jairath et al. 2011). These three studies all used mortality: in-hospital mortality (Ferguson et al. 2014; Jairath et al. 2011); in and out of hospital mortality (Bray et al. 2013); and other adverse events (Ferguson et al. 2014). Using logistic regression (Ferguson et al. 2014; Jairath et al. 2011) and survival analysis (Bray et al. 2013), all three studies found no

correlation between mortality and the weekend admissions. Ferguson et al. (2014) also found no weekend effect on adverse events. Process measures of care were not assessed in these studies.

Aylin et al. (2013) studied the weekend effect on 30 day in-hospital and out of hospital mortality for all elective admissions using logistic regressions. They found for all elective admissions assessed together that mortality on the weekend were higher than weekdays. They extended their analysis to high risk procedures, and analysed them together and separately. From the condition specific regressions, weekend mortality was higher for: excision of colon; repair of abdominal aortic aneurysm; and excision of lung. When analysing high risk procedures together, the authors did not find a weekend effect. The issue with analysing weekend admissions for elective surgery is the case mix of patients, as I would expect hospitals to be selecting patients to treat at weekends. The authors note that 48 hours after surgery is the time where most complications will occur (Aylin et al. 2013) and therefore will select patients who were easier to treat on the weekends. Hospital Trusts selecting patients to treat creates a sample bias for weekends. This may be the reason why high risk procedures do not exhibit a weekend effect on mortality whereas overall elective admissions do.

The three studies in England which have assessed the weekend effect on all emergency admissions were Mohammed et al. (2012), Aylin et al. (2010) and Freemantle et al. (2012). Two studies used logistic regression (Aylin et al. 2010; Mohammed et al. 2012) and one used hazard proportional Cox models (Freemantle et al. 2012). All three sets of authors found that there was an increase in the mortality if patients were admitted on the weekend when compared to the weekday. Aylin et al. (2010) and Freemantle et al. (2012) conducted analysis of specific conditions. Aylin et al. (2010) and Freemantle et al. (2012) found that mortality was higher on weekend admissions for AMI but not pneumonia. Mohammed et al. (2012), Aylin et al. (2010) and Freemantle et al. (2012) did not account for the volume of patients presenting themselves at Accident and Emergency who were then admitted to receive care.

5.1.7. Summary of the literature

From reviewing these 33 studies, there seems to be mixed evidence on the weekend effect. I found that when studies aggregate conditions together, they found a weekend effect. However, when conditions were assessed separately, there does not seem to be a weekend effect for pneumonia and the results from assessing AMI were mixed worldwide. However, studies based in England indicate that there is a weekend effect for AMI.

5.2 Aim

The aim of this study was to find whether the quality of health care, process measures of care, under Advancing Quality exhibits a weekend effect. I also aimed to examine if there was a weekend effect on mortality for patients in the North West of England, and if so, were there variations in the quality of health care received, process measures of care.

Using the Principal Agent theory (Ross 1973), I hypothesised that the quality of health care under an Advancing Quality should not differ over the days of the week as hospitals were being financially incentivised to provide process measures of care. This means that hospital may exert more effort to providing patients process measures of care throughout all days of the week.

I defined weekend as calendar days, Saturday and Sunday in my proposed study. I did not have patients' admission time in the data which means out of usual hours of care could not be analysed.

5.2.1. Potential contributions to literature

I identified three potential contributions this study will add to the literature:

1. Our main potential contribution was to test whether quality of care was lower for patients admitted on weekends. This built on the literature using mortality as the main outcome variable, as mortality was not solely determined by activities within the hospital, and therefore the weekend effect on mortality may not be attributed fully to the hospital. Conversely, process measures were more under the control of hospitals. Other papers have used different quality measures such as: adverse events (Needleman et al. 2002; Ferguson et al. 2014) which were not necessarily developed on the weekend; and hospital financial data as a proxy of quality (Dorn et al. 2010).

However, weekend admissions may exhibit higher costs due to longer inpatient stays as patients admitted on weekends were unlikely to be discharged on the weekend.

2. I proposed to analyse the bank holiday effect as well as the weekend effect. This work has been conducted by Smith et al. (2014) who compared public holidays to non-public holidays. I proposed to add to this literature by also comparing weekends and public holidays to non-weekend and non-public holidays. This distinction was made as I expected hospital services on public holidays to be similar to weekends (Phillips et al. 2004).
3. I proposed to use two additional variables that may capture aspects of patient case mix. I included volume of AMI and pneumonia patients, as weekend admissions were lower in numbers compared to the weekdays. A lower number of admissions may indicate that patient case mix was different between admission to hospital on weekdays and weekends. The second variable was the distance between the patient's residential area and hospital site. Distance was seen as a barrier to hospital care, where the further away a patient lives, the greater the barrier of care from an emergency department. The case mix variables contained in administrative data may under represent the true case mix of patients (Bottle et al. 2014).

I did not test whether the proportion of patient exclusions fluctuate by day of the week. I assume that the level of exclusions through the days of the week is constant. This research potentially improves on previous research by using individual level process measures of care to examine whether quality of care exhibits a weekend effect without examining exclusions also.

5.3. Methods

I modelled the relationship between the receipt of process measures of care and day of the week. I performed a logistic regression and, after modelling, I obtained average marginal effects (Baum 2006) to aid interpretability of the results. In order to test whether there was a mortality weekend effect for AMI and pneumonia patients in Advancing Quality, I perform survival analysis and logistic regression. I did not examine out of hours care as admission time is not routinely collected in Secondary Users Service, and therefore we

were unable to attribute time of admission for each patient. Days of the week and weekends were strictly defined using calendar days.

5.3.1. Econometric Models

I used survival analysis to analyse survival probability based on admission day, weekday or weekends, and bank holidays. I then used logistic regressions to examine if there was a difference in the provision of process measures of care on the weekend. I also used logistic regression to examine what explains the weekend effect on mortality.

5.3.1.1. Survival analysis

The aim of this analysis was to find where the probability of survival was different between weekdays and weekends. To conduct survival analysis, I used data from the SUS extract which contains all subsequent hospital admissions for patients who have been admitted as an Advancing Quality patient. This allows me to track a patient through time and capture any death in any hospital in the North West of England in the follow-up period. Then I produced a Kaplan-Meier (Kaplan and Meier 1958) curve with in-hospital death as the failure outcome. I plotted the survival probability over time stratified by: days of the week; weekdays and weekends with bank holidays included; weekdays with weekends; and weekdays, weekends and bank holidays separately.

Due to having spell level data, a patient could have multiple admissions to hospital between, being admitted as an Advancing Quality patient and another spell which ends in the patient dying. As there were multiple admissions between the initial spell and the last spell, the multiple records may complicate the findings. Therefore I did not keep the spells between the first and last spell. This means that my survival curves represent survival after the first patient spell, taking the day of admission from the first patient admission. I assumed that the occurrence of subsequent admissions is potentially affected by care at the initial stay.

If patients died out of hospital, I was unable to attribute the death within my dataset as I could not follow patients to track health outcomes outside of a hospital in the North West of England. If patients were not discharged as dead and were not observed in the data, I assumed this patient was alive.

5.3.1.2. Logistic regression

The aim of this analysis was to examine whether the provision of process measures of care differed by days of the week. This analysis was conducted on a univariate and full models.

Our dependent variables were process measures of care and in-hospital patient mortality. My main explanatory variables were days of the week. I also controlled for known covariates. The model I estimated is given by the following general equation:

$$\boldsymbol{y_{ijt} = \alpha + \beta X_{ijq} + \gamma p_{ijq} + \emptyset v_t + h_j + t_q + e_{ijq}} \quad \text{Equation 15}$$

In equation 16, *i* denotes spell, *j* denotes Trust, whereas *q* denotes time. y_{ijt} is a dichotomous outcome variable for a process measure of care or mortality. α is the constant term; $\boldsymbol{X_{ijt}}$ is a vector of days of the week. $\boldsymbol{\beta}$ are the coefficients of interest. $\boldsymbol{p_{ijt}}$ is a vector of patient characteristics: age; sex; ethnicity; distance to hospital; and Elixhauser comorbidities. γ is a vector of coefficients on patient characteristics. v_t is the volume of Advancing Quality patients admitted to North West Trusts stratified by AMI and pneumonia. $\emptyset$ is the coefficient on volume. h_j are hospital variables: hospital type; volume; foundation Trust status; and Quality of care commission quality of care. t_q are quarterly 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 ($\pi/3$).

I also modelled a univariate models which is given by the following equation:

$$\boldsymbol{y_{ijt} = \alpha + \beta X_{ijq} + e_{ijq}} \quad \text{Equation 16}$$

In equation 17, *i* denotes spell, *j* denotes Trust, whereas *q* denotes time. y_{ijt} is a dichotomous outcome variable for a process measure of care or mortality. α is the

constant term. $\boldsymbol{X_{ijt}}$ is a vector of days of the week and e_{ijt} is the error term which is assumed to be *iid* with a mean of zero and a constant variance ($\pi/3$).

I also ran models without individual days of the week variables, but a dummy variable for weekends and bank holidays for a robustness check. For the appropriate care scores which were proportions, I ran analysis using Ordinary Least squares regression. All analysis was conducted using Stata 13.

5.3.2. Data

The data included hospital records at spell level from the Secondary Uses Service (SUS) (NCHOD 2015) containing patient and Trust-level characteristics and data from the programme's Quality Measures Reporter (QMR) (AQ 2015), which records delivery of the process measures of care for each patient. The linked data contained 39,917 spells for patients admitted with AMI and 85,655 spells for patients admitted with pneumonia across 14 quarters, October 2009 until April 2013. This encompassed the entire population of AMI and pneumonia patients qualifying for inclusion in the Advancing Quality programme from the 24 Trusts in the North West of England.

The introduction of Advancing Quality meant that Trusts who all participated to be a part of the scheme in the North West of England 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 the quality of data collected was not as complete in the first year as it was in subsequent years. In the first year of data, the link between the unique patient identifiers in SUS and Advancing Quality were not saved, and therefore, I was unable to match the Advancing Quality dataset with the rich SUS data directly. I decided to drop the first year of data for my analysis.

5.3.2.1. Process measures of care

For this study, I did not account for the exclusion of patients from process measures of care. I was interested in how the quality of care was distributed between the days of the

week. By comparing proportion of achievement against failing to achieve captures the quality of care on the weekends without including excluded patients. I assumed that the probability of patients being excluded was fixed throughout all days of the week.

I made a distinction between arrival process measures of care and discharge process measures of care as discharge process measures of care were likely to be given on weekdays rather than weekends as most patients admitted on weekends would be discharged during weekdays. By having admission and discharge process measures of care, I was able to test whether it was important to provide immediate quality of care for patient health outcomes, or where there is a more cumulative effect of total healthcare quality affecting health outcomes. For discharge measures, as majority of patients admitted on the weekend would be discharged during weekdays, the provision of discharge process measures should not be affected by admission during the weekend. If I believe that the weekend effect is explained by the quality of care at admission, I should find effects when examining admission process measures of care, but not for discharge measures.

The following two conditions were the focus of this study: AMI and pneumonia.

5.3.2.1.1. AMI

There were seven process measures of care for AMI under Advancing Quality. Three of these were process measures given when patients arrive at hospital: aspirin at arrival; percutaneous coronary intervention (PCI) within 90 minutes of hospital arrival; and fibrinolytic therapy within 30 minutes of hospital arrival. Four process measures were given on discharge or during the hospital stay: aspirin prescribed at discharge; beta blocker at discharge; angiotensin converting enzyme (ACE) inhibitors for left ventricle systolic dysfunction (LVSD); and smoking cessation advice.

For each process measure of care and each patient I knew whether the target was achieved. Each process measure of care was a dichotomous variable and each will be an outcome for each individual.

5.3.2.1.2. Pneumonia

Four of the five process measures of care for pneumonia under Advancing Quality were given to patients at arrival: oxygenation assessment; initial antibiotic selection for community acquired pneumonia in immunocompetent patients; blood cultures in A&E before initial antibiotics received in hospital; and initial antibiotic received within six hours of hospital arrival. Smoking cessation advice is a process measure given either at discharge or during the hospital stay. Each indicator was a dichotomous variable which takes the value of one if the patients have been given the process measure of care.

5.3.2.2. Appropriate care scores

I created two appropriate care scores which combined process measures of care to generate a single composite quality of care metric from the process measures of care. This captured the hospital's provision of the incentivised quality of care from the quality improvement programme.

1. I create an appropriate care score which was a proportion of all process measures, admission and discharge measures, which a patient had been given, removing any process measures of care which the patient had been excluded from.
2. A proportion for all admission process measures of care given to an individual with excluded process measures not included in the calculation. I removed discharge process measures of care as the quality of care for discharge should not be affected by the day of admission.

5.3.2.3. Mortality

The mortality measure I used was spell mortality which was generated using the discharge method from the SUS dataset. This dichotomous variable takes a value of one if a patient was discharged dead from a spell. I did not use the day of death of a patient, I linked whether a patient had died to the day of admission.

5.3.2.4. Day of the week

I created a categorical day of the week variable which was generated using the admission date of spells from the SUS dataset. Each day of the week was specified as a calendar day from 12 midnight to 12 midnight. This categorical variable is the main variable of interest. Analysis cannot extend to out of hours care as time of admission was not routinely collected in Secondary Uses Service, and therefore we were unable to attribute an admission for each patient.

5.3.2.5. Weekends

For my analysis I used a weekend variable where 'weekend' is defined as the full days, Saturday and Sunday. For sensitivity analysis I added bank holidays to weekends as hospital services on weekends were comparable with bank holidays (Alspach 2010; Smith et al. 2014). Also from a patient perspective, the bank holidays will also bring about reduced transport service. Over my time period from 1st October 2008 until 31st March 2013, I observed 31 bank holidays. For sensitivity analysis, I also included a separate dichotomous variable just for bank holidays.

5.3.2.6. Elixhauser comorbidities

Elixhauser comorbidities were a set of dummy variables which captured a patients comorbidities (Danielle A Southern 2004). Elixhauser comorbidities will capture unobserved variations in whether a patient would receive process measures of care or not in hospital as morbidities may alter the complexity of treatment for an individual. Using the primary and secondary diagnosis fields in the SUS dataset, which were given as four digit International Classification of Disease codes version 10 (ICD10), I followed the coding algorithms described by Quan et al. (2005) to create a list of 31 Elixhauser co-morbidities. I included the co-morbidities of patients to control for the complexity of patients. A study by Bottle et al. (2014) found that using Elixhauser co-morbidities can appropriately capture patients with multi morbidities.

For sensitivity analysis, I also used the Charlson Index instead of the Elixhauser comorbidities. I did this sensitivity analysis to see whether the method of accounting for comorbidities will affect results. This method of accounting for comorbidities used the same ICD10 codes as the Elixhauser method but, the Charlson index is a single score where the higher value represents more comorbidities (Charlson et al. 1987). A study by Gutacker, Bloor, and Cookson (2015) compared the predictability of co-morbidities using Elixhauser and Charlson index. The authors used data from five countries including England. The clinical conditions AMI, CABG and stroke were studied. The authors found that Elixhauser Comorbidities consistently outperformed the Charlson Index when accounting for co-morbidities. I decided to use Elixhauser Comorbidities over Charlson index as many previous studies comparing both methods on predicting health outcomes found that the Elixhauser Comorbidities were the best predictors (Sharabiani, Aylin, and Bottle 2012; Menendez et al. 2014; Southern 2004; Lieffers et al. 2011; Chu, Ng, and Wu 2010). The advantage of the Charlson index being a single score is that it is easier to interpret the effect of comorbidities on health outcomes.

5.3.2.6.1. ICD10 diagnostic codes

To adjust for the types of AMI or pneumonia, I created dummy variables for the main chapters of the ICD10 codes relating to the clinical conditions. These sets of variables captured differences within the clinical conditions and aimed to capture some severity of the condition. Further differences in severity were not captured.

The main ICD10 chapters for AMI were I210, I211, I212, I214, I219, I220, I221, I228, I229, and I251. The main ICD10 chapters for pneumonia were J13, J14, J15, and J18.

5.3.2.7. Patient distance measures

I captured patient distance to hospital as it was a proxy variable to capture patient self-selection into hospital care. Distances from the patient's area of residence to hospital sites were calculated using SUS data, and Lower Super Output Area (LSOA) centroid coordinates from the Ordinance Survey and a postcode to (x,y) coordinate look up. Each

coordinate was given as a northing (*y*) and easting (*x*). The northing coordinate (*y*) was the distance in metres above the most south western point on the OS grid for Great Britain. The easting coordinate (*x*) was measured by metres east of the most south western part of Britain using the OS grid. The LSOA centroid was the population weighted centroid of an LSOA and the data was given as northing and easting coordinates.

The distance between hospital site and patient LSOA centroid was calculated using the Pythagoras theorem. I took the difference between the northing coordinates and the difference between the easting coordinates. I then took the square root of the sum of the squared differenced coordinates. See equation 15:

$$\boldsymbol{distance\ to\ hospital} = \sqrt{((\boldsymbol{x_h} - \boldsymbol{x_p})^2 + (\boldsymbol{y_h} - \boldsymbol{y_p})^2)} \qquad \text{Equation 17}$$

Where x_h is hospital site easting is coordinate, x_p is population weighted LSOA easting coordinate, y_h is hospital site northing coordinate and y_p is the population weighted LSOA northing coordinate. To calculate this distance to hospital, I used locations of hospital sites within the Trust, as a Trust may have several sites spread over a large area.

5.3.2.8. Volume measure

I accounted for variations in the daily volume of patients admitted using the Advancing Quality dataset by counting the total number of patients admitted to all 24 Trusts on each day. This value was not Trust specific as small Trusts have a low number of patients attending hospital for each clinical condition. I performed a robustness check using a Trust specific emergency volume measure which was calculated using Hospital Episode Statistics (HES) (HSCIC 2012) data. I counted the number of emergency admitted patients for each day of the week by Trust. As the extract of available HES data spans the financial years of 2009-2012, the emergency volume variable was not available for the final four quarters of the Advancing Quality data.

5.3.2.9. Bed utilisation

Bed utilisation was calculated from HES data, from 1st October 2008 to 31st March 2013, using spell admission date and length of stay. I was able to calculate the dates which each patient were present in the hospital and subsequently calculate the number of patients present in each day in the hospital. The Trust utilisation variable was the ratio of each day's volume divided by the maximum utilisation I observed at that Trust.

This variables captures healthcare supply as Trust utilisation indicated a higher use of hospital services. This also captures workload of hospital staff, such as nurses and doctors who worked on Trust wards.

5.3.2.10. Trust characteristics

Trust characteristics were obtained from a variety of sources. Trust type was obtained from the National Centre for Health Outcomes Development (NCHOD). Foundation Trust status was obtained from Monitor website (Monitor 2014). Hospital quality was obtained from 2007/2008 annual health check report by the Healthcare Commission. I used the 2007/2008 annual health check as Advancing Quality may impact on the quality scores for the Advancing Quality hospitals.

5.4. Results

In this section I present the results from models for AMI and pneumonia separately. I begin by describing the descriptive statistics. I then described the results on the Kaplan-Meier survival curves. I began to describe regression results based on univariate regressions (equation 17) before describing the full models (equation 16). When describing the full models I initially described process measures of care first. I then described results for the appropriate care score, then those for weekend and bank holidays.

I then presented results from my robustness tests on the full models. The order of the robustness tests are: models including distance between patient residence and hospital

site; full model including bed utilisation; the full model using the Charlson index rather than the Elixhauser co-morbidities; and using various in-hospital mortality measures.

5.4.1. Descriptive statistics

5.4.1.1. AMI

Table 31 shows the descriptive statistics for all of the variables I used in the analysis for AMI. Mortality was one percentage point higher for patients admitted on the weekends. This weekend effect was not mirrored when I looked at specific process measures of care. I found little variation when examining days of the week for process measures of care. Proportion of process measures for AMI seem to exhibit ceiling effects fluctuating by between 0-4 percentage points across the days of the week. There were variations however, between process measures of care, with fibrinolytic therapy having the lowest achievement proportions.

There was little variation in patient ages by admission day with a higher average patient age on weekends. I found that patients on the weekend were older than patients admitted on the weekday by around half a year. This supported my assumption that case mix of patients admitted on weekends was different to the case mix of patients admitted on weekdays.

There was no variation across the days of the week when I looked at Elixhauser co-morbidities, main ICD 10 chapters, and distance travelled. The case mix of individuals may not be affected by the distance a patient has to travel to the hospital.

I did not find variations over the days of the week for Trust characteristics with the exception of volume for AMI and hospital utilisation. I found that a lower number of patients were admitted with AMI over the weekend. This was also reflected by the lower number of observations over the weekend. I also found that hospital utilisation on the weekend was lower than the weekdays. Hospital bed utilisation reduces over the days of the week. Friday had the lowest level of utilisation, before increasing on the weekends.

5.4.1.2. Pneumonia

Table 32 shows the descriptive statistics for all of the variables I used in the analysis for pneumonia.

<Insert Table 31>

I found that patient mortality does not fluctuate throughout the weekdays and increased by one percentage point on Saturday but returns to the levels of the weekdays on Sunday. This indicated a modest increase in mortality proportion on weekends.

Analysing the specific process measures of care, I found variations in proportions throughout all days of the week, with no indication that weekends yields a lower proportion of achievement. For one process measure, antibiotics received within six hours of arrival, I found that the achievement proportions were slightly higher over the weekend. Unlike achievement proportions for process measures of care under AMI I found that achievement proportions for pneumonia were lower. For pneumonia, the lower achievement proportions meant that there was more variation between patients who were given and not given process measures of care.

I found that patients attending hospital on the weekend were on average half a year older on the weekend when compared to the weekdays. I did not find that other patient characteristics exhibit much variation over the days of the week.

The descriptive statistics for Trust characteristics from pneumonia population were similar to the AMI population. The only variables which exhibit weekend effects were hospital utilisation and volume measures. I found that volumes of patients with pneumonia and all emergencies were lower on the weekends.

5.4.2. Survival analysis

5.4.2.1. AMI

In Figure 8, I found that the survival proportions on weekdays were higher than weekends. This was consistent across all stratifications of days of the week in figure 10. However, I also found that the higher probability of survival when admitted on weekdays

was persistent over long time periods. Analysing the probability of survival by admission days of the week, I found that there was not only a weekend effect for survival, but also that there were different survival proportions when considering each individual day of the week. Friday, Saturday and Sunday had consistently the lowest level of survival, and Tuesdays and Thursdays have consistently the highest. When I stratified by weekdays, weekends and bank holidays in figure 9, I found that survival proportion on bank holidays was consistently higher than on weekdays and weekends.

<Insert Figure 8>

<Insert Figure 9>

<Insert Figure 10>

5.4.2.2. Pneumonia

In Figure 11, I found that there was little difference in survival proportions for patients with pneumonia from the Kaplan-Meier curves. This included comparisons of weekdays with week end and comparisons of weekdays with weekends including bank holidays in figure 12. I found that the survival proportions were similar throughout out observed time period. Examining survival at individual days of the week in figure 13, I found slight variations over the days of the week. Saturday and Sunday exhibit lower proportions of survival. When I explored the survival proportions for bank holidays, I found that survival proportions for patients on the bank holidays were higher and consistently higher over the observed time period. This higher survival on bank holidays only emerged 30 days after admission.

<Insert Figure 11>

<Insert Figure 12>

<Insert Figure 13>

5.4.3. Univariate regression results

Table 33 and Table 34 show the univariate regression results for each process measure of care and in-hospital mortality for AMI and pneumonia respectively. I did not find any weekend effect in the process measures of care (Table 33). Patients have a lower probability of receiving primary PCI, when admitted on Sunday compared to Wednesday. However, this effect was not large enough to generate a statistically significant weekend effect. There was a positive effect on the death probability for patients admitted on the weekends, with and without bank holidays, by 0.6 and 0.8 percentage points respectively.

<Insert Table 33>

<Insert Table 34>

I found no statistically significant differences in the death probability for pneumonia patients by admission days of the week, weekends or bank holidays. For blood cultures, I found that patients have a lower probability of receiving the process measure of care if admitted on Monday (2.8 percentage points lower), Saturday (2.7 percentage points lower) and Sunday (3.7 lower) compared to Wednesday. This result is not repeated when I look at probability of receiving blood cultures process measure on the combination of weekends and bank holidays. However, I found a negative relationship between being admitted on weekends with no bank holidays on the receipt of blood cultures by 1.5 percentage points.

I found that patients admitted on the weekend were more likely to receive initial antibiotics by 2.4 percentage points on weekends with and without bank holiday admissions.

5.4.4. Process measures of care

5.4.4.1. Full models

Table 35 and Table 36 show the results from the full models for process measures of care for AMI and pneumonia respectively.

<Insert Table 35>

<Insert Table 36>

In Table 35, I did not find any statistically significant differences in receipt of all process measures of care when admitted throughout the days of the week with the exception of smoking cessation advice which is a discharge process measure. The results were largely consistent with the univariate regression results. For smoking cessation, I found that patients were statistically more likely to receive the advice if admitted on Tuesdays and Saturdays compared to Wednesday by 2.4 and 2.3 percentage points respectively

In Table 36 I found that there were differentials in the probability of receiving process measures of care for pneumonia patients if admitted on different days of the week. Consistent with the univariate regression results, the probability of receiving blood cultures if admitted on Monday, Friday, Saturday and Sunday were statistically different to Wednesdays by -2.3, -2.6, -3.3 and -2.7 percentage points respectively. Patients were 0.3 percentage points more likely to receive oxygenation assessment if admitted on Tuesday compared to Wednesday. Patients have different probabilities of receiving initial antibiotics if admitted on different days throughout the week. Sunday had a +1.9 percentage point difference when compared to Wednesday. Tuesday had a -1.7 percentage point difference when compared to Wednesday.

5.4.4.2. Appropriate care score

Table 37 shows the results using different version of the appropriate care score as the measure of quality. I did not find any statistically significant differences in the quality of care between the days of the week for all composite scores for process measures for AMI. I found some statistical variation in appropriate care scores between days of the week. Patients were more likely to receive a lower appropriate care score if admitted on Friday when compared to Wednesday, by -0.9 and -0.7 percentage points for appropriate care score for all process measures and appropriate care score for admission process measures respectively for pneumonia patients.

<Insert Table 37>

5.4.4.3. Weekends and bank holidays

Table 38 and Table 39 show the probability of receiving process measures of care if admitted on weekend or bank holidays for AMI and pneumonia, respectively, compared to weekdays. In Table 38, I did not find a weekend effect. The sizes of the statistically insignificant coefficients for the weekend variable were almost identical to the results from the univariate regressions.

<Insert Table 38>

<Insert Table 39>

In Table 39 I found that the coefficients on the weekend variable were identical to the univariate regressions to three decimal places. This includes the statistically significant result for initial antibiotics, showing that patients admitted on weekends had an increased probability of receiving by 2.4 percentage points.

5.4.4.4. Only weekends and only bank holidays

Table 38 and Table 39 show the results for the probability of receiving process measures of care if admitted on weekends compared to weekdays only. The weekend variable does not capture bank holidays. Table 38 and Table 39 also show the probability of receiving process measures of care if admitted on bank holidays when compared to non-bank holidays including weekends.

In Table 39, I found no statistically significant results, suggesting that there was no difference in the likelihood of receiving process measures of care if patients have been admitted on weekdays or on the weekend. By controlling for patient and Trust characteristics, the provision of blood cultures and initial antibiotic selections was not statistically significantly different between admissions on weekdays and weekends.

From Table 38 and Table 39, I did not find statistically significant results on patients admitted on bank holiday. This indicated that compared to non-bank holidays, the probability of receiving process measures of care does not differ by day of admission for both AMI and pneumonia.

5.4.5. Robustness tests

I conducted robustness checks to see whether results change if I changed the method in how I create proxy variables in the full models (Equation 16).

5.4.5.1. Continuous patient to hospital distance measure

Table 40 and Table 41 show the results of the full models including the continuous patient distance variable rather than categorical bands.

<Insert Table 40>

<Insert Table 41>

Table 40 shows results consistent with Table 35, where the effect sizes and signs of the probability to receive care were similar, even when the sample size for the continuous measure had reduced. The reduced sample size was the result of missing northing and easting coordinates for patient area LSOA. For smoking cessation advice, patients admitted on Tuesdays were 2.4 percentage points more likely to receive care compared to Wednesday. I also found that patients admitted on Friday were 1.7 percentage points more likely to receive ACEI for LVSD when compared to Wednesday admissions. A one kilometre increase in patient distance was associated with 0.1 percentage point increase in receiving primary PCI. The result was statistically significant to the 5 percent level.

Results from Table 41 for oxygenation assessments, smoking cessation and initial antibiotics were similar to results from Table 36. Patients admitted on Saturdays had a 3.5 percentage point lower likelihood of receiving blood cultures when compared to Wednesday, and 1.4 percentage point lower likelihood of receiving antibiotic selection if admitted on Fridays compared to Wednesdays.

A one kilometre increase in patient distance was associated with 0.1 percentage point decrease in smoking cessation advice. The result was statistically significant to 0.1 percentage level for both AMI and PN clinical conditions.

5.4.5.2. Bed utilisation

Table 42 and Table 43 show the full model specification, but, with the inclusion of daily bed utilisation in terms of total number of patients staying overnight for AMI and pneumonia conditions respectively.

<Insert Table 42>

< Insert Table 43>

The results in Table 42 show that admission day of the week did not affect the probability of receiving process measures of care for AMI when compared to being admitted Wednesday. The only exception was for ACEI or ARB for LSDV where patients admitted on Friday have a 2.3 percentage point increase in probability of receiving the process measure of care when compared to Wednesday. Bed utilisation did not affect the probability of receiving process measures of care when statistical significance was measured at the 5 percent level.

Results from Table 43 show consistent results for oxygenation assessment, blood cultures, antibiotic selection and smoking cessation with Table 36. I found no differences in admission days of the week and receipt of initial antibiotics. A one percentage point increase in bed utilisation had a 0.014 and 0.184 percentage point increase in oxygenation assessment and smoking cessation respectively. The results were statistically significant to the 5 and 1 percent level respectively. A one percentage point increase in bed utilisation has a 0.081 percentage point decrease in the probability of receiving initial antibiotics in 6 hours of hospital arrival. The result was statistically significant at the 5 percent level.

5.4.5.3. Charlson Index

Table 44 and Table 45 show the full model specification with the Elixhauser comorbidities replaced by the weighted Charlson index for AMI and pneumonia respectively.

<Insert Table 44>

<Insert Table 45>

In Table 44 I found that the higher the Charlson index, the lower the probability of receiving: aspirin at arrival by 0.2 percentage points; smoking cessation by 0.9 percentage points and fibrinolytic therapy by 3 percentage points. I found that the results from using the Charlson index were consistent with the Elixhauser comorbidities.

In Table 45 I did not find a relationship between the process measures of care and Charlson index. The results were consistent with the base results in table 35.

5.4.5.4. In-hospital mortality

Table 46 and Table 47 show results of five regression models where each specification controls for more confounders for AMI and pneumonia respectively.

<Insert Table 46>

<Insert Table 47>

In Table 46, I found that there was an initial weekend effect for deaths in the univariate regression results. However, once I adjusted for the volume of patients admitted, I found no weekend effect on mortality. This build-up of the models suggested that it was patient and Trust characteristics that explained the differences in mortality by days of the week.

In Table 47, I did not find any difference in in-hospital mortality by admissions on weekends compared to the weekday. This was consistent throughout the build-up of the full model, KM curves and sensitivity when using an emergency volume measure.

5.5. Discussion

5.5.1 Findings

Survival analysis showed that there was a weekend effect among patients admitted for AMI. The probability of survival was lower when admitted on the weekends, and the

effect on the survival risk is persistent over a long time period. This indicated that hospitals may have different levels of quality between weekdays and the weekend. I also found for AMI that each day of the week has different survival probabilities, with Friday having survival probability which mirror the weekends. This may be due to out of hours care on Fridays, where patients may have to wait Friday evening and over the weekend before hospitals return to usual operating capacity. I found that Tuesdays and Thursdays have the highest survival probability when compared to other days of the week.

The results from the survival analysis for pneumonia patients showed that there was no weekend effect on mortality for pneumonia patients. This suggested that either the immediate level of care for pneumonia patients was not as important as immediate care was for AMI or that the quality of care received if patients were admitted on weekends was not different to care when compared to the weekdays. I found that the survival probability for each day of admission does differ slightly, but this was not as pronounced as the case for AMI patients.

With respect to bank holidays, survival probabilities were higher than for weekends and weekdays for both AMI and pneumonia. This is counter intuitive as I would expect bank holidays will be similar to weekends due to levels of staff in hospital and availability of public transport being similar. The consistent finding may be due to a systematic difference in the patients attending hospitals on the weekend. Another finding was that the survival probability for patients admitted on bank holidays for pneumonia was higher only after 30 days following admission.

I found that patient and Trust characteristics explained the difference between process measures of care by day of the week. As day of admission is not linked with the probability of receiving care, I did not find differences between admission measures and discharge process measures of care.

From the results for mortality, I found that there is no weekend effect on mortality once I accounted for volume of admissions. The only significant result I found is from the univariate regression for AMI patients when the mortality proportion is not adjusted for confounding factors. The Trust and patient characteristics explain the elevated mortality proportion of patients for AMI. My finding for pneumonia is even stronger, where I found that the univariate regression results of mortality and being admitted on weekend has no statistically significant relationship.

I found that the only process measure which consistently was statistically significant different across all the models is initial antibiotics, where patients admitted on weekends were more likely to receive this measure. Blood cultures also exhibited a negative weekend effect. Blood cultures and antibiotic selection process measures form a pathway of care for pneumonia patients. I would expect that blood cultures were conducted less on a weekend as laboratories conducting these tests would not be fully operational. As blood cultures were not taken, I found that the probability for a patient to have appropriate antibiotics selected was lower, although not statistically significant. Patients on the weekend were more likely to be given antibiotics within six hours of hospital arrival. This meant that patients who were admitted on weekends with pneumonia were more likely to receive antibiotics, although not the most appropriate antibiotics for the patient. Using Charlson index to account for comorbidities rather than Elixhauser comorbidities as the sizes of the marginal effects and statistical significance were largely unchanged.

5.5.2. Implications

It was not observable from my data what was driving the mortality proportions and weekend effects for both patient and Trust effects. I suspected that there may be two levels of selection. For Trusts, there may be higher admission thresholds on the weekends due to reduced staffing levels. Therefore, a patient needs to have a more severe condition to be admitted by a Trust. This will therefore increase the mortality proportion on the weekend, as a lower number of more severe patients were admitted. Controlling for co-morbidities and ICD 10 chapters was linked to patient complexity. There may be additional variation in complexity of the incident that was not captured.

For patients, I suspected that on weekends fewer patients will go to hospital with mild AMI and wait until the weekday to present themselves in accident and emergency (Phillips et al. 2004). Patients with mild AMI may mistake the condition, such as heart burn, and therefore inadvertently become more severe patients when they were admitted on Monday. Using admission proportions I partially control for aspects of severity through hospital patient selection. I tried to control for the patient self-selection by adding the distance between the patient's home and hospital site. However, this does not capture patient self-selection on the weekends as the distance measure does not change by the day

of the week. To control for patient self-selection to hospital and the potential difficulties arising from patients going to hospitals on weekends (Flynn 2013), ideally my distance measure will be travel time by day as travel arrangements/services may change over the weekend.

5.5.3. How this relates to the literature

Our results agree with the previous body of research which does not find a weekend effect on patient mortality for patients admitted with pneumonia (Aylin et al. 2010; Freemantle et al. 2012; Cram et al. 2004). My results for AMI indicated that there was a weekend effect on mortality for patients with AMI which agrees with previous studies (Kostis et al. 2007; Freemantle et al. 2012; Aylin et al. 2010; Becker 2007; Cram et al. 2004). However, after controlling for the volume of admissions to address patient condition severity, I found no weekend effect on mortality. I also add to this body of research by testing whether the quality of care on weekends is different.

Like all previous research I did not find what explains the mortality effect. However, I had ruled out one possibility in process measures of care. I had also added to the literature by analysing each day of the week to see the quality of care was different throughout the week. This will be beneficial as the differences by the days of the week may help identify the explanation for the weekend effect.

5.5.4. Strengths

The main strength of this study was the linkage of process measures of care to spells and outcomes at a patient level. I was therefore able to assess whether process measures also exhibited a weekend effect using a large administrative dataset. This tests whether the quality of incentivised care was different between weekdays and the weekend.

The second strength of this study was analysing bank holidays as well as weekends. I was able to test if the quality of care and mortality were different between weekdays, weekends and bank holidays. Hospital services on weekend and bank holidays were

similar. My findings suggested that quality of care and mortality differences were not attributable to hospital services or hospital staff.

The third strength of this study was capturing aspects of patient case mix which were not usually recorded. Using administrative data, I used volume of patients and hospital distance to control for unobserved patient case mix factors.

5.5.5. Limitations

One limitation of this study was my definition of the weekend being just calendar dates. Without time of admission my study does not differentiate between usual operating hours and out of hours care. Other studies such as: Al-Asadi and Kadhim (2014); Kostis et al. (2007) and Coiera et al. (2014) included the evening on Friday as the start of the weekend due to a reduction of staffing during the end of the working week. This would make my study more sensitive to variations in quality between weekdays and weekends, especially when my results have identified modest Friday effects, this means that my study under-estimates the weekend effect. However this issue cannot be addressed with my datasets as the time of admission was not captured. To minimise the effect on not having time of admission, I compared all days to Wednesday.

Another limitation of this study was the sample size of each indicator of care. There were two ways in which my sample suffers from a problem of sample size. The first was the low number of spells where certain indicators have been given, fibrinolytic therapy and primary PCI. These two indicators exclude a high proportion of patients from the measures. This drop in sample size limits the number of explanatory variables and does not allow us to use Trust fixed effects as logistic regressions did not converge. The second issue with sample was the high proportion of patients receiving process measures. As the proportion at which patients were given process measures of care was so high, there was little variation to exploit between giving a process measure or not. My more robust analysis came from patients with pneumonia, where achievement proportions were more modest and not as subject to the ceiling effects, which yielded more significant results.

A limitation in the survival analysis is that patients who died out of hospital were not identified. My analysis counts them as patients who survived to the end of the period. A limitation of this piece of work and all the reviewed literature is the lack of a patient severity measure. Having a true measure of patient severity would enable us to look more closely into patient selection into care over the weekends. However, I did control for variables which may proxy for severity. I considered distance between place of residence and hospital and patient day volume. I also controlled for comorbidities using the Elixhauser comorbidities and ICD 10 main chapters (Bottle et al. 2014).

I did not capture all differences in the quality of care between weekdays and weekends. The quality of care I observed was the quality of care incentivised through a quality incentive scheme.

I assumed that the exclusion proportion of patients was the same on all days of the week. I found that the achievement proportion of process measures of care did not vary by days of the week. This may be driven by higher proportions of exclusion on the weekend. Further research could examine the weekend effect on the proportion of patient exclusions.

5.5.6. Policy implications

Our results suggested that the weekend effect on mortality cannot be addressed by improving the quality of care through a quality incentive scheme. Instead, policy makers needed more research into case-mix over the weekend to try to address why possible case-mix differences were occurring.

Policy makers need to take more time in deciding whether extending hospital hours is needed. The lack of the weekend effect found in this paper is consistent with a large number of studies internationally and in England (Cram et al. 2004; Hoh et al. 2010; Laupland, Ball, and Kirkpatrick 2009; Needleman et al. 2002; Schmulewitz, Proudfoot, and Bell 2005). The push for NHS hospitals to conduct 7 day working and 24 hours care may be accredited by one negative piece of research. Studies indicate that the quality of the care on the weekend is lower than the weekday (NHS 2013). Greater pressure is put upon the NHS to address the quality issue. My findings were inconsistent between bank

holiday and weekends. I expected that hospitals operate the same during weekend and bank holidays, the supply of health care was expected to be the same. I found that admissions on bank holidays were associated with better health outcomes than admissions on the weekend. This implies that the difference in mortality was not a supply side factor; staffing levels may not be the cause of the weekend effect on mortality.

I found that higher patient volumes results in lower mortality which was consistent with literature on volume of patients and patient outcomes (Glance et al. 2006; Kontos et al. 2013; Tu, Austin, and Chan 2001). This effect may be due to a higher average severity of patients on the weekends. Policy makers could investigate the type of patients attending accident and emergency on weekend, and investigate why patient volume was lower.

5.6. Conclusion

In this chapter I found that the distribution of incentivised quality of care did not vary between days of the week. The weekend effect on mortality was not explained by differences in incentivised quality of care. I found that mortality was lower on Bank Holidays when compared to the weekend, indicating that mortality was not driven by reduced hospital services or inexperienced staff. I found that the volume of patients on weekends was lower, indicating a difference in patient case mix. The weekend effect on mortality may be driven by patient case mix. The next chapter I will conclude the findings from this thesis.

6. Discussion

The aim of this thesis was to examine whether incentivised improvements in quality of health care were associated with improved patient outcomes and how that quality of care was distributed. There was a rich literature focusing on hospital quality of care but the literature on quality of care from pay-for-performance schemes was more limited. There was a particular gap in the literature using data on quality of care on a patient level. I identified three areas of research which addressed the gaps in the literature and answered the following empirical questions:

1. *Are financially-incentivised improvements in quality of care associated with better patient outcomes?* This question addressed whether there was a direct effect between the process measures of care from the quality incentive scheme, and health outcomes such as mortality and readmissions. (See Chapter 3)
2. *"Rich or poor, who gets more"? The distribution care from a quality improvement program*. I examined whether the distribution of care under the quality incentive scheme was equitable. Equity was a core goal of the NHS and I examined whether quality incentive schemes impact on health care inequity through increased selection of patients as hospitals strive to meet targets. (See Chapter 4)
3. *Is the weekend effect on hospital mortality attributable to lower quality of care?* I aimed to test whether the quality of care was driving the weekend effect on mortality, and if not, what were the drivers behind the increased mortality at the weekend. This question assessed how the quality of care from a quality incentive scheme was distributed by through the days of the week. (See Chapter 5)

6.2. Data

The main source of the data I used for this thesis was obtained directly from the Advancing Quality initiative. Process measures of care were incentivised under the quality incentive scheme. Trusts within the scheme began to record the provision of process measures on a spell-process measure level. This data was linked to spells from the Secondary Uses Service dataset using a unique spell level identifier.

I obtained the dataset used to create the performance scores which Advancing Quality publishes online. The data encompasses the entire population of Advancing Quality patients in an un-aggregated format and contained 252,284 spells. The unit of observation in Chapter three was a spell. For Chapters four and five, the unit of observation was a care opportunity of which there were several in a single spell.

I supplemented this dataset with publically available data. For all studies, I obtained Trust characteristics and Foundation Trust status. For Chapters three and four, I obtained income deprivation at lower super output area level. For Chapters three and five, I also used Hospital Episode Statistics to test for robustness.

6.3. What I found

6.3.1. Chapter 3

I found that the quality of care incentivised by the quality incentive scheme was associated with lower mortality for AMI and pneumonia patients, and lower readmissions for hip and knee replacement patients, for the study period, 1st October 2008 until 31st March 2013. My calculation of the effect of Advancing Quality on health outcomes using Trust level analysis was 876 lives saved and 1122 readmissions avoided over four and a half year period. My estimates using patient level analysis calculated that Advancing Quality saved 129 lives and avoided 121 readmissions. The individual level results were my preferred results, as this was the direct effect of process measures in patient outcomes.

I found that not all process measures of care lead to improvements in health outcomes. The results were driven by certain process measures of care. One of these (blood cultures taken for pneumonia patients) was removed from the incentive programme in October 2012. Quality improvement teams should research the effect of process measures before removing them from incentive schemes.

6.3.2. Chapter 4

Overall, I found that the process measures of care had a pro poor distribution. This result is primarily driven by the giving of advice, smoking cessation and discharge instructions. The results for smoking cessation advice were driven by the exclusion criteria. The results for discharge instructions were driven by exclusion and achievement.

Process measures of care under emergencies were distributed pro rich. This was driven by patients from poorer areas have more severe conditions. The prescriptions of drugs and diagnostic tests for patients were distributed fairly. This did not change over time.

For process measures of care for elective surgeries, provision of process measures of care with direct impact on health outcomes and complications such as antibiotics had no evidence of unequitable distribution of care. However ordering and selecting antibiotics had a pro poor distribution in the first year but eventually become pro rich in the final year.

6.3.3. Chapter 5

I found that the mortality proportion was higher on the weekend for AMI patients when compared to the weekday. This relationship had been found by many studies including studies using patient level administrative data from the UK (Aylin & Freemantle). I found that the quality of care from the quality incentive scheme on the weekend was the same as the quality of care on the weekdays. Therefore the difference in mortality proportion on the weekend was not due to lower levels of these types of quality of care.

I found that patients admitted on bank holidays did not have a higher mortality proportion than patients admitted on weekdays. This suggests that the weekend effect on mortality was not due to reduced hospital services and staffing levels.

I found that the volume of patients explained the weekend mortality effect. Higher volume of patients was associated with lower patient mortality. I found that the volume of patients admitted on the weekend for AMI and pneumonia were lower than the volume of patients admitted on weekdays. This suggested that the case mix of patients attending Trusts on weekends was different to patients admitted on weekdays. This difference in patient case mix may be driving the weekend effect.

6.3.4. Overall findings

The aim of this thesis was to examine whether incentivised improvements in quality of health care were associated with improved patient outcomes and how quality of care is distributed. Findings from this thesis show that health care providers distribute process measures of care equitably and evenly across the days of the week and process measures of care incentivised through Advancing Quality were positively correlated with health outcomes. The selection of process measures of care from policy makers was indicative of improved health outcomes and, my findings suggest that Trusts responded by exerting more effort to provide incentivised process measures of care.

The aim of the Incentive scheme was to improve the quality of care provided by Trusts in the North West of England and also to reduce the quality of care differential between Trusts in the North West of England. My research suggested that the incentivised quality of care was linked to improved health outcomes which implied that the incentive scheme had positive effects on health outcomes, shown in Chapter 3. Findings from Chapter 3 provided evidence that the quality of care is correlated to patient mortality, but this did not mean that Advancing Quality led to an overall improvement in the quality of care as this aspect was not testable with our data. However, Processes of care which Trusts were incentivised to perform were associated with improved health outcomes.

Once I found in Chapter 3 that the provision of process measures of care was correlated to improved health outcomes, the distribution of the process measures of care became important empirical question which examines how the quality of care was distributed. Distribution of care in this thesis was examined through patient's area income level (Chapter 4) and between the days of the week (Chapter 5). If process measures of care were not correlated with health outcomes, the distribution of process measures under Advancing Quality becomes a trivial exercise as the distribution of process measures would be not be considered as quality of care (Kohn et al. 2001; Campbell, Roland, and Buetow 2000).

Providing quality of care equitably to patents from different income areas indicates that financial incentives such as Advancing Quality did not incentivise trusts to implement 'cream skimming' (Barros 2003). 'Cream skimming' is when Trusts provide care to

easier to treat patients before treating more difficult patients. As patients from more deprived areas are generally harder to treat, cream skimming may affect patients from poorer areas generating a pro rich bias in the quality of care. Process measures of care under Advancing Quality were provided equitably where the achievement proportion was generally high from all Trusts. However, the research question in Chapter 4 did not empirically test whether the quality of care differential between Trusts in the North West of England narrowed over time. This would test if Advancing Quality would have met another aim of reducing quality of care differential. Instead, I focussed on within trust variation of quality. With achievement proportions being high from Trusts, this is indicative that the quality of care from Advancing Quality is similar between Trusts.

Chapter 5 found that the quality of care was distributed evenly across the days of the week. This chapter examined whether weekend quality of care was lower than quality of care on week days. Using data from a financial incentive scheme may not be appropriate to test the differential in the quality of care between week days and weekends in general. Financial incentives should universally increase the provision of process measures of care from all Trusts. Extra effort exerted from trusts to provide process measures of care may have led to the findings in Chapter 5. Distribution of care could only be examined within the financial incentive scheme with limited generalisability outside of financial incentives.

6.4. Strengths

The main strength of this thesis was the dataset used for all three studies. I obtained a unique patient level dataset which was linked at a patient level to process measures of care and health outcomes directly from the quality incentive scheme. For the first time, access to incentivised spell process measures of care level dataset was available. Previously there was only access to aggregated process measures of care under a financial incentive scheme.

With access to quality of care received on a patient level, I was able to test whether the quality of care exhibits a weekend effect similar to that observed for mortality. The assumption of a reduction of quality of care on weekends driving increases in mortality

can therefore be tested. Ultimately, I examined how the increase in quality of care under the scheme was associated with health outcomes.

All Trusts in the North West of England participated in the quality incentive scheme. This strength of universal uptake meant that there was not self-selection bias in my sample (Scott et al. 2011). Hospitals which participate in Premier Hospital Quality Initiative Demonstration self-select into the scheme (5%). This may have lead to a selection bias with better performing providers self-selecting into the incentive scheme.

6.5. Limitations

There were several limitations which affected all research questions under study. First and foremost, it was not mandatory for all Trusts in England to collect data on process measures of care before or after the implementation of Advancing Quality. This limitation meant that I could only assess the quality of care within the quality incentive scheme.

If the collection of process measures of care was captured and integrated with a large administrative dataset such as Hospital Episode Statistics or Secondary Uses Service dataset, this would benefit Chapter 3 as I would be possible to link all spells with the provision of process measures of care. I then would be able to create a patient's 30 day mortality in-hospital mortality proportion for my entire sample period, which would further help remove the endogeneity between health outcomes and process measures of care. Furthermore, with a link to Hospital Episode Statistics, I would be able to link patients to Office of National Statistics death records. This will enable us to link process measures of care for each individual patient that died in and out of hospital.

Another limitation to this thesis was that it relied on data from an administrative dataset. As with all administrative datasets, there may be data quality issues, including improper coding of the data and measurement error (Groen 2012). These errors have the potential to bias results through the loss of observations which may not be random. This limitation, however, applies to all administrative datasets and does not mean that the dataset was not suitable for estimation. The advantages of administrative datasets such as Advancing Quality over survey data are: the large sample size which directly covers the Advancing Quality population; and low probability of sampling bias.

Administrative datasets such as Advancing Quality lack information on patients' individual or household-level incomes, which were captured in survey data such as Understanding Society. For the empirical study reported in Chapter 4, I used area level income of patients as a proxy for patients' income. I therefore assumed that a patient's area income level was the patient's income score area level. A study by McLean et al. (2008) found that aggregated data for deprivation, under-estimates the true relationship of patient level income on distribution of care. Therefore my findings were a modest estimate which sets a lower bound to the relationship of how care has been distributed between different income groups.

Another limitation I encountered with the Advancing Quality dataset was that there were no measures of patient condition severity. This variable was not captured in administrative datasets and was a limitation for all empirical Chapters. The severity of condition was a determining factor for clinicians to decide to provide care for patients or not, and were linked to health outcomes where a high condition severity will be linked to higher mortality and readmission risk (Bottle et al. 2014). As this variable was not captured in administrative data from the English NHS (Bottle et al. 2014), I could not account for this. However, it would be a valuable variable for researchers. One problem of capturing severity as a variable is that there is no universal method to account for severity across all clinical conditions.

Collecting time of admission of patients was widely collected in administrative Accident and Emergency datasets such as Hospital Episode Statistics and Secondary Users Service datasets. Time of admission can be used to calculate whether patients have met a 4 hour waiting time target in Accident and Emergency. Time of admission however was not available for all Trusts in administrative inpatient datasets such as Advancing Quality. Chapter 5 would benefit from having admission time as robustness checks could be performed to compare the distribution of care between out of hours care with weekend. Reduced services in hospitals during out of hours will mirror that of hospital, and therefore it would be valuable test.

The overall generalisability of my findings may be limited to Advancing Quality. The North West of England was chosen for the quality incentive scheme as the quality of care in the North West of England was believed to be lower than the rest of England (Ledward, Horne, and Butterworth 2008). As the region was selected based on poor quality of care, the distribution of quality from this scheme may be unique to the region as

the region may be affected by similar attributes which were addressed under the scheme. Furthermore, this scheme affected 24 Trusts in England. This small sample of Trusts in a deprived area of the country may not be representative to generalise to other Trusts in England.

6.6. Policy implications

The process measures of care incentivised under the scheme have been distributed broadly in an equitable way, which was in line with the policy of the NHS where people of an equal need should receive equal care. Furthermore, I found that the provision of these process measures of care were associated with improved health outcomes.

From the empirical study reported in Chapter three, I found that policy makers need to weight and wait. The trends for the provision of blood cultures were increasing over time and I found that the provision of blood cultures led to improved patient outcomes. This indicator was removed before research was conducted on the association between process measures of care and health outcomes. I found that not all process measures of care lead to better health outcomes. Basing the financial award on the effect on improving health may be a more powerful way to improve health outcomes.

From the empirical study reported in Chapter four, I found that provision of care was distributed pro poor and the distribution was based on patient severity or case mix. An implication is that process measures of care should be more tailored towards patient socioeconomic backgrounds.

From the empirical study reported in Chapter five, I found that the weekend effect on mortality was not attributable to reduced likelihood of receiving selected process measures at weekends. Instead, it may be explained by patient case mix, in particularly patient severity. The Department of Health in the UK aimed to improve quality of care by increasing hospital operating hours on weekdays and weekends (NHS 2013). The findings of this research implicate that there may be an accessibility of care issue. If patients were being selected based on higher severity criteria for hospital admission, then improving operating hours may have a positive effect on mortality. If accessibility to care was the issue, then schemes such as 24/7 may be limited.

6.7. Further research

From the empirical study reported in Chapter five I found that a lower number of patients are attending Accident and Emergency on the weekend, I exploited this as a potential sign of patient severity. An area of further research may be to find the reason behind the lower number of Accident and Emergency admissions for AMI and pneumonia on the weekend compared to weekdays. A potential starting point will be to use Accident and Emergency administrative dataset from Hospital Episodes Statistics and find whether the number of patients attending Accident and Emergency by ambulance is different on the weekend. Then I can identify whether there is a reduction in General Practitioner referrals to hospital with AMI or pneumonia conditions on the weekends when compared to the weekdays. The next step would be to use Office of National Statistics patient mortality records linked to Hospital Episode Statistics to find the effect of attendance by Ambulance and General Practitioner referrals on in and out of hospital patient mortality.

Another area of further research is to assess the impact of Advancing Quality on the provision of process measures of care. This will build on the studies by Kristensen et al. (2014), and Sutton et al. (2012) which used mortality as the outcome variable. This will also build on Chapters 4 and 5 of this thesis. Using data from Myocardial Ischemia National Audit Project, process measures of care are collected from all Trusts in England that provide healthcare to acute coronary syndrome patients, such as AMI (UCL 2013). Data is collected for each patient including process measures of care. Data collection for all Trusts spans from 2003 to 2013 (UCL 2013). By using a triple difference design, I could find the impact of Advancing Quality on distribution of care between patients of different socioeconomic backgrounds and by days of the week. I did not use audit data for my studies for two main reasons. Firstly the achievement proportions of process measures for AMI are high. This high achievement proportion means that little variation in the distribution of care between patients or by days of the week. This data may not help explain distributions of care. Secondly, the submission of data for audit is voluntary. The quality of information may not be consistent throughout the long sample period.

Trusts within the Advancing Quality scheme have potential to ‘game’ the incentive scheme in two ways. The first is that the incentive scheme is allowed to remove patients

from a patient list which the Advancing Quality team with Lesley Kitchen circulates. The circulated patient list identifies patients with AQ qualifying condition using both disease category codes, ICD 10, and procedure codes. The patients on which Trusts will be assessed for a financial incentive from Advancing Quality will be based on the patients on the circulated patient list. Managers of Trusts are allowed to add and remove patients from this circulated patient list as patients may have incorrect ICD 10 or procedure codes. As managers of Trusts have the autonomy to control which patients are included in Advancing Quality, the managers of Trusts under Advancing Quality have an incentive to remove patients which were not offered process measures of care. An area of further research is to assess the differences between the patients who were included in Advancing Quality with patients who were not included but have to relevant procedure and ICD 10 code. This will highlight whether Trusts systematically removed certain patient types from Advancing Quality.

The second way a Trust can 'game' the incentive scheme is through excluding patients from process measures of care. Trusts are allowed to exclude patients with patient safety in mind; excluding patients that failed to receive process measures will positively impact the proportion of eligible patients a Trust has provided process measures of care to. Further research can evaluate the differences in patients who receive and are excluded from process measures of care. In particularly, does the time of the year impact on the proportion of patients who are exception reported. As process measures of care are reported at the end of each financial year, would more exclusion happen closer to the end of the financial year, which would indicate gaming by Trusts?

In chapter 5 I assessed the quality of care on the weekend when compared to weekdays. I assumed that the proportion of patients being excluded on weekends is the same as the week day and therefore I do not include exclusions in my models. An area of further research is to explore exclusion to find if there is a difference between the weekend and weekday. With a lower number of staff on wards on weekends, more patients may be excluded on weekends when compared to the week day.

6.8. Conclusion

Process measures of health care which were financially incentivised through the Advancing Quality scheme were associated with lower mortality and readmissions.

Process measures of health care were also distributed equitably between patients and by days of the week. Policy makers aiming to introduce new process measures of care should select process measures which have a direct correlation with health improvement. Once process measures have been selected, using an Advancing Quality incentive scheme structure may ensure that the quality of care is provided equitably between patients and days of the week.

Achievement of process measures were high overall and began to plateau leaving little variation between days of the week or patients from different income backgrounds. The Advancing Quality team may need to introduce new process measures of care which are linked to improved patient outcomes to continue to incentivise quality of care improvement.

References

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

Al-Asadi, J. N., and F. N. Kadhim. 2014. "Day of Admission and Risk of Myocardial Infarction Mortality in a Cardiac Care Unit in Basrah, Iraq." *Nigerian Journal of Clinical Practice* 17 (5): 579–84. doi:10.4103/1119-3077.141422.

Alshamsan, Riyadh, Azeem Majeed, Mark Ashworth, Josip Car, and Christopher Millett. 2010. "Impact of Pay for Performance on Inequalities in Health Care: Systematic Review." *Journal of Health Services Research & Policy* 15 (3): 178–84. doi:10.1258/jhsrp.2010.009113.

Amemiya, Takeshi. 1975. "Qualitative Response Models." In *Annals of Economic and Social Measurement*, 3rd ed., 4:363–3 72.

AQ. 2015. "Advancing Quality." December 10. http://www.advancingqualitynw.nhs.uk/.

———. 2015a. "Conditions." Accessed August 15. http://www.advancingqualitynw.nhs.uk/measures/.

———. 2015b. "FAQs." Accessed February 19. http://www.advancingqualitynw.nhs.uk/about/faq.php.

Asha, Stephen Edward, Kathryn Titmuss, and Deborah Black. 2011. "No Effect of Time of Day at Presentation to the Emergency Department on the Outcome of Patients Who Are Admitted to the Intensive Care Unit." *Emergency Medicine Australasia* 23 (1): 33–38. doi:10.1111/j.1742-6723.2010.01371.x.

Ashworth, Mark, Jibby Medina, and Myfanwy Morgan. 2008. "Effect of Social Deprivation on Blood Pressure Monitoring and Control in England: A Survey of Data from the Quality and Outcomes Framework." *British Medical Journal* 337 (October): a2030. doi:10.1136/bmj.a2030.

Ashworth, Mark, Peter Schofield, Paul Seed, Stevo Durbaba, Maria Kordowicz, and Roger Jones. 2011. "Identifying Poorly Performing General Practices in England: A Longitudinal Study Using Data from the Quality and Outcomes Framework." *Journal of Health Services Research & Policy* 16 (1): 21–27. doi:10.1258/jhsrp.2010.010006.

Ashworth, Mark, Paul Seed, David Armstrong, Stevo Durbaba, and Roger Jones. 2007. "The Relationship between Social Deprivation and the Quality of Primary Care: A National Survey Using Indicators from the UK Quality and Outcomes Framework." *British Journal of General Practice* 57 (539): 441–48.

Aylin, P., R. Alexandrescu, M. H. Jen, E. K. Mayer, and A. Bottle. 2013. "Day of Week of Procedure and 30 Day Mortality for Elective Surgery: Retrospective Analysis of Hospital Episode Statistics." *BMJ* 346 (may28 1): f2424–f2424. doi:10.1136/bmj.f2424.

Aylin, P., A. Yunus, A. Bottle, A. Majeed, and D. Bell. 2010. "Weekend Mortality for Emergency Admissions. A Large, Multicentre Study." *Quality and Safety in Health Care*, January, qshc.2008.028639. doi:10.1136/qshc.2008.028639.

Barba, R., J. E. Losa, M. Velasco, C. Guijarro, G. García de Casasola, and A. Zapatero. 2006. "Mortality among Adult Patients Admitted to the Hospital on Weekends." *European Journal of Internal Medicine* 17 (5): 322–24. doi:10.1016/j.ejim.2006.01.003.

Barros, Pedro Pita. 2003. "Cream-Skimming, Incentives for Efficiency and Payment System." *Journal of Health Economics* 22 (3): 419–43. doi:10.1016/S0167-6296(02)00119-4.

Bartus, Tamás. 2005. “Estimation of Marginal Effects Using Margeff.” *The Stata Journal* 5 (3): 309–29.

Baum, Christopher F. 2006. *An Introduction to Modern Econometrics Using Stata*. Stata Press.

Beattie, Michelle, William Lauder, Iain Atherton, and Douglas J. Murphy. 2014. “Instruments to Measure Patient Experience of Health Care Quality in Hospitals: A Systematic Review Protocol.” *Systematic Reviews* 3 (1): 4. doi:10.1186/2046-4053-3-4.

Becker, David J. 2007. “Do Hospitals Provide Lower Quality Care on Weekends?” *Health Services Research* 42 (4): 1589–1612. doi:10.1111/j.1475-6773.2006.00663.x.

Bell, Chaim M., and Donald A. Redelmeier. 2001. “Mortality among Patients Admitted to Hospitals on Weekends as Compared with Weekdays.” *New England Journal of Medicine* 345 (9): 663–68. doi:10.1056/NEJMsa003376.

Bennett, Paula. 2012. “Nursing Quality Indicators: The next Step in Enhancing Quality in Emergency Care.” *International Emergency Nursing* 20 (3): 179–86. doi:10.1016/j.ienj.2012.04.001.

Blumenthal, David, and Jennifer Dixon. 2012. “Health-Care Reforms in the USA and England: Areas for Useful Learning.” *Lancet* 380 (9850): 1352–57.

Borland, Barry L., and Joseph P. Rudolph. 1975. “Relative Effects of Low Socio-Economic Status, Parental Smoking and Poor Scholastic Performance on Smoking among High School Students.” *Social Science & Medicine (1967)* 9 (1): 27–30. doi:10.1016/0037-7856(75)90155-9.

Bottle, A, R Gaudoin, R Goudie, S Jones, and P. Aylin. 2014. “Can Valid and Practical Risk-Prediction or Casemix Adjustment Models, Including Adjustment for Comorbidity, Be Generated from English Hospital Administrative Data (Hospital Episode Statistics)? A National Observational Study.” Health Services Delivery and Research.

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): 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.

Buis, Maarten L. 2011. “The Consequences of Unobserved Heterogeneity in a Sequential Logit Model.” *Research in Social Stratification and Mobility* 29 (3): 247–62. doi:10.1016/j.rssm.2010.12.006.

Campbell, S. M, M. O Roland, and S. A Buetow. 2000. “Defining Quality of Care.” *Social Science & Medicine* 51 (11): 1611–25. doi:10.1016/S0277-9536(00)00057-5.

Campbell, Stephen, Kerin Hannon, and Helen Lester. 2011. “Exception Reporting in the Quality and Outcomes Framework: Views of Practice Staff – a Qualitative Study.” *British Journal of General Practice* 61 (585): e183–89. doi:10.3399/bjgp11X567117.

Cassel, Christine K., Patrick H. Conway, Suzanne F. Delbanco, Ashish K. Jha, Robert S. Saunders, and Thomas H. Lee. 2014. “Getting More Performance from

Performance Measurement." *New England Journal of Medicine* 371 (23): 2145–47. doi:10.1056/NEJMp1408345.

Charlson, Mary E., Peter Pompei, Kathy L. Ales, and C.Ronald MacKenzie. 1987. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Diseases* 40 (5): 373–83. doi:10.1016/0021-9681(87)90171-8.

Cheng, Simon, and J. Scott Long. 2007. "Testing for IIA in the Multinomial Logit Model." *Sociological Methods & Research* 35 (4): 583–600. doi:10.1177/0049124106292619.

Christianson, Jon B, David J Knutson, and Roger S Mazze. 2006. "Physician Pay-For-Performance: Implementation and Research Issues." *Journal of General Internal Medicine* 21 (Suppl 2): S9–13. doi:10.1111/j.1525-1497.2006.00356.x.

Chung, Eugene S., Lin Guo, Donald E. Casey, Cheryl Bartone, Santosh Menon, Syed Saghir, Anubhav Mital, and Dean J. Kereiakes. 2008. "Relationship of a Quality Measure Composite to Clinical Outcomes for Patients With Heart Failure." *American Journal of Medical Quality* 23 (3): 168–75. doi:10.1177/1062860608315337.

Chu, Yu-Tseng, Yee-Yung Ng, and Shiao-Chi Wu. 2010. "Comparison of Different Comorbidity Measures for Use with Administrative Data in Predicting Short- and Long-Term Mortality." *BMC Health Services Research* 10 (1): 140. doi:10.1186/1472-6963-10-140.

CMS. 2013. "Centers for Medicare and Medicaid Services (CMS) Premier Hospital Quality Incentive Demonstration." April 10. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/index.html?redirect=/HospitalQualityInits/35_HospitalPremier.asp.

Coiera, Enrico, Ying Wang, Farah Magrabi, Oscar Perez Concha, Blanca Gallego, and William Runciman. 2014. "Predicting the Cumulative Risk of Death during Hospitalization by Modeling Weekend, Weekday and Diurnal Mortality Risks." *Bmc Health Services Research* 14 (May): 226. doi:10.1186/1472-6963-14-226.

Concha, Oscar Perez, Blanca Gallego, Ken Hillman, Geoff P Delaney, and Enrico Coiera. 2014. "Do Variations in Hospital Mortality Patterns after Weekend Admission Reflect Reduced Quality of Care or Different Patient Cohorts? A Population-Based Study." *BMJ Quality & Safety* 23 (3): 215–22. doi:10.1136/bmjqs-2013-002218.

CQC. 2009. "2008/09 Scoring Methodology for Quality of Services and Quality of Financial Management." London: Care quality commission.

Cram, Peter, Stephen L Hillis, Mitchell Barnett, and Gary E Rosenthal. 2004. "Effects of Weekend Admission and Hospital Teaching Status on in-Hospital Mortality." *The American Journal of Medicine* 117 (3): 151–57. doi:10.1016/j.amjmed.2004.02.035.

Crawley, Danielle, Anthea Ng, Arch G. Mainous, Azeem Majeed, and Christopher Millett. 2009. "Impact of Pay for Performance on Quality of Chronic Disease Management by Social Class Group in England." *Journal of the Royal Society of Medicine* 102 (3): 103–7. doi:10.1258/jrsm.2009.080389.

Danielle A Southern, Hude Quan. 2004. "Comparison of the Elixhauser and Charlson/Deyo Methods of Comorbidity Measurement in Administrative Data." *Medical Care* 42 (4): 355–60. doi:10.1097/01.mlr.0000118861.56848.ee.

Darzi, Ara. 2008. *High Quality Care for All: NHS next Stage Review Final Report*. Vol. 7432. The Stationery Office.

Dawson, D, H Gravelle, M O'Mahony, A Street, M Weale, A Castelli, R Jacobs, et al. 2005. "Developing New Approaches to Measuring NHS Outputs and Productivity." CHE Research Paper. York, UK: Centre for Health Economics.

Dixon, Anna, Artak Khachatryan, and Yang Tian. 2012. "Socioeconomic Differences in Case Finding among General Practices in England: Analysis of Secondary Data." *Journal of Health Services Research & Policy* 17 (April): 18–22. doi:10.1258/jhsrp.2012.011112.

DOH. 2009. "High Quality Care for All : Department of Health - Health Care." http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/healthcare/highqualitycareforall/index.htm.

———. 2010. "Equity and Excellence: Liberating the NHS." London: Department of Health.

———. 2015. "Best Practice Tariffs." Article. Accessed August 15. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/NHSFinancialReforms/DH_105080.

Doran, Tim, Catherine Fullwood, Evangelos Kontopantelis, and David Reeves. 2008. "Effect of Financial Incentives on Inequalities in the Delivery of Primary Clinical Care in England: Analysis of Clinical Activity Indicators for the Quality and Outcomes Framework." *Lancet* 372 (9640): 728–36. doi:10.1016/S0140-6736(08)61123-X.

Doran, Tim, Catherine Fullwood, David Reeves, Hugh Gravelle, and Martin Roland. 2008. "Exclusion of Patients from Pay-for-Performance Targets by English Physicians." *New England Journal of Medicine* 359 (3): 274–84. doi:10.1056/NEJMsa0800310.

Doran, Tim, Evangelos Kontopantelis, Catherine Fullwood, Helen Lester, Jose M. Valderas, and Stephen Campbell. 2012. "Exempting Dissenting Patients from Pay for Performance Schemes: Retrospective Analysis of Exception Reporting in the UK Quality and Outcomes Framework." *British Medical Journal* 344 (April): e2405. doi:10.1136/bmj.e2405.

Dorn, Spencer D., Nilay D. Shah, Bjorn P. Berg, and James M. Naessens. 2010. "Effect of Weekend Hospital Admission on Gastrointestinal Hemorrhage Outcomes." *Digestive Diseases and Sciences* 55 (6): 1658–66. doi:10.1007/s10620-009-0914-1.

Dow, J. K., and J. W. Endersby. 2004. "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research." *Electoral Studies* 23 (1): 107–22. doi:10.1016/S0261-3794(03)00040-4.

Doyle, Cathal, Laura Lennox, and Derek Bell. 2013. "A Systematic Review of Evidence on the Links between Patient Experience and Clinical Safety and Effectiveness." *BMJ Open* 3 (1). doi:10.1136/bmjopen-2012-001570.

Dr Foster Intelligence. 2011. "Inside Your Hospital. Dr Foster Hospital Guide 2001–2011." http://drfosterintelligence.co.uk/wp-content/uploads/2011/11/Hospital_Guide_2011.pdf.

Echaiz, J. F., J. P. Henderson, D. K. Warren, and J. Marschall. 2014. "Weekend Diagnosis of Escherichia Coli Urinary Tract Infection Does Not Predict Poor Outcome." *Epidemiology and Infection* 142 (7): 1422–24. doi:10.1017/S095026881300229X.

Elwyn, G., S. Buetow, J. Hibbard, and M. Wensing. 2007. "Respecting the Subjective: Quality Measurement from the Patient's Perspective." *BMJ* 335 (7628): 1021–22. doi:10.1136/bmj.39339.490301.AD.

Epstein, Arnold M., Karen E. Joynt, Ashish K. Jha, and E. John Orav. 2014. "Access to Coronary Artery Bypass Graft Surgery Under Pay for Performance Evidence

From the Premier Hospital Quality Incentive Demonstration." *Circulation: Cardiovascular Quality and Outcomes*, August, CIRCOUTCOMES.114.001024. doi:10.1161/CIRCOUTCOMES.114.001024.

Falk, Armin, Ernst Fehr, and David Huffman. 2008. "The Power and Limits of Tournament Incentives." *Work in Progress*. http://www.researchgate.net/profile/Armin_Falk/publication/228982197_The_power_and_limits_of_tournament_incentives/links/004635244d1ca277f8000000.pdf.

Ferguson, Henry JM, Nigel J. Hall, Aneel Bhangu, I. G. Panagiotopoulou, N. Chatzizacharias, M. Rana, K. Rollins, et al. 2014. "A Multicentre Cohort Study Assessing Day of Week Effect and Outcome from Emergency Appendicectomy." *BMJ Quality & Safety* 23 (9): 732–40. doi:10.1136/bmjqs-2013-002290.

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): 1635–42. doi:10.2105/AJPH.2011.300153.

Flynn, Paul. 2013. "Should the NHS Work at Weekends as It Does in the Week? No." *Bmj-British Medical Journal* 346 (February): f622. doi:10.1136/bmj.f622.

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): 61–70. doi:10.1001/jama.297.1.61.

Francis, R. 2013. "Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry." London: The Stationery Office.

Freedman, D. A. 2001. "Ecological Inference and the Ecological Fallacy." In *International Encyclopedia of the Social Behavioral Sciences*, 6:4027–30. New York: Elsevier.

Freemantle, N., M. Richardson, J. Wood, D. Ray, S. Khosla, D. Shahian, W. R. Roche, I. Stephens, B. Keogh, and D. Pagano. 2012. "Weekend Hospitalization and Additional Risk of Death: An Analysis of Inpatient Data." *Journal of the Royal Society of Medicine* 105 (2): 74–84. doi:10.1258/jrsm.2012.120009.

Gallerani, Massimo, Davide Imberti, Eduardo Bossone, Kim A. Eagle, and Roberto Manfredini. 2012. "Higher Mortality in Patients Hospitalized for Acute Aortic Rupture or Dissection during Weekends." *Journal of Vascular Surgery* 55 (5): 1247–54. doi:10.1016/j.jvs.2011.11.133.

Glance, Laurent G., Yue Li, Turner M. Osler, Andrew Dick, and Dana B. Mukamel. 2006. "Impact of Patient Volume on the Mortality Rate of Adult Intensive Care Unit Patients." *Critical Care Medicine* 34 (7): 1925–34. doi:10.1097/01.CCM.0000226415.93237.84.

Goddard, Maria, and Peter Smith. 2001. "Equity of Access to Health Care Services:: Theory and Evidence from the UK." *Social Science & Medicine* 53 (9): 1149–62. doi:10.1016/S0277-9536(00)00415-9.

Goldstein, Seth D., Dominic J. Papandria, Jonathan Aboagye, Jose H. Salazar, Kyle Van Arendonk, Khaled Al-Omar, Gezzer Ortega, Maria Grazia Sacco Casamassima, and Fizan Abdullah. 2014. "The 'Weekend Effect' in Pediatric Surgery - Increased Mortality for Children Undergoing Urgent Surgery during the Weekend." *Journal of Pediatric Surgery* 49 (7): 1087–91. doi:10.1016/j.jpedsurg.2014.01.001.

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): 858–65. doi:10.1016/j.amjmed.2005.01.070.

Gravelle, Hugh, Matt Sutton, and Ada Ma. 2010. "Doctor Behaviour under a Pay for Performance Contract: Treating, Cheating and Case Finding?*." *The Economic Journal* 120 (542): F129–56. doi:10.1111/j.1468-0297.2009.02340.x.

Groen, Jeffrey A. 2012. "Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures." *Journal of Official Statistics* 28 (2): 173.

Groves, Elliott M., Mahdi Khoshchehreh, Christine Le, and Shaista Malik. 2014. "Effects of Weekend Admission on the Outcomes and Management of Ruptured Aortic Aneurysms." *Journal of Vascular Surgery* 60 (2): 318–24. doi:10.1016/j.jvs.2014.02.052.

Gutacker, Nils, Karen Bloor, and Richard Cookson. 2015. "Comparing the Performance of the Charlson/Deyo and Elixhauser Comorbidity Measures across Five European Countries and Three Conditions." *The European Journal of Public Health* 25 (suppl 1): 15–20. doi:10.1093/eurpub/cku221.

Hall, J. A., D. L. Roter, and N. R. Katz. 1988. "Meta-Analysis of Correlates of Provider Behavior in Medical Encounters." *Medical Care* 26 (7): 657–75.

Handel, Adam E., Sunil V. Patel, Andrew Skingsley, Katrina Bramley, Roma Sobieski, and Sreeram V. Ramagopalan. 2012. "Weekend Admissions as an Independent Predictor of Mortality: An Analysis of Scottish Hospital Admissions." *Bmj Open* 2 (6): e001789. doi:10.1136/bmjopen-2012-001789.

Hart, Jt. 1971. "Inverse Care Law." *Lancet* 1 (7696): 405 – &.

Haustein, Knut-Olaf. 2006. "Smoking and Poverty." *European Journal of Cardiovascular Prevention and Rehabilitation: Official Journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology* 13 (3): 312–18.

Hawkes, Nigel. 2009. "Measuring up the NHS." *BMJ* 338 (February): b703. doi:10.1136/bmj.b703.

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.

Heckman, Jj. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46 (4): 931–59. doi:10.2307/1909757.

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): 247. doi:10.1186/1472-6963-10-247.

Hoh, Brian L., Yueh-Yun Chi, Michael F. Waters, J. Mocco, and Fred G. Barker. 2010. "Effect of Weekend Compared With Weekday Stroke Admission on Thrombolytic Use, In-Hospital Mortality, Discharge Disposition, Hospital Charges, and Length of Stay in the Nationwide Inpatient Sample Database, 2002 to 2007." *Stroke* 41 (10): 2323–28. doi:10.1161/STROKEAHA.110.591081.

Horwich, Tamara B., Adrian F. Hernandez, Li Liang, Nancy M. Albert, Kenneth A. LaBresh, Clyde W. Yancy, and Gregg C. Fonarow. 2009. "Weekend Hospital Admission and Discharge for Heart Failure: Association with Quality of Care and Clinical Outcomes." *American Heart Journal* 158 (3): 451–58. doi:10.1016/j.ahj.2009.06.025.

HSCIC. 2012. “Hospital Episode Statistics.” Standard. May 28. http://www.hscic.gov.uk/hes.

———. 2015. “Secondary Uses Service (SUS) - Health and Social Care Information Centre.” *Secondary Uses Service*. December 10.

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

Jairath, V., B. C. Kahan, R. F. A. Logan, S. A. Hearnshaw, S. P. L. Travis, M. F. Murphy, and K. R. Palmer. 2011. “Mortality From Acute Upper Gastrointestinal Bleeding in the United Kingdom: Does It Display a ‘Weekend Effect’?” *American Journal of Gastroenterology* 106 (9): 1621–28. doi:10.1038/ajg.2011.172.

James, Matthew T., Ron Wald, Chaim M. Bell, Marcello Tonelli, Brenda R. Hemmelgarn, Sushrut S. Waikar, and Glenn M. Chertow. 2010. “Weekend Hospital Admission, Acute Kidney Injury, and Mortality.” *Journal of the American Society of Nephrology: JASN* 21 (5): 845–51. doi:10.1681/ASN.2009070682.

Jha, Ashish K., E. John Orav, and Arnold M. Epstein. 2010. “The Effect of Financial Incentives on Hospitals That Serve Poor Patients.” *Annals of Internal Medicine* 153 (5): 299–U47. doi:10.7326/0003-4819-153-5-201009070-00004.

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): 1104–10. doi:10.1377/hlthaff.26.4.1104.

Jones, Andrew M. 2007. *Applied Econometrics for Health Economists: A Practical Guide*. Radcliffe Publishing.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47 (2): 263–91.

Kahn, Lawrence M., and Kimio Morimune. 1979. “Unions and Employment Stability: A Sequential Logit Approach.” *International Economic Review* 20 (1): 217–35. doi:10.2307/2526427.

Kaplan, E. L., and Paul Meier. 1958. “Nonparametric Estimation from Incomplete Observations.” *Journal of the American Statistical Association* 53 (282): 457–81. doi:10.2307/2281868.

Kohn, Linda T., Janet M. Corrigan, Molla S. Donaldson, and others. 2001. “Crossing the Quality Chasm: A New Health System for the 21st Century.” *Washington, DC: Committee on Quality of Health Care in America, Institute of Medicine*.

Kontopantelis, Evangelos, David A. Springate, Mark Ashworth, Roger T. Webb, Iain E. Buchan, and Tim Doran. 2015. “Investigating the Relationship between Quality of Primary Care and Premature Mortality in England: A Spatial Whole-Population Study.” *The BMJ* 350 (March): h904. doi:10.1136/bmj.h904.

Kontos, Michael C., Stephen L. Rennyson, Anita Y. Chen, Karen P. Alexander, Eric D. Peterson, and Matthew T. Roe. 2014. “The Association of Myocardial Infarction Process of Care Measures and in-Hospital Mortality: A Report from the NCDR®.” *American Heart Journal* 168 (5): 766–75. doi:10.1016/j.ahj.2014.07.005.

Kontos, Michael C., Yongfei Wang, Sarwat I. Chaudhry, George W. Vetrovec, Jeptha Curtis, and John Messenger. 2013. “Lower Hospital Volume Is Associated With Higher In-Hospital Mortality in Patients Undergoing Primary Percutaneous Coronary Intervention for ST-Segment–Elevation Myocardial Infarction A Report From the NCDR.” *Circulation: Cardiovascular Quality and Outcomes* 6 (6): 659–67. doi:10.1161/CIRCOUTCOMES.113.000233.

Kostis, William J., Kitaw Demissie, Stephen W. Marcella, Yu-Hsuan Shao, Alan C. Wilson, and Abel E. Moreyra. 2007. "Weekend versus Weekday Admission and Mortality from Myocardial Infarction." *New England Journal of Medicine* 356 (11): 1099–1109. doi:10.1056/NEJMoa063355.

Kristensen, Søren Rud, Rachel Meacock, Alex J. Turner, Ruth Boaden, Ruth McDonald, Martin Roland, and Matthew Sutton. 2014. "Long-Term Effect of Hospital Pay for Performance on Mortality in England." *New England Journal of Medicine* 371 (6): 540–48. doi:10.1056/NEJMoa1400962.

Laupland, Kevin B, Chad G Ball, and Andrew W Kirkpatrick. 2009. "Hospital Mortality among Major Trauma Victims Admitted on Weekends and Evenings: A Cohort Study." *Journal of Trauma Management & Outcomes* 3: 8. doi:10.1186/1752-2897-3-8.

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

Lee, Jonathan S., Wato Nsa, Leslie R. M. Hausmann, Amal N. Trivedi, Dale W. Bratzler, Dana Auden, Maria K. Mor, Kristie Baus, Fiona M. Larbi, and Michael J. Fine. 2014. "Quality of Care for Elderly Patients Hospitalized for Pneumonia in the United States, 2006 to 2010." *JAMA Internal Medicine* 174 (11): 1806–14. doi:10.1001/jamainternmed.2014.4501.

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): 1175.e9–17. doi:10.1016/j.amjmed.2011.05.029.

Lieffers, Jessica R., Vickie E. Baracos, Marcy Winget, and Konrad Fassbender. 2011. "A Comparison of Charlson and Elixhauser Comorbidity Measures to Predict Colorectal Cancer Survival Using Administrative Health Data." *Cancer* 117 (9): 1957–65. doi:10.1002/cncr.25653.

Lindenauer, Peter K., Denise Remus, Sheila Roman, Michael B. Rothberg, Evan M. Benjamin, Allen Ma, and Dale W. Bratzler. 2007. "Public Reporting and Pay for Performance in Hospital Quality Improvement." *New England Journal of Medicine* 356 (5): 486–96. doi:10.1056/NEJMsa064964.

Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata, Second Edition*. Stata Press.

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): 201–10. 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): 413–21. doi:10.1093/intqhc/mzg055.

Maddala, G. S., and Lung-Fei Lee. 1976. "Recursive Models with Qualitative Endogenous Variables." NBER Chapters. National Bureau of Economic Research, Inc. https://ideas.repec.org/h/nbr/nberch/10494.html.

Maly, Rose C., Yihang Liu, Elaine Kwong, Amardeep Thind, and Allison L. Diamant. 2009. "Breast Reconstructive Surgery in Medically Underserved Women with Breast Cancer: The Role of Patient-Physician Communication." *Cancer* 115 (20): 4819–27. doi:10.1002/cncr.24510.

Maxwell, R J. 1984. "Quality Assessment in Health." *British Medical Journal (Clinical Research Ed.)* 288 (6428): 1470–72.

McDonald, R, Ruth Boaden, Martin Roland, Søren Kristensen, R Meacock, Y Lau, T Mason, A Turner, and M Sutton. 2014. “Evaluation of the Advancing Quality Pay for Performance Programme in the NHS North West.” Health Services and Delivery Research.

McGovern, M. P., D. J. Williams, P. C. Hannaford, M. W. Taylor, K. E. Lefevre, M. A. Boroujerdi, and C. R. Simpson. 2008. “Introduction of a New Incentive and Target-Based Contract for Family Physicians in the UK: Good for Older Patients with Diabetes but Less Good for Women?” *Diabetic Medicine: A Journal of the British Diabetic Association* 25 (9): 1083–89. doi:10.1111/j.1464-5491.2008.02544.x.

McLean, Gary, Bruce Guthrie, Graham Watt, Mark Gabbay, and Catherine A. O’Donnell. 2008a. “Practice Postcode versus Patient Population: A Comparison of Data Sources in England and Scotland.” *International Journal of Health Geographics* 7 (July): 37. doi:10.1186/1476-072X-7-37.

———. 2008b. “Practice Postcode versus Patient Population: A Comparison of Data Sources in England and Scotland.” *International Journal of Health Geographics* 7 (July): 37. doi:10.1186/1476-072X-7-37.

Meacock, Rachel, Soren Rud Kristensen, and Matt Sutton. 2014a. “The Cost-Effectiveness of Using Financial Incentives to Improve Provider Quality: A Framework and Application.” *Health Economics* 23 (1): 1–13. doi:10.1002/hec.2978.

Meacock, Rachel, Søren Kristensen, and Matt Sutton. 2014b. “Paying for Improvements in Quality: Recent Experience in the NHS in England.” *Nordic Journal of Health Economics* 2 (1). https://www.journals.uio.no/index.php/NJHE/article/view/794.

Medicare, Centers for, Medicaid Services 7500 Security Boulevard Baltimore, and Md21244 Usa. 2013. “HospitalProcessOfCareMeasures.” January 9. http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalProcessOfCareMeasures.html.

Menendez, Mariano E., Valentin Neuhaus, C. Niek van Dijk, and David Ring. 2014. “The Elixhauser Comorbidity Method Outperforms the Charlson Index in Predicting Inpatient Death After Orthopaedic Surgery.” *Clinical Orthopaedics and Related Research®* 472 (9): 2878–86. doi:10.1007/s11999-014-3686-7.

Millett, Christopher, Jeremy Gray, Sonia Saxena, Gopalakrishnan Netuveli, Kamlesh Khunti, and Azeem Majeed. 2007. “Ethnic Disparities in Diabetes Management and Pay-for-Performance in the UK: The Wandsworth Prospective Diabetes Study.” *Plos Medicine* 4 (6): 1087–93. doi:10.1371/journal.pmed.0040191.

Millett, Christopher, Gopalakrishnan Netuveli, Sonia Saxena, and Azeem Majeed. 2009. “Impact of Pay for Performance on Ethnic Disparities in Intermediate Outcomes for Diabetes: A Longitudinal Study.” *Diabetes Care* 32 (3): 404–9. doi:10.2337/dc08-0912.

Mohammed, Mohammed A., Khesh S. Sidhu, Gavin Rudge, and Andrew J. Stevens. 2012. “Weekend Admission to Hospital Has a Higher Risk of Death in the Elective Setting than in the Emergency Setting: A Retrospective Database Study of National Health Service Hospitals in England.” *BMC Health Services Research* 12 (1): 87. doi:10.1186/1472-6963-12-87.

“Monitor | Sector Regulator for Health Services in England.” 2014. May 14. http://www.monitor-nhsft.gov.uk/.

———. 2014. Accessed May 14. http://www.monitor-nhsft.gov.uk/.

NCHOD. 2015. “HSCIC Indicator Portal.” https://indicators.ic.nhs.uk/webview/.

Needleman, Jack, Peter Buerhaus, Soeren Mattke, Maureen Stewart, and Katya Zelevinsky. 2002. “Nurse-Staffing Levels and the Quality of Care in Hospitals.”

New England Journal of Medicine 346 (22): 1715–22. doi:10.1056/NEJMsa012247.

NHS. 2013. "NHS Services Open Seven Days a Week: Every Day Counts." http://www.nhsiq.nhs.uk/resource-search/publications/every-day-counts-seven-day-services.aspx.

———. 2015. "Commissioning for Quality and Innovation (CQUIN) Payment Framework - NHS Institute for Innovation and Improvement." Accessed August 15. http://www.institute.nhs.uk/world_class_commissioning/pct_portal/cquin.html.

NHS Choices. 2013. "Authorities and Trusts - The NHS in England - NHS Choices." http://www.nhs.uk/NHSEngland/thenhs/about/Pages/authoritiesandtrusts.aspx.

NHS England, 2015. 2015. "NHS England » How Will We Measure Success?" Accessed March 16. http://www.england.nhs.uk/about/imp-our-mission/success/.

O'brien, Robert M. 2007. "A Caution Regarding Rules of Thumb for Variance Inflation Factors." *Quality & Quantity* 41 (5): 673–90. doi:10.1007/s11135-006-9018-6.

Palmer, R. H., and M. C. Reilly. 1979. "Individual and Institutional Variables Which May Serve as Indicators of Quality of Medical Care." *Medical Care* 17 (7): 693–717.

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): 1912–20. doi:10.1001/jama.295.16.1912.

Phillips, David P., Jason R. Jarvinen, Ian S. Abramson, and Rosalie R. Phillips. 2004. "Cardiac Mortality Is Higher Around Christmas and New Year's Than at Any Other Time The Holidays as a Risk Factor for Death." *Circulation* 110 (25): 3781–88. doi:10.1161/01.CIR.0000151424.02045.F7.

Quan, Hude, Vijaya Sundararajan, Patricia Halfon, Andrew Fong, Bernard Burnand, Jean-Christophe Luthi, L. Duncan Saunders, Cynthia A. Beck, Thomas E. Feasby, and William A. Ghali. 2005. "Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data." *Medical Care* 43 (11): 1130–39.

Reeves, David, Tim Doran, Jose M. Valderas, Evangelos Kontopantelis, Paul Trueman, Matt Sutton, Stephen Campbell, and Helen Lester. 2010. "How to Identify When a Performance Indicator Has Run Its Course." *BMJ* 340 (April): c1717. doi:10.1136/bmj.c1717.

Roberts, A, L Marshall, and A Charlesworth. 2012. "The Funding Pressures Facing the NHS from 2010/11 to 2021/22. A Decade of Austerity?" London: The Nuffield Trust.

Ross, Stephen A. 1973. "The Economic Theory of Agency: The Principal's Problem." *The American Economic Review* 63 (2): 134–39.

Ryan, Andrew M., and Jan Blustein. 2011. "The Effect of the MassHealth Hospital Pay-for-Performance Program on Quality." *Health Services Research* 46 (3): 712–28. doi:10.1111/j.1475-6773.2010.01224.x.

Ryan, Andrew M., Jan Blustein, Tim Doran, Marilyn D. Michelow, and Lawrence P. Casalino. 2012. "The Effect of Phase 2 of the Premier Hospital Quality Incentive Demonstration on Incentive Payments to Hospitals Caring for Disadvantaged Patients." *Health Services Research* 47 (4): 1418–36. doi:10.1111/j.1475-6773.2012.01393.x.

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): 274–90.

Ryan, Andrew M., and Tim Doran. 2012. “The Effect of Improving Processes of Care on Patient Outcomes: Evidence From the United Kingdom’s Quality and Outcomes Framework.” *Medical Care March 2012* 50 (3): 191–99. doi:10.1097/MLR.0b013e318244e6b5.

Saposnik, Gustavo, Akerke Baibergenova, Neville Bayer, and Vladimir Hachinski. 2007. “Weekends: A Dangerous Time for Having a Stroke?” *Stroke* 38 (4): 1211–15. doi:10.1161/01.STR.0000259622.78616.ea.

Saxena, Sonia, Josip Car, Darren Eldred, Michael Soljak, and Azeem Majeed. 2007. “Practice Size, Caseload, Deprivation and Quality of Care of Patients with Coronary Heart Disease, Hypertension and Stroke in Primary Care: National Cross-Sectional Study.” *Bmc Health Services Research* 7 (June): 96. doi:10.1186/1472-6963-7-96.

Schilling, Peter L., Darrell A. Jr Campbell, Michael J. Englesbe, and Matthew M. Davis. 2010. “A Comparison of In-Hospital Mortality Risk Conferred by High Hospital Occupancy, Differences in Nurse Staffing Levels, Weekend Admission, and Seasonal Influenza.” *Medical Care March 2010* 48 (3): 224–32. doi:10.1097/MLR.0b013e3181c162c0.

Schmulewitz, L., A. Proudfoot, and D. Bell. 2005. “The Impact of Weekends on Outcome for Emergency Patients.” *Clinical Medicine* 5 (6): 621–25. doi:10.7861/clinmedicine.5-6-621.

Schwartz, S. 1994. “The Fallacy of the Ecological Fallacy - the Potential Misuse of a Concept and the Consequences.” *American Journal of Public Health* 84 (5): 819–24. doi:10.2105/AJPH.84.5.819.

Scott, Anthony, Peter Sivey, Driss Ait Ouakrim, Lisa Willenberg, Lucio Naccarella, John Furler, and Doris Young. 2011. “The Effect of Financial Incentives on the Quality of Health Care Provided by Primary Care Physicians.” In *Cochrane Database of Systematic Reviews*, edited by The Cochrane Collaboration. Chichester, UK: John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/14651858.CD008451.pub2.

Sharabiani, Mansour T. A., Paul Aylin, and Alex Bottle. 2012. “Systematic Review of Comorbidity Indices for Administrative Data.” *Medical Care* 50 (12): 1109–18. doi:10.1097/MLR.0b013e31825f64d0.

Sharp, Adam L., HwaJung Choi, and Rod A. Hayward. 2013. “Don’t Get Sick on the Weekend: An Evaluation of the Weekend Effect on Mortality for Patients Visiting US EDs.” *American Journal of Emergency Medicine* 31 (5): 835–37. doi:10.1016/j.ajem.2013.01.006.

Simpson, Colin R., Philip C. Hannaford, Karen Lefevre, and David Williams. 2006. “Effect of the UK Incentive-Based Contract on the Management of Patients with Stroke in Primary Care.” *Stroke* 37 (9): 2354–60. doi:10.1161/01.STR.0000236067.37267.88.

Simpson, Colin R., Philip C. Hannaford, Matthew McGovern, Michael W. Taylor, Paul N. Green, Karen Lefevre, and David J. Williams. 2007. “Are Different Groups of Patients with Stroke More Likely to Be Excluded from the New UK General Medical Services Contract? A Cross-Sectional Retrospective Analysis of a Large Primary Care Population.” *BMC Family Practice* 8: 56. doi:10.1186/1471-2296-8-56.

Smith, Stacy, Ananda Allan, Nicola Greenlaw, Sian Finlay, and Chris Isles. 2014. “Emergency Medical Admissions, Deaths at Weekends and the Public Holiday Effect. Cohort Study.” *Emergency Medicine Journal* 31 (1): 30–34. doi:10.1136/emermed-2012-201881.

Square, T. 2015. “Patient Reported Outcome Measures.” Standard. March 12. http://www.hscic.gov.uk/proms.

Stewart, M. 1984. “Patient Characteristics Which Are Related to the Doctor-Patient Interaction.” *Family Practice* 1 (1): 30–36.

Street, R. L. 1991. “Information-Giving in Medical Consultations: The Influence of Patients’ Communicative Styles and Personal Characteristics.” *Social Science & Medicine (1982)* 32 (5): 541–48.

———. 1992. “Communicative Styles and Adaptations in Physician-Parent Consultations.” *Social Science & Medicine (1982)* 34 (10): 1155–63.

Sucov, Andrew, Jonathan Valente, and Steven E. Reinert. 2013. “Time to First Antibiotics for Pneumonia Is Not Associated with In-Hospital Mortality.” *The Journal of Emergency Medicine* 45 (1): 1–7. doi:10.1016/j.jemermed.2012.11.018.

Suissa, Samy, Sophie Dell’Aniello, Daniel Suissa, and Pierre Ernst. 2014. “Friday and Weekend Hospital Stays: Effects on Mortality.” *The European Respiratory Journal* 44 (3): 627–33. doi:10.1183/09031936.00007714.

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): 1821–28. doi:10.1056/NEJMsa1114951.

Sutton, M., and G. McLean. 2006. “Determinants of Primary Medical Care Quality Measured under the New UK Contract: Cross Sectional Study.” *British Medical Journal* 332 (7538): 389–90. doi:10.1136/bmj.38742.554468.55.

Taira, D. A., D. G. Safran, T. B. Seto, W. H. Rogers, and A. R. Tarlov. 1997. “The Relationship between Patient Income and Physician Discussion of Health Risk Behaviors.” *JAMA* 278 (17): 1412–17.

Travaglia, J, and D Debono. 2009. “Clinical Indicators: A Comprehensive Review of the Literature.” University of New South Wales, Sydney: Centre for Clinical Governance Research in Health.

Tu, J. V., P. C. Austin, and B. T. Chan. 2001. “Relationship between Annual Volume of Patients Treated by Admitting Physician and Mortality after Acute Myocardial Infarction.” *JAMA* 285 (24): 3116–22.

UCL. 2013. “NICOR: National Institute for Cardiovascular Outcomes Research: Transparency Agenda.” http://www.ucl.ac.uk/nicor/audits/minap/transparency.

Verlinde, Evelyn, Nele De Laender, Stéphanie De Maesschalck, Myriam Deveugele, and Sara Willems. 2012. “The Social Gradient in Doctor-Patient Communication.” *International Journal for Equity in Health* 11 (1): 12. doi:10.1186/1475-9276-11-12.

Werner, Rachel M. 2010. “Does Pay-for-Performance Steal From the Poor and Give to the Rich?” *Annals of Internal Medicine* 153 (5): 340–41. doi:10.7326/0003-4819-153-5-201009070-00010.

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): 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): 1464–84. doi:10.1111/j.1475-6773.2007.00817.x.

WHO. 2015. “WHO | International Classification of Diseases (ICD).” *WHO*. http://www.who.int/classifications/icd/en/.

Wooldridge, Jeffrey. 2013. *Introductory Econometrics*. International ed of 5th revised ed edition. Mason, OH: South-Western College Publishing.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wu, Wen-Chih, Lan Jiang, Peter D. Friedmann, and Amal Trivedi. 2014. "Association between Process Quality Measures for Heart Failure and Mortality among US Veterans." *American Heart Journal* 168 (5): 713–20.e3. doi:10.1016/j.ahj.2014.06.024.

Tables and Figures

Table 1: Process measures of care

Process measure	Target patient population	Description of process measure	Action required
Aspirin at arrival	AMI	Aspirin is an antiplatelet drug, used to thin the blood and prevent a blood clot from forming.	Patients must be prescribed aspirin or any other antiplatelet drug within 24 hours before or after hospital arrival, or when symptoms have started to surface
Aspirin at discharge	AMI or CABG	Aspirin is an antiplatelet drug, used to thin the blood and deter blood from forming a clot.	Patients must be prescribed aspirin or any other antiplatelet drug at discharge from hospital.
ACEI/ARB for LVSD	AMI or heart failure	Angiotensin Converting Enzyme Inhibitors (ACEIs) reduce blood pressure by reducing the production of an enzyme, angiotensin II, which causes blood vessels to contract. The contraction is natural in the human body as it aids blood circulation. Angiotensin Receptor Blockers (ARBs) also reduce blood pressure by reducing the effect of the enzyme angiotensin II. Left Ventricle Systolic Dysfunction (LVSD) occurs when this section of the heart weakens which reduces the heart's function to circulate blood around the body. ACEIs or ARBs therefore allow the heart to pump the blood around the body easier.	Either an ACEI or ARB must be prescribed to the target population who are not currently on either drug

Smoking Cessation Advice	AMI; heart failure; or pneumonia	Smoking cession advice or counselling involves informing patients about the harms of smoking and suggesting that patients should quit. A patient who smokes is defined by as having smoked a cigarette in the year prior to the admission date	Advice or counselling must be given to patients during the hospital stay
Beta blocker at discharge	AMI	Beta blockers are drugs used to manage high blood pressure and heart attacks by reducing the effects of stimulants to the heart such as adrenalin.	Beta blockers must be prescribed to patients at discharge from hospital.
Fibrinolytic therapy	AMI	Fibrinolytic therapy is treatment with drugs which help break down blood clots that have formed in the blood stream.	Patients must receive fibrinolytic therapy within 30 minutes of hospital arrival.
Primary PCI received	AMI	Primary Percutaneous Coronary Intervention (PCI) involves widening the coronary artery usually using a balloon, to allow blood flow. A stent is usually placed in the artery to keep the artery dilated.	Patients must receive a primary PCI within 90 minutes of hospital arrival. This process measure of care was introduced one year after the introduction of Advancing Quality scheme.
Prophylactic antibiotics received	CABG or hip and knee replacements.	Prophylactic antibiotics are drugs which reduce the risk of a patient acquiring an infection from surgery.	A patient must receive initial antibiotics up to one hour prior to surgical incision. However if vancomycin or a fluoroquinolone were administered, these must be administered two hours before surgical incision.
Prophylactic antibiotic selection	CABG or hip and knee replacements.	Prophylactic antibiotics are drugs which reduce the risk of a patient acquiring an infection from surgery.	Trusts must select prophylactic antibiotics following clinical guidelines for each condition and procedure before surgical incision.

Prophylactic antibiotics discontinued	CABG or hip and knee replacements.	Prophylactic antibiotics do not offer further benefit to a patient after the patient's wound has been closed for several hours. Prolonged doses of prophylactic antibiotics may increase the risk of infections of the digestive system	Prophylactic antibiotics must be discontinued within 24 hours after surgery
Evaluation of LVS	Heart failure	Evaluation of Left Ventricle Systolic (LVS) function is a diagnostic test that assesses the heart's ability to circulate blood around the body. The results are useful in deciding which medications will be suitable for the patient.	Evaluation of LVS must be conducted: before hospital arrival; during spell; or scheduled after spell.
Discharge instructions	Heart failure	Discharge instructions are information given to the patient, or a care giver, which educates the patient how to manage their health condition. The information must consist of: appropriate exercise; diet; medications on discharge; weight monitoring; future appointments; and information if the condition worsens.	Patients admitted with heart failure. To achieve this process measure, information must be given to the patient during hospital stay or at discharge.
VTE prophylactics ordered	Hip and knee replacements	Venous thromboembolism (VTE) is a term describing a blood clot in the vein which is potentially fatal, and is a complication resulting from operations.	Clinicians need to order drugs which prevent VTE.
VTE prophylactics received	Hip and knee replacements	Venous thromboembolism (VTE) is a term describing a blood clot in the vein which is potentially fatal, and is a complication resulting from operations.	VTE prophylactics must be administered with a period beginning 24 hours before surgery and ending 24 hours after surgery.

Oxygenation assessment	Pneumonia.	Oxygenation assessment tests for the level of oxygen in the blood stream from arteries for patients with pneumonia. A lack of oxygen in the blood is a mortality risk factor.	A patient must have the oxygenation levels in the blood tested by arterial blood gas or pulse oximetry within 24 hours of hospital arrival, or during hospital arrival.
Antibiotic selection	Pneumonia.	Antibiotics are drugs which reduce the risk of a patient acquiring an infection from surgery.	To achieve this measure, hospitals must select antibiotics following clinical guidelines.
Blood cultures	Pneumonia.	Blood samples are taken from patients to ascertain the type of infection which causes patients to have a certain condition.	Blood samples must be taken in the emergency ward before the first dose of initial antibiotics. This was removed in the 16^{th} quarter of the data sample period
Antibiotics received	Pneumonia.	Antibiotics are drugs which reduce the risk of a patient acquiring an infection from surgery.	A patient must receive the first dose of antibiotics within 6 hours of hospital arrival.

Table 2: Exclusion rules

	Process measures for each condition																								
	AMI							CABG				Heart failure				Hip and knee					Pneumonia				
List of exclusion rules	Aspirin Arrival	Aspirin Discharge	ACEI/ARB for LSVD	Smoking Advice	Beta Blocker Arrival	Fibrinolytic Therapy	Primary PCI	Aspirin Discharge	Antibiotic Received	Antibiotic Selection	Antibiotic Discontinued	Evaluation LVS	ACEI/ARB for LSVD	Discharge Instructions	Smoking Advice	Antibiotic Received	Antibiotic Selection	Antibiotic Discontinued	VTE Prophylaxis Ordered	Received VTE Prophylaxis	Oxygenation	Antibiotic Selection	Blood Cultures	Antibiotics Received	Smoking Advice
Patients who died or were a still-birth on the day of, or the day after, arrival	x	x	x	x	x			x				x	x	x	x						x	x	x	x	x
Patients involved in clinical trials	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Patients who were transferred between or within hospitals	x	x	x	x	x	x	x	x					x	x	x						x	x	x	x	
Patients receiving C*omfort Measures Only or palliative care*	x	x	x	x	x	x	x	x				x	x	x	x						x	x	x	x	x
Patients less than 18 years of age	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Allergy or other medical reasons	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			x				x		x	x						x	x	x	x	x
Other medical condition			x										x												
Severe Condition or reasons to delay		x				x	x						x												
Patients discharged on day of arrival	x																				x	x	x	x	
Patients who do not smoke				x											x										x
Patients who did not require treatment						x	x																		
Patients who were receiving antibiotics within 24 hours prior to arrival									x	x	x					x	x	x						x	
Patients who had a principal diagnosis suggestive of preoperative infectious diseases									x	x	x					x	x	x							
Patients whose ICD-9-CM (OPCS) principal procedure occurred prior to the Admission Date									x	x	x					x	x	x	x	x					

Patients with documented infection prior to surgical procedure of interest																					
Patients whose principal procedure was performed entirely by laparoscope																					
Patients who were receiving antibiotics more than 24 hours prior to surgery																					
Patients who had receipt procedures requiring general or spinal anaesthesia during hospital stay																					
Patients who did not receive any antibiotics during this hospitalization																					
Allergy or other medical reasons																					
Patients who were diagnosed with infections within three days after surgery end date																					
Patients who did not receive any antibiotics before or during surgery, or within 24 hours after surgery end time																					
Patients who had a hysterectomy and a caesarean section performed during this hospitalization																					
Patients who had a left ventricular assistive device or heart transplant procedure during hospital stay procedure code for LVAD and heart transplant																					
Burn patients																					
Patients who are on warfarin prior to admission																					
Patients with contraindications to both mechanical and pharmacological prophylaxis																					
Patients whose total surgery time is less than or equal to 30 minutes																					
Patients who stayed less than or equal to 24 hours postop																					
Patients with contraindications to both mechanical and pharmacological prophylaxis																					
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																					
Patients with Cystic Fibrosis																					
Patients who did not receive antibiotics or a blood culture																					

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

Figure 1: AMI achievement proportions over time

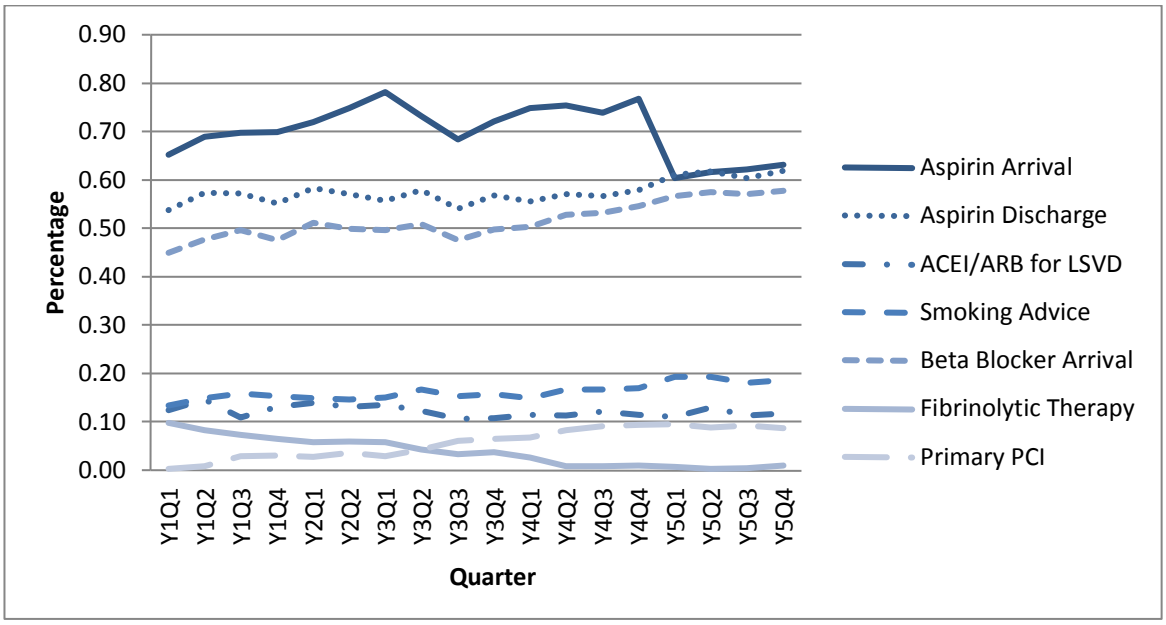


Figure 2: CABG achievement proportions over time

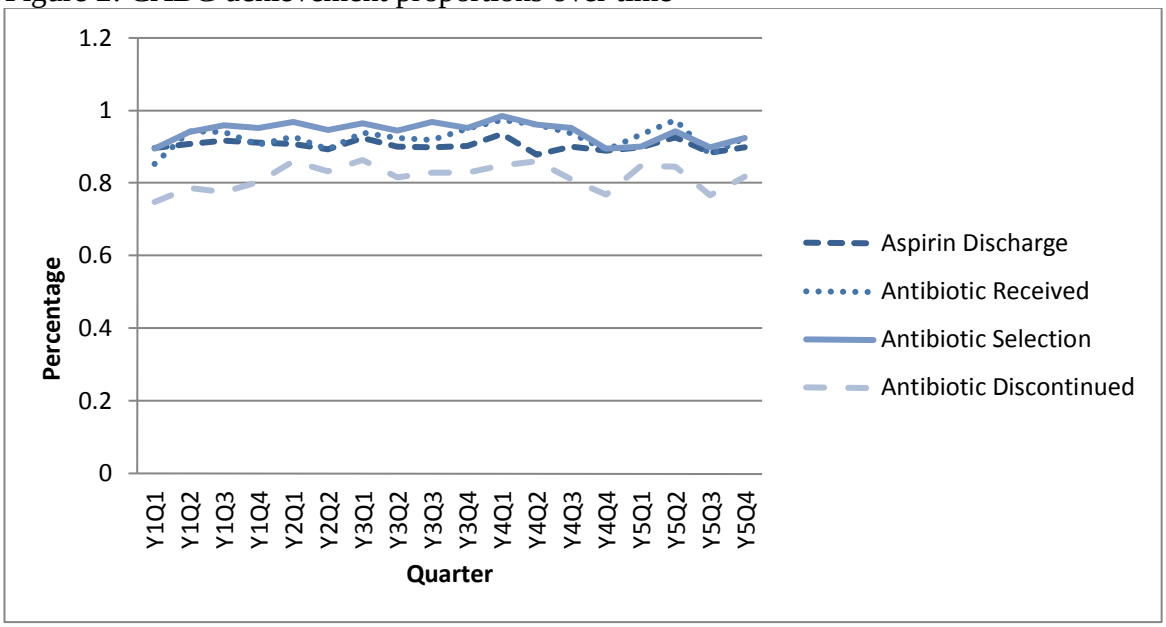


Figure 3: Hip and knee achievement proportions over time

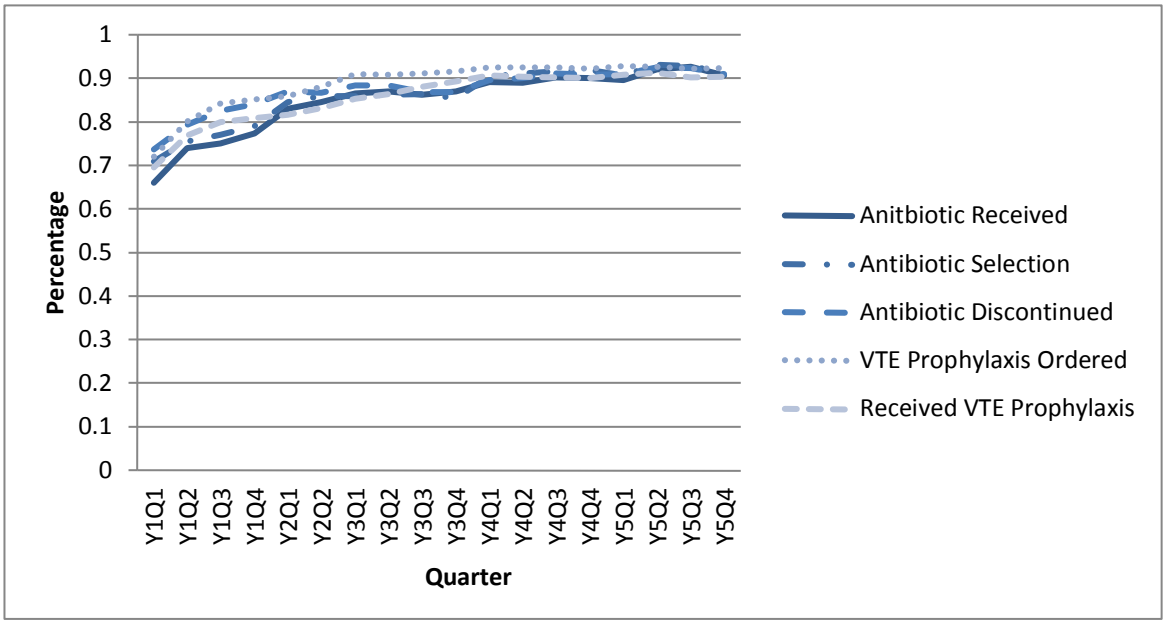


Figure 4: Pneumonia achievement proportions over time

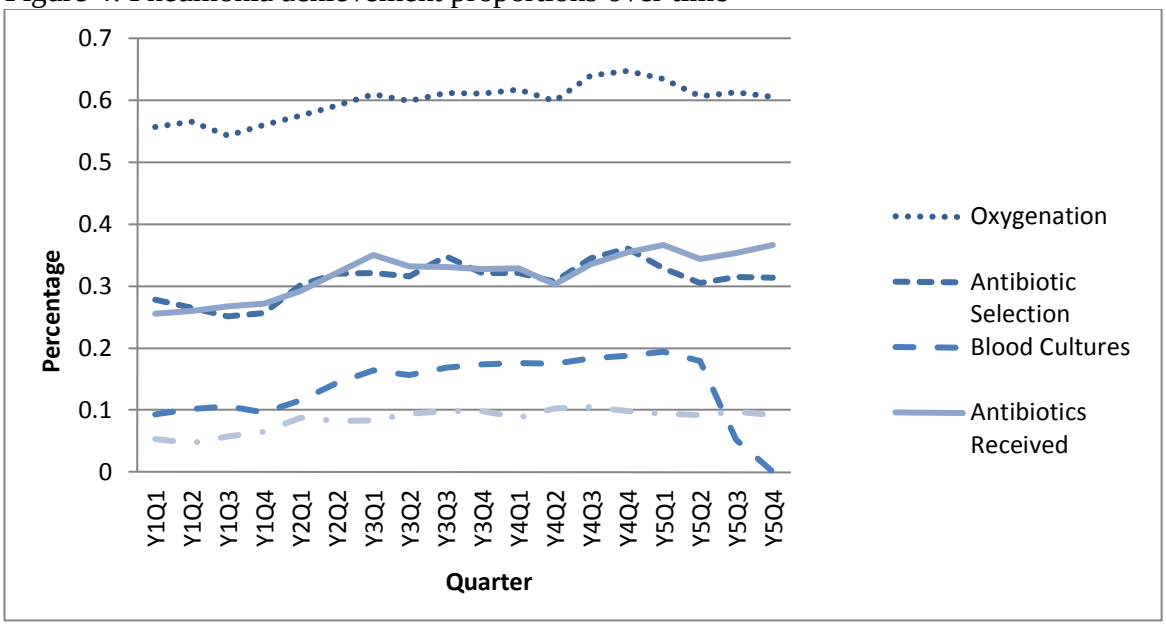


Figure 5: Mortality and readmission proportions over time

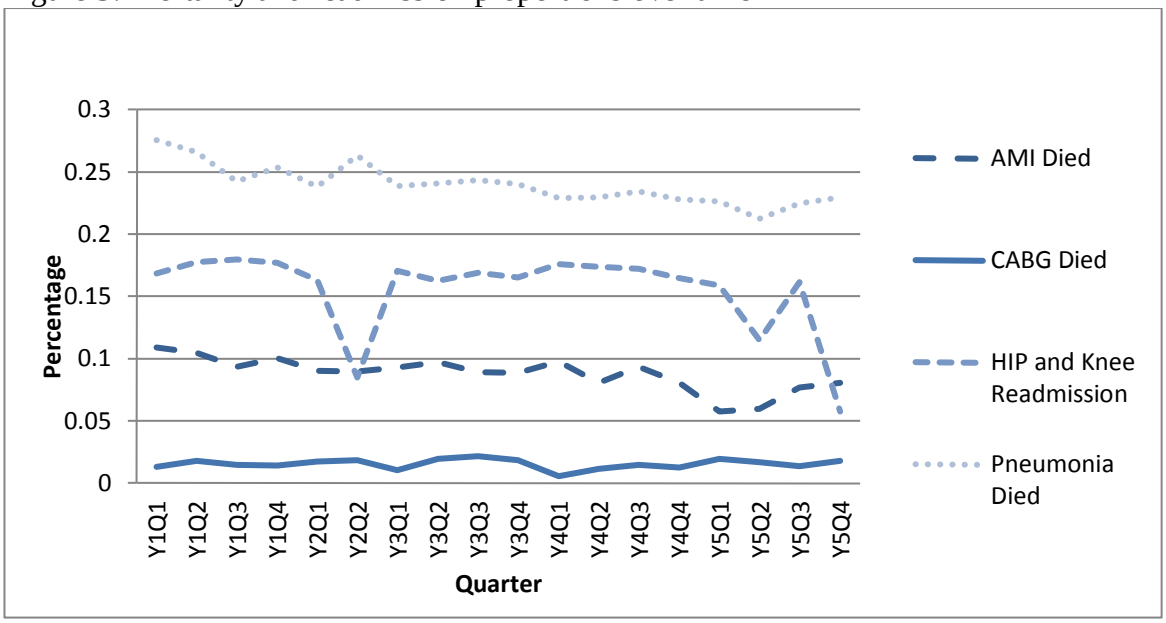


Table 3: Descriptive statistics

		Individual Level		Trust Level			
		Mean	SD	Mean	SD		
					Overall	Between	Within
Overall	**Died (AMI, CABG & PN)**	0.175	0.380	0.156	0.040	0.031	0.026
	Readmitted (HIPKNEE)	0.155	0.362	0.157	0.050	0.021	0.046
	Age (years)	72.12	14.255	72.642	2.381	2.218	0.951
	Male	0.503	0.500	0.520	0.057	0.048	0.032
	White	0.878	0.327	0.871	0.133	0.075	0.112
	Non-White	0.044	0.204	0.042	0.047	0.045	0.015
	Missing	0.078	0.269	0.087	0.130	0.058	0.118
	Income Deprivation	0.182	0.133	0.177	0.048	0.048	0.008
	Observations	226386		1335			
AMI	**Died**	0.086	0.281	0.092	0.043	0.021	0.038
	Age (years)	69.660	14.173	70.268	3.589	2.725	2.420
	Male	0.620	0.485	0.611	0.071	0.040	0.059
	White	0.859	0.349	0.846	0.143	0.093	0.112
	Non-White	0.057	0.232	0.051	0.060	0.053	0.029
	Missing	0.084	0.278	0.103	0.139	0.076	0.119
	Income Deprivation	0.180	0.131	0.179	0.051	0.049	0.016
	Appropriate Care Score	0.974	0.123	0.971	0.036	0.023	0.029
	Achievement of process measures						
	Aspirin at Arrival	0.693	0.461	0.728	0.141	0.110	0.091
	Aspirin at Discharge	0.580	0.494	0.554	0.156	0.123	0.099
	ACEI or ARB for LVSD	0.121	0.326	0.131	0.082	0.055	0.064
	Smoking Cessation Advice	0.165	0.371	0.147	0.090	0.077	0.051
	Beta Blocker Discharge	0.524	0.499	0.495	0.152	0.121	0.095
	Fibrinolytic Therapy	0.032	0.177	0.034	0.055	0.032	0.045
	Primary PCI	0.062	0.242	0.039	0.100	0.096	0.037
	Exclusions of process measures						
	Aspirin at Arrival	0.295	0.456	0.259	0.139	0.111	0.088
	Aspirin at Discharge	0.415	0.493	0.440	0.155	0.122	0.099
	ACEI or ARB for LVSD	0.876	0.329	0.866	0.084	0.057	0.065
	Smoking Cessation Advice	0.822	0.383	0.839	0.091	0.076	0.054
	Beta Blocker Discharge	0.460	0.498	0.486	0.153	0.123	0.095
	Fibrinolytic Therapy	0.962	0.192	0.959	0.062	0.037	0.050
	Primary PCI	0.935	0.247	0.960	0.104	0.099	0.041
	Observations for AMI	**45719**		**426**			
CABG	**Died**	0.015	0.122	0.015	0.011	0.008	0.010
	Age	65.691	9.806	70.268	3.589	2.725	2.420
	Male	0.809	0.393	0.611	0.071	0.040	0.059
	White	0.850	0.357	0.846	0.143	0.093	0.112
	Non-White	0.068	0.252	0.051	0.060	0.053	0.029
	Missing	0.081	0.273	0.103	0.139	0.076	0.119

	Income Deprivation	0.172	0.131	0.179	0.051	0.049	0.016
	Appropriate Care Score	0.972	0.096	0.972	0.033	0.025	0.025
	Achievement of process measures						
	Aspirin at Discharge	0.904	0.295	0.869	0.184	0.452	0.047
	Antibiotics Received <1 Surgery	0.928	0.259	0.887	0.194	0.459	0.064
	Antibiotic Selection	0.943	0.232	0.902	0.199	0.466	0.059
	Antibiotics Discontinued	0.817	0.387	0.781	0.230	0.437	0.076
	Exclusions of process measures						
	Aspirin at Discharge	0.094	0.291	0.128	0.185	0.453	0.047
	Antibiotics Received <1 Surgery	0.044	0.204	0.085	0.194	0.470	0.052
	Antibiotic Selection	0.029	0.168	0.070	0.196	0.476	0.047
	Antibiotics Discontinued	0.143	0.350	0.185	0.225	0.446	0.071
	Observations for CABG	**9683**			**76**		
HIP and Knee	**Readmitted**	0.155	0.362	0.157	0.050	0.021	0.046
	Age	71.131	11.401	71.461	1.905	1.703	0.903
	Male	0.387	0.487	0.383	0.045	0.018	0.041
	White	0.914	0.280	0.887	0.139	0.061	0.126
	Non-White	0.029	0.167	0.030	0.036	0.032	0.016
	Missing	0.057	0.233	0.083	0.137	0.051	0.129
	Income Deprivation	0.157	0.122	0.158	0.046	0.046	0.011
	Appropriate Care Score	0.944	0.143	0.940	0.078	0.037	0.069
	Achievement of process measures						
	Antibiotics Received <1 Surgery	0.857	0.350	0.851	0.131	0.058	0.118
	Antibiotic Selection	0.867	0.340	0.874	0.146	0.075	0.126
	Antibiotics Discontinued	0.878	0.328	0.876	0.096	0.036	0.089
	VTE Prophylactic Ordered	0.895	0.307	0.894	0.110	0.039	0.103
	VTE Prophylactic Received	0.864	0.342	0.864	0.128	0.057	0.115
	Exclusions of process measures						
	Antibiotics Received <1 Surgery	0.070	0.255	0.067	0.075	0.025	0.071
	Antibiotic Selection	0.068	0.252	0.066	0.078	0.027	0.073
	Antibiotics Discontinued	0.098	0.298	0.099	0.082	0.031	0.076
	VTE Prophylactic Ordered	0.079	0.269	0.077	0.069	0.024	0.065
	VTE Prophylactic Received	0.079	0.269	0.077	0.069	0.024	0.065
	Observations for Hip and knee	**72213**			**408**		
Pneumonia	**Died**	0.238	0.426	0.244	0.047	0.030	0.037
	Age	72.757	16.481	72.298	3.352	2.632	2.234
	Male	0.498	0.500	0.503	0.072	0.022	0.069
	White	0.898	0.302	0.879	0.128	0.072	0.108
	Non-White	0.043	0.202	0.041	0.054	0.048	0.026
	Missing	0.059	0.236	0.081	0.125	0.058	0.112
	Income Deprivation	0.201	0.138	0.190	0.054	0.053	0.015
	Appropriate Care Score	0.858	0.235	0.846	0.077	0.036	0.068
	Achievement of process measures						
	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 of process measures						
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 for pneumonia	**98771**			**425**		

Values in the table of 0.01 signify 1%. There are variations between trust and individual statistics as the trust means are not weighted means by individual, therefore the averaging of the trust data will differ from the individual means, this is due to the trust level analysis takes each trust as a unit of observation.

Table 4: Effect of process measures of care on AMI mortality

	Trust Level				Individual level (#)			
	Random Effects		Fixed Effects		Trust-quarter level process measures		Individual level process measures	
Age	0.008***	(0.001)	0.008***	(0.001)	0.005***	(0.000)	0.005***	(0.000)
Age Squared	-0.0002*	(0.000)	-0.0001	(0.000)	0.00002**	(0.000)	0.00002*	(0.000)
Male	-0.060	(0.032)	-0.046	(0.031)	0.001	(0.003)	0.004	(0.003)
Not White Ethnicity	0.104**	(0.037)	-0.013	(0.061)	0.001	(0.006)	0.005	(0.006)
Missing Ethnicity	-0.007	(0.015)	-0.024	(0.016)	0.008	(0.005)	0.001	(0.005)
Income Deprivation	0.151***	(0.046)	0.152	(0.111)	0.048***	(0.010)	0.046***	(0.010)
Achievement								
Aspirin at Arrival	-0.385***	(0.079)	-0.429***	(0.080)	-0.191*	(0.090)	-0.074***	(0.009)
Aspirin at Discharge					-0.107	(0.164)		
ACEI or ARB for LVSD					0.284	(0.259)		
Smoking Cessation Advice					0.182	(0.099)		
Beta Blocker Discharge					0.004	(0.080)		
Fibrinolytic Therapy	0.179	(0.166)	0.097	(0.171)	0.067	(0.162)	-0.031	(0.017)
Primary PCI					0.243	(0.288)	0.060	(0.042)
Exclusions								
Aspirin at Arrival	-0.311***	(0.083)	-0.381***	(0.084)	-0.190*	(0.092)	0.047***	(0.009)
Aspirin at Discharge	-0.013	(0.037)	0.103*	(0.040)	-0.024	(0.168)		
ACEI or ARB for LVSD	-0.023	(0.023)	-0.017	(0.028)	0.275	(0.260)		
Smoking Cessation Advice	-0.065	(0.043)	-0.082	(0.043)	0.168	(0.098)		
Beta Blocker Discharge	0.088*	(0.037)	0.038	(0.042)	0.010	(0.080)		
Fibrinolytic Therapy	0.061	(0.149)	-0.056	(0.157)	0.073	(0.152)	-0.011	(0.015)
Primary PCI	-0.058*	(0.029)	0.046	(0.042)	0.227	(0.267)	0.066	(0.041)
Constant	-0.096	(0.171)	-0.017	(0.187)				
Observations	**426**		**426**		**45719**		**45719**	
Mean Variance Inflation Factor	**8.09**							
R-squared (within)	**0.273**		**0.3276**					
R-squared (between)	**0.5571**		**0.2589**					
R-squared (overall)	**0.3347**		**0.2328**		0.107		0.189	
Hausman test, Chi-squared statistic	**91.73**							
	p <0.001							

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for individual level regressions. Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 5: Effect of Appropriate Care Score on AMI mortality

	Trust Level		Individual level (#)					
			Probit		Probit		Bivariate probit	
	Fixed Effects		Trust level ACS		Individual level ACS		Individual level ACS	
Age	0.005***	(0.001)	0.005***	(0.000)	0.004***	(0.000)	0.003***	(0.000)
Age Squared	-0.0001	(0.000)	0.00002**	(0.000)	0.00001	(0.000)	0.000005	(0.000)
Male	-0.060	(0.032)	0.001	(0.003)	0.005	(0.002)	0.003*	(0.002)
Non-White Ethnicity	0.021	(0.048)	0.001	(0.006)	0.006	(0.006)	0.004	(0.004)
Missing Ethnicity	-0.012	(0.016)	0.009	(0.005)	-0.0001	(0.005)	-0.001	(0.003)
Income Deprivation	0.205**	(0.064)	0.048***	(0.011)	0.037***	(0.010)	0.022***	(0.006)
ACS	-0.278***	(0.060)	-0.115*	(0.058)	-0.141***	(0.002)	-0.027**	(0.009)
Constant	-0.023	(0.089)						
Observations	**426**		**45719**		**45719**		**45719**	
R-squared (within)	**0.1863**							
R-squared (between)	**0.4754**							
R-squared (overall)	**0.2534**		**0.1053**		**0.2261**			
Hausman test, Chi-squared statistic	**5.54 p = 0.594**							
Joint significance of exclusion restrictions								
Chi squared							**1077.38**	
Prob>chi squared							**0.0000**	
Walt test of rho = 0								
Chi squared							**55.2116**	
Prob>chi squared							**0.0000**	
Rho (correlation of errors)							**-0.4866**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit and bivariate probit regressions. # Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. ACS denoted appropriate care score. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 6: Effect of process measures of care on CABG mortality

	Trust Level				Individual level (#)			
	Random Effects		Fixed Effects		Trust-quarter level process measures		Individual level process measures	
Age	0.002	(0.001)	0.001	(0.001)	0.002***	(0.000)	0.002***	(0.000)
Age Squared	-0.0003*	(0.000)	-0.001*	(0.000)	0.00004***	(0.000)	0.00004***	(0.000)
Male	0.028	(0.040)	0.027	(0.042)	-0.002	(0.003)	-0.002	(0.003)
Not White Ethnicity	0.044	(0.037)	0.034	(0.045)	0.015*	(0.007)	0.014*	(0.007)
Missing Ethnicity	-0.008	(0.011)	-0.006	(0.014)	-0.000	(0.005)	-0.000	(0.005)
Income Deprivation	0.038	(0.059)	-0.149	(0.123)	0.006	(0.009)	0.009	(0.009)
Achievement								
Aspirin at Discharge					-0.490	(0.332)		
Antibiotics Received <1 Surgery	0.0004	(0.041)	0.0004	(0.042)	0.003	(0.057)	-0.004	(0.007)
Antibiotic Selection					0.015	(0.077)	-0.012	(0.007)
Antibiotics Discontinued	0.038	(0.026)	0.014	(0.036)	0.005	(0.039)	-0.010	(0.006)
Exclusions								
Aspirin at Discharge	0.002	(0.052)	-0.027	(0.057)	-0.468	(0.321)		
Antibiotics Received <1 Surgery	0.033	(0.094)	0.022	(0.110)	-0.045	(0.131)	0.018	(0.010)
Antibiotic Selection	-0.036	(0.069)	-0.039	(0.080)	-0.024	(0.163)	-0.018	(0.009)
Antibiotics Discontinued	0.023	(0.024)	0.021	(0.028)	0.016	(0.035)	0.001	(0.007)
Constant	-0.189*	(0.094)	-0.078	(0.112)				
Observations	**76**		**76**		**9679**		**9679**	
Mean Variance Inflation Factor	**32.12**							
R-squared (within)	**0.0725**		**0.1275**					
R-squared (between)	**0.977**		**0.6906**					
R-squared (overall)	**0.2821**		**0.0669**		0.1034		0.1415	
Hausman test, Chi-squared statistic	**5.01**							
	p = 0.9576							

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit regressions. Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 7: Effect of Appropriate Care Score on CABG mortality

	Trust Level		Individual level (#)					
			Probit		Probit		Bivariate probit	
	Fixed Effects		Trust level ACS		Individual level ACS		Individual level ACS	
Age	-0.000	(0.002)	0.002***	(0.000)	0.002***	(0.000)	0.001***	(0.000)
Age Squared	-0.001**	(0.001)	0.00004***	(0.000)	0.00004***	(0.000)	0.00003***	(0.000)
Male	0.021	(0.040)	-0.002	(0.003)	-0.001	(0.003)	-0.001	(0.002)
Not White Ethnicity	0.036	(0.042)	0.015*	(0.007)	0.014*	(0.007)	0.012*	(0.006)
Missing Ethnicity	-0.009	(0.013)	-0.001	(0.005)	-0.001	(0.005)	-0.0003	(0.004)
Income Deprivation	-0.049	(0.113)	0.006	(0.009)	0.007	(0.009)	0.004	(0.008)
ACS	0.042	(0.051)	0.058	(0.053)	-0.016***	(0.004)	0.015*	(0.007)
Predicted Error								
Constant	-0.019	(0.118)						
Observations	73		9679		9679		9679	
R-squared (within)	0.1494							
R-squared (between)	0.6879							
R-squared (overall)	0.0373		0.1008		0.1126			
Hausman test, Chi-squared statistic	7.65 p = 0.3643							
Joint significance of exclusion restrictions								
Chi squared							326.13	
Prob>chi squared							0.0000	
Walt test of rho = 0								
Chi squared							12.6344	
Prob>chi squared							0.0004	
Rho (correlation of errors)							-0.5674	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit and bivariate probit regressions. # Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. ACS denoted appropriate care score. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 8: Effect of process measures of care on hip and knee readmission proportions

	Trust Level				Individual level (#)			
	Random Effects		Fixed Effects		Trust-quarter level process measures		Individual level process measures	
Age	0.005*	(0.002)	0.006*	(0.003)	0.004***	(0.000)	0.003***	(0.000)
Age Squared	-0.001	(0.001)	-0.0002	(0.001)	0.0001***	(0.000)	0.0001***	(0.000)
Male	0.002	(0.057)	0.023	(0.058)	0.031***	(0.003)	0.031***	(0.003)
Not White Ethnicity	0.027	(0.109)	0.023	(0.151)	0.006	(0.008)	0.006	(0.008)
Missing Ethnicity	-0.031	(0.021)	-0.027	(0.022)	-0.032***	(0.005)	-0.032***	(0.005)
Income Deprivation	0.238*	(0.099)	0.084	(0.224)	0.067***	(0.012)	0.065***	(0.012)
Achievement								
Antibiotics Received <1 Surgery	-0.055	(0.043)	-0.059	(0.047)	-0.014	(0.032)	-0.015**	(0.005)
Antibiotic Selection	-0.043	(0.034)	-0.048	(0.036)	-0.050**	(0.019)	-0.003	(0.006)
Antibiotics Discontinued	-0.046	(0.108)	-0.038	(0.115)	0.027	(0.072)	-0.005	(0.009)
VTE Prophylactic Ordered	-0.033	(0.058)	-0.025	(0.060)	0.021	(0.042)	0.030**	(0.011)
VTE Prophylactic Received	0.029	(0.043)	0.018	(0.046)	0.006	(0.030)	-0.016*	(0.008)
Exclusions								
Antibiotics Received <1 Surgery	-0.110	(0.140)	-0.148	(0.146)	-0.129	(0.097)	-0.028*	(0.012)
Antibiotic Selection	0.112	(0.146)	0.147	(0.156)	0.049	(0.106)	0.002	(0.012)
Antibiotics Discontinued	-0.089	(0.148)	-0.084	(0.159)	0.017	(0.103)	0.029*	(0.012)
VTE Prophylactic Ordered	-0.029	(0.086)	-0.041	(0.089)	0.079	(0.059)	0.029**	(0.010)
VTE Prophylactic Received								
Constant	-0.074	(0.177)	-0.164	(0.223)			0.170*	(0.068)
Observations	**408**		**408**		**72213**		**72213**	
Mean Variance Inflation Factor	**6.6**							
R-squared (within)	**0.0613**		**0.0662**					
R-squared (between)	**0.2437**		**0.0321**					
R-squared (overall)	**0.0909**		**0.0461**		0.0327		0.0339	
Hausman test, Chi-squared statistic	**10.02**							
	p = 0.8186							

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit regressions. Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 9: Effect of Appropriate Care Score on hip and knee readmission proportions

	Trust Level		Individual level (#)					
			Probit		Probit		Bivariate probit	
	Fixed Effects		Trust level ACS		Individual level ACS		Individual level ACS	
Age	0.006*	(0.003)	0.004***	(0.000)	0.004***	(0.000)	0.003***	(0.000)
Age Squared	-0.000	(0.001)	0.0001***	(0.000)	0.0001***	(0.000)	0.00005***	(0.000)
Male	0.013	(0.057)	0.031***	(0.003)	0.032***	(0.003)	0.026***	(0.002)
Non-White Ethnicity	0.041	(0.148)	0.006	(0.008)	0.006	(0.008)	0.006	(0.007)
Missing Ethnicity	-0.029	(0.020)	-0.032***	(0.005)	-0.033***	(0.005)	-0.027***	(0.004)
Income Deprivation	0.066	(0.222)	0.067***	(0.012)	0.067***	(0.012)	0.051***	(0.009)
ACS	-0.148***	(0.036)	-0.074**	(0.026)	-0.016***	(0.003)	-0.015	(0.008)
Predicted Error								
Constant	-0.140	(0.207)						
Observations	408		72213		72213		72213	
R-squared (within)	0.056							
R-squared (between)	0.0208							
R-squared (overall)	0.0375		0.0326		0.0328			
Hausman test, Chi-squared statistic	8.78							
	p = 0.2687							
Joint significance of exclusion restrictions								
Chi squared							5023.57	
Prob>chi squared							P<0.001	
Walt test of rho = 0								
Chi squared							0.08283	
Prob>chi squared							0.7735	
Rho (correlation of errors)							0.0078	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit and bivariate probit regressions. # Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. ACS denoted appropriate care score. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 10: Effect of process measures of care on pneumonia mortality

	Trust Level				Individual level (#)			
	Random Effects		Fixed Effects		Trust-quarter level process measures		Individual level process measures	
Age	0.008***	(0.002)	0.007***	(0.002)	0.008***	(0.000)	0.006***	(0.000)
Age Squared	0.00003	(0.000)	0.00004	(0.000)	0.00005	(0.000)	-0.00005	(0.000)
Male	-0.052	(0.035)	-0.038	(0.034)	0.019***	(0.003)	0.018***	(0.002)
Not White Ethnicity	-0.081	(0.075)	-0.250**	(0.095)	-0.032***	(0.007)	-0.038***	(0.006)
Missing Ethnicity	0.050*	(0.023)	0.034	(0.023)	0.016**	(0.006)	0.016**	(0.005)
Income Deprivation	0.046	(0.090)	0.516**	(0.156)	-0.001	(0.010)	0.041***	(0.010)
Achievement								
Oxygenation Assessment					0.251	(0.150)	-0.075***	(0.019)
Antibiotic Selection	0.033	(0.067)	-0.051	(0.069)	0.010	-0.044	0.002	(0.009)
Blood Cultures	-0.285***	(0.083)	-0.233**	(0.084)	-0.083	(0.059)	-0.048***	(0.008)
Antibiotics Received	0.025	(0.059)	0.035	(0.058)	0.082*	(0.042)	0.021***	(0.005)
Smoking Cessation					-0.112	(0.066)	0.033	(0.030)
Exclusions								
Oxygenation Assessment	0.089*	(0.037)	0.121**	(0.039)	0.309*	(0.150)	0.097***	(0.019)
Antibiotic Selection	-0.012	(0.078)	-0.099	(0.081)	-0.010	(0.052)	0.060***	(0.009)
Blood Cultures	-0.262***	(0.073)	-0.233**	(0.073)	-0.064	(0.051)	-0.022**	(0.008)
Antibiotics Received	-0.015	(0.053)	0.028	(0.055)	0.069	(0.039)	0.024***	(0.005)
Smoking Cessation	-0.147	(0.079)	-0.033	(0.084)	0.068	(0.066)	0.507***	(0.024)
Constant	-0.025	(0.138)	-0.078	(0.136)				
Observations	**425**		**425**		**98771**		**98771**	
Mean Variance Inflation Factor	**5.88**							
R-squared (within)	**0.1244**		**0.1737**					
R-squared (between)	**0.6785**		**0.0138**					
R-squared (overall)	**0.3474**		**0.0555**		**0.0738**		**0.1927**	
Hausman test, Chi-squared statistic	**64.73**							
	P<0.001							

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit regressions. Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 11: Effect of Appropriate Care Score on pneumonia mortality

	Trust Level		Individual level (#)					
			Probit		Probit		Bivariate probit	
	Random Effects		Trust level ACS		Individual level ACS		Individual level ACS	
Age	0.005***	(0.001)	0.008***	(0.000)	0.008***	(0.000)	0.003***	(0.000)
Age Squared	-0.000	(0.000)	0.00001	(0.000)	0.00001	(0.000)	0.0000003	(0.000)
Male	0.039	(0.045)	0.019***	(0.003)	0.021***	(0.003)	0.009***	(0.001)
Non-White Ethnicity	-0.068	(0.084)	-0.032***	(0.007)	-0.031***	(0.007)	-0.008**	(0.003)
Missing Ethnicity	0.046**	(0.018)	0.015*	(0.006)	0.013*	(0.006)	0.004	(0.002)
Income Deprivation	-0.031	(0.093)	-0.001	(0.010)	-0.002	(0.010)	-0.003	(0.004)
ACS	-0.109***	(0.028)	0.043	(0.029)	-0.165***	(0.003)	-0.017**	(0.007)
Predicted Error								
Constant	-0.047	(0.099)						
Observations	409		98771		98771		98771	
R-squared (within)	0.1042							
R-squared (between)	0.0036							
R-squared (overall)	0.0158		0.0734		0.1082			
Hausman test, Chi-squared statistic	24.89							
	p = 0.0004							
Joint significance of exclusion restrictions								
Chi squared							2389.79	
Prob>chi squared							P<0.001	
Walt test of rho = 0								
Chi squared							48.0705	
Prob>chi squared							P<0.001	
Rho (correlation of errors)							-0.2599	
Lives saved	537						84	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. # Marginal effects displayed for probit and bivariate probit regressions. Marginal effects are average marginal effects. Models also include quarter dummies and Trust fixed effects. Fixed effects estimations did not remove patient fixed variables such as gender as the case mix of hospitals changes over time. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 12: Effect of process measures of care and Appropriate Care Score on alternative AMI mortality proportions

	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	
Achievement												
Aspirin at Arrival	-0.404***	(0.084)	-0.443***	(0.087)	-0.349***	(0.078)	-0.340***	(0.082)	-0.277***	(0.072)	-0.280***	(0.076)
Fibrinolytic Therapy	0.171	(0.172)	0.088	(0.179)	0.194	(0.162)	0.180	(0.168)	0.096	(0.150)	0.104	(0.157)
Exclusions												
Aspirin at Arrival	-0.288**	(0.089)	-0.354***	(0.093)	-0.321***	(0.083)	-0.302***	(0.088)	-0.242**	(0.077)	-0.247**	(0.082)
Aspirin at Discharge	-0.000	(0.043)	0.082	(0.048)	0.008	(0.041)	0.028	(0.045)	-0.001	(0.039)	0.017	(0.042)
ACEI or ARB for LVSD	-0.036	(0.027)	-0.028	(0.032)	-0.056*	(0.026)	-0.031	(0.030)	-0.026	(0.025)	-0.016	(0.028)
Smoking Cessation Advice	-0.047	(0.047)	-0.072	(0.048)	-0.084	(0.044)	-0.094*	(0.045)	-0.053	(0.040)	-0.066	(0.042)
Beta Blocker Discharge	0.069	(0.043)	0.043	(0.049)	0.085*	(0.042)	0.060	(0.046)	0.064	(0.039)	0.049	(0.043)
Fibrinolytic Therapy	0.083	(0.156)	-0.043	(0.164)	0.083	(0.147)	0.057	(0.155)	-0.030	(0.137)	-0.036	(0.144)
Primary PCI	-0.026	(0.036)	0.029	(0.048)	-0.069	(0.036)	-0.082	(0.046)	-0.057	(0.034)	-0.046	(0.042)
R-squared (within)	0.2838		0.316		0.1665		0.1735		0.1391		0.142	
R-squared (between)	0.5452		0.2211		0.5612		0.379		0.177		0.0495	
R-squared (overall)	0.3448		0.2188		0.2852		0.2331		0.1474		0.1175	
Hausman test, Chi-squared statistic	35.7				19.13				5.42			
	p = 0.001				p = 0.160				p = 0.9879			
Appropriate care score	-0.249***	(0.063)	-0.283***	(0.071)	-0.194***	(0.058)	-0.199**	(0.063)	-0.199***	(0.054)	-0.213***	(0.058)
R-squared (within)	0.1508		0.1589		0.0702		0.0735		0.0632		0.0652	
R-squared (between)	0.5182		0.3493		0.5289		0.4048		0.1597		0.1024	
R-squared (overall)	0.2375		0.1908		0.2063		0.1739		0.0827		0.0696	
Hausman test, Chi-squared statistic	8.65				9.24				2.6			
	p = 0.2789				p = 0.2359				p = 0.9194			
Observations	**334**											
Lives saved	**279**		**317**		**217**		**223**		**223**		**239**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. Coefficients signify proportions; a 0.01 coefficient signifies 1%

Table 13: Effect of process measures of care and Appropriate Care Score on alternative pneumonia mortality proportions

	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	
Achievement												
Antibiotic Selection	0.055	(0.077)	-0.059	(0.083)	-0.006	(0.062)	-0.049	(0.061)	-0.008	(0.057)	-0.027	(0.060)
Blood Cultures	-0.270**	(0.090)	-0.216*	(0.095)	-0.199**	(0.072)	-0.125	(0.069)	-0.233***	(0.066)	-0.238***	(0.068)
Antibiotics Received	0.059	(0.066)	0.078	(0.067)	0.060	(0.052)	0.065	(0.049)	0.030	(0.048)	0.034	(0.048)
Exclusions												
Oxygenation Assessment	0.089*	(0.041)	0.102*	(0.045)	0.030	(0.033)	0.062	(0.033)	-0.024	(0.030)	-0.012	(0.032)
Antibiotic Selection	0.031	(0.086)	-0.078	(0.094)	-0.018	(0.070)	-0.058	(0.069)	0.035	(0.064)	0.002	(0.068)
Blood Cultures	-0.180*	(0.081)	-0.149	(0.087)	-0.161*	(0.065)	-0.099	(0.064)	-0.187**	(0.060)	-0.181**	(0.063)
Antibiotics Received	-0.007	(0.057)	0.035	(0.060)	0.007	(0.045)	0.032	(0.044)	-0.004	(0.042)	0.019	(0.043)
Smoking Cessation	-0.176*	(0.086)	-0.110	(0.096)	-0.217**	(0.070)	-0.108	(0.070)	-0.102	(0.065)	-0.105	(0.069)
R-squared (within)	0.1865		0.2157		0.0531		0.0907		0.1094		0.1342	
R-squared (between)	0.7259		0.2432		0.7626		0.0769		0.1109		0.0199	
R-squared (overall)	0.4005		0.2101		0.4536		0.0713		0.0991		0.0067	
Hausman test, Chi-squared statistic	26.48				120.01				29.08			
	p = 0.0146				p = 0.0000				p = 0.006			
Appropriate care score	-0.081**	(0.031)	-0.081*	(0.032)	-0.064*	(0.030)	-0.060*	(0.030)	-0.086**	(0.029)	-0.093**	(0.030)
R-squared (within)	0.1026		0.1144		0.0741		0.0844		0.0873		0.0963	
R-squared (between)	0.2592		0.0074		0.1848		0.0154		0.0016		0.0404	
R-squared (overall)	0.1706		0.0392		0.1189		0.0042		0.0421		0.0029	
Hausman test, Chi-squared statistic	15.56				29.16				3			
	p = 0.0295				p = 0.0001				p = 0.8849			
Observations	**333**											
Lives saved	**238**		**238**		**188**		**177**		**253**		**274**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Fixed effects estimations did not remove patient fixed variables such as gender as case mix of hospitals in each quarter change. Coefficients signify proportions; a 0.01 coefficient signifies 1%.

Table 14: Estimated Effect of Advancing Quality

Effect of Advancing Quality

	Lives saved				Readmissions avoided Hip and knee replacements	
	AMI		Pneumonia			
	Trust	Individual	Trust	Individual	Trust	Individual
Full model (2008-2013)	467	45	537	84	1122	121
Number of Patients	**45719**		**98771**		**72213**	
Percentage of patients	**1.021**	**0.099**	**0.544**	**0.085**	**1.553**	**0.168**
Robustness (2008-2012)	**Trust (% of patients)**		**Trust (% of patients)**			
In-Hospital Mortality	317 (0.984)		238 (0.331)			
30 day in-hospital mortality	223 (0.692)		188 (0.261)			
Risk adjusted 30 day in-hospital mortality	239 (0.741)		274 (0.380)			
Patients	**32243**		**72038**			

Results presented in this Table are from: Table 4; Table 6; Table 10; Table 11; and Table 12.

Figure 6: Decision tree under the multinomial logit

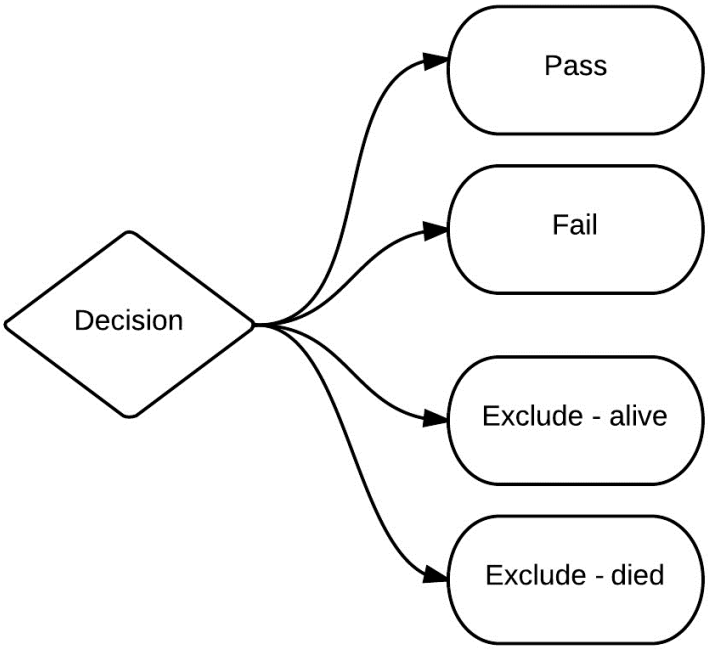


Figure 7: Decision tree under sequential logit

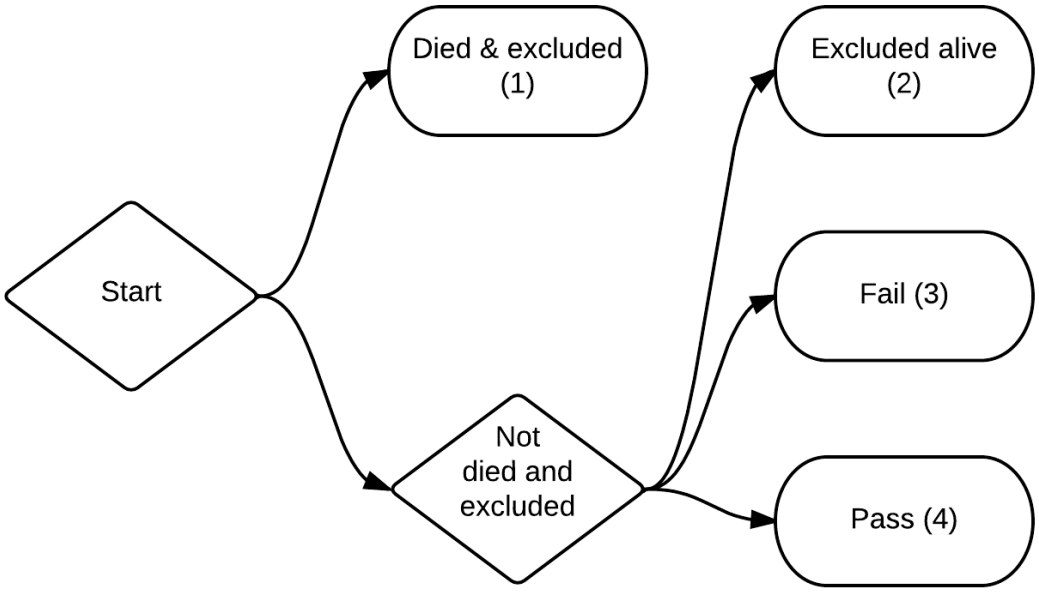


Table 15: Descriptive statistics

	AMI	Heart Failure	Hip and knee	Pneumonia
Excluded and died	0.081	0.159	0.005	0.201
Excluded and alive	0.631	0.385	0.074	0.481
Failed	0.009	0.114	0.049	0.049
Achieved	0.279	0.342	0.872	0.27
Income score (Standard deviation)	0.820 (0.131)	0.816 (0.131)	0.843 (0.122)	0.800 (0.137)
Age in years (Standard deviation)	69.7 (14.2)	77.6 (11.6)	71.1 (11.4)	72.7 (16.5)
Age squared (Standard deviation)	2.066 (2.612)	1.646 (2.155)	1.307 (2.105)	2.734 (4.391)
Male	0.62	0.524	0.387	0.498
Ethnicity				
Black and minority ethnic group	0.057	0.058	0.029	0.044
Missing	0.084	0.054	0.058	0.061
Care Quality Commission - quality				
Good	0.484	0.551	0.596	0.543
Poor	0.204	0.151	0.111	0.124
Care Quality Commission - financial				
Good	0.333	0.298	0.305	0.253
Poor	0.229	0.241	0.258	0.266
Trust type				
Medium	0.318	0.315	0.301	0.358
Large	0.345	0.41	0.455	0.336
Specialist/teaching	0.228	0.162	0.139	0.194
Mean Trust achievement proportion (Standard deviation)	0.820 (0.045)	0.819 (0.048)	0.822 (0.050)	0.812 (0.052)
Foundation Trust status	0.584	0.549	0.549	0.482
Oct 2008 - Dec 2008	0.031	0.04	0.039	0.039
Jan 2009 - Mar 2009	0.04	0.057	0.055	0.053
Apr 2009 - Jun 2009	0.041	0.056	0.053	0.044
Jul 2009 - Sept 2009	0.053	0.05	0.053	0.038
Oct 2009 - Dec 2009	0.059	0.054	0.058	0.053
Jan 2010 - Mar 2010	0.042	0.044	0.044	0.04
Apr 2010 - Jun 2010	0.054	0.061	0.057	0.051
Jul 2010 - Sept 2010	0.052	0.054	0.057	0.043
Oct 2010 - Dec 2010	0.061	0.058	0.057	0.065
Jan 2011 - Mar 2011	0.055	0.057	0.054	0.07
Apr 2011 - Jun 2011	0.056	0.061	0.055	0.057
Jul 2011 - Sept 2011	0.052	0.055	0.06	0.048
Oct 2011 - Dec 2011	0.055	0.055	0.061	0.058
Jan 2012 - Mar 2012	0.056	0.06	0.061	0.069
Apr 2012 - Jun 2012	0.072	0.06	0.054	0.062
Jul 2012 - Sept 2012	0.07	0.054	0.058	0.052
Oct 2012 - Dec 2012	0.08	0.061	0.062	0.07
Jan 2013 - Mar 2013	0.073	0.062	0.063	0.089
Observations	365752	146256	363260	561012

Numbers of observations are the number of spell-process measure level observations. Values in the table of 0.01 signify 1%.

Table 16: Descriptive statistics of process measures

		Excluded and died	Excluded and alive	Failed	Achieved
AMI	Aspirin at arrival	0.050	0.245	0.011	0.693
	Fibrinolytic therapy	0.084	0.878	0.006	0.032
	Smoking cessation advice	0.086	0.736	0.013	0.165
	Aspirin at discharge	0.086	0.329	0.005	0.580
	ACEI/ARB for LVSD	0.086	0.790	0.002	0.121
	Beta blocker at discharge	0.086	0.373	0.017	0.524
Heart Failure	Evaluation of LVS function	0.159	0.129	0.054	0.659
	ACEI/ARB for LVSD	0.159	0.527	0.019	0.295
	Discharge instructions	0.158	0.128	0.359	0.355
	Smoking cessation advice	0.159	0.757	0.024	0.061
Hip and knee	Antibiotics received	0.005	0.066	0.074	0.856
	Antibiotic selection	0.005	0.064	0.065	0.866
	Antibiotics discontinued	0.007	0.091	0.024	0.877
	VTE prophylaxis ordered	0.005	0.074	0.027	0.894
	VTE prophylaxis received	0.005	0.074	0.057	0.864
Pneumonia	Oxygenation assessment	0.155	0.241	0.006	0.598
	Antibiotic selection	0.199	0.452	0.040	0.309
	Blood cultures	0.217	0.614	0.038	0.132
	Antibiotics received	0.181	0.392	0.105	0.322
	Smoking cessation	0.238	0.618	0.059	0.086
	Observations - AMI		**45719**		
	Observations - Heart failure		**36564**		
	Observations - Hip and knee		**72652**		
	Observations - Pneumonia		**93502**		

Numbers of observations are the number of spell level observations. Values in the table of 0.01 signify 1%.

Table 17: Multinomial logit results for condition specific regressions - Patient characteristics

	AMI			Heart failure			Hip and knee replacement			Pneumonia		
	Excluded died vs Pass	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**
Income	-0.038	0.069**	-0.006**	-0.034	0.098***	0.009	-0.007*	0.042***	-0.010	0.023	0.054**	-0.027***
	(0.022)	(0.023)	(0.002)	(0.033)	(0.027)	(0.011)	(0.003)	(0.013)	(0.007)	(0.021)	(0.021)	(0.005)
Interaction between income and year												
Base category – year 1												
Year 2	-0.036	0.021	0.002	-0.048	-0.054	0.026	-0.002	-0.077***	0.006	-0.005	-0.021	0.015
	(0.035)	(0.036)	(0.004)	(0.055)	(0.044)	(0.020)	(0.006)	(0.023)	(0.011)	(0.036)	(0.034)	(0.009)
Year 3	0.018	0.015	0.001	0.027	-0.010	-0.056***	-0005	-0.010***	0.012	-0.004	-0.044	0.017*
	(0.029)	(0.029)	(0.004)	(0.045)	(0.036)	(0.017)	(0.005)	(0.018)	(0.010)	(0.027)	(0.027)	(0.007)
Year 4	-0.016	0.070*	-0.017	0.008	-0.049	-0.055**	-0.003	-0.103***	0.018	-0.032	-0.017	0.016*
	(0.028)	(0.029)	(0.004)	(0.045)	(0.036)	(0.018)	(0.005)	(0.018)	(0.012)	(0.028)	(0.028)	(0.007)
Year 5	-0.005	0.022	-0.0002	-0.091*	0.002	0.0.03	-0.001	-0.049**	0.001	-0.040	-0.004	0.023**
	(0.028)	(0.028)	(0.004)	(0.044)	(0.035)	(0.019)	(0.004)	(0.019)	(0.012)	(0.027)	(0.027)	(0.007)
Male	0.001	-0.003	-0.001**	0.015***	-0.029***	-0.012***	0.002***	0.008***	-0.001	0.012***	-0.017***	-0.001
	(0.003)	(0.002)	(0.000)	(0.004)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Age	0.049***	-0.042***	0.001***	0.056***	0.002	-0.006***	0.008***	0.016***	0.004***	0.070***	-0.038***	-0.008***
	(0.001)	(0.001)	(0.000)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Age squared	-0.001	-0.003***	0.000***	0.005***	0.003***	-0.0005	-0.000	0.003***	0.001***	-0.002***	0.001	-0.0002*
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Ethnicity												
Base category – White												
Not-white	0.001	-0.009	0.002	-0.035***	0.020**	0.030***	-0.000	-0.023***	-0.007**	-0.026***	0.007	0.001
	(0.006)	(0.005)	(0.001)	(0.008)	(0.006)	(0.004)	(0.001)	(0.004)	(0.003)	(0.006)	(0.006)	(0.002)
Missing	0.011*	0.011*	0.001	0.020*	0.002	0.011**	0.002	-0.006	0.006**	0.020***	-0.023***	0.001
	(0.005)	(0.005)	(0.001)	(0.009)	(0.007)	(0.004)	(0.001)	(0.003)	(0.002)	(0.005)	(0.005)	(0.001)
Observations		**365752**			**146256**			**363260**			**561012**	
Pseudo R-squared		**0.0407**			**0.0249**			**0.0575**			**0.03555**	
Log likelihood		**-313262.52**			**-181699.3**			**-167103.46**			**-636065.73**	
Hausman test, number of outcomes failed		**1**			**0**			**2**			**0**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 18: Multinomial logit results for condition specific regressions - Trust characteristics

	AMI			Heart Failure			Hip and knee replacement			Pneumonia		
	Excluded died vs Pass	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**
Foundation Trust	-0.003	0.027***	0.000	0.001	-0.011	-0.010***	-0.000	0.040***	0.036***	0.022***	0.015***	0.016***
	(0.005)	(0.005)	(0.001)	(0.007)	(0.006)	(0.003)	(0.001)	(0.003)	(0.002)	(0.004)	(0.004)	(0.001)
Care quality commission - quality												
Base category – excellent												
Good	-0.007*	0.022***	-0.000	-0.003	-0.005	-0.015***	0.002***	0.016***	0.024***	-0.000	0.040***	-0.006***
	(0.003)	(0.003)	(0.001)	(0.005)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)
Poor	-0.005	-0.056***	-0.002***	-0.011	-0.020***	-0.010**	0.002	-0.018***	-0.009***	-0.005	-0.004	0.001
	(0.005)	(0.005)	(0.001)	(0.007)	(0.006)	(0.003)	(0.001)	(0.003)	(0.001)	(0.005)	(0.005)	(0.001)
Care quality commission - financial resources												
Base category – excellent												
Good	-0.014***	0.018***	-0.000	-0.001	0.010*	0.005*	0.001	0.034***	0.029***	0.004	0.003	0.008***
	(0.003)	(0.003)	(0.001)	(0.005)	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)
Poor	0.006	-0.006	-0.001	-0.013	0.021***	0.002	-0.000	0.029***	0.043***	0.015**	0.013**	0.008***
	(0.005)	(0.005)	(0.001)	(0.007)	(0.006)	(0.003)	(0.001)	(0.003)	(0.002)	(0.005)	(0.005)	(0.001)
Trust type												
Base category – small												
Medium	-0.021***	0.075***	-0.002**	0.011	-0.005	-0.057***	0.002*	0.046***	0.017***	-0.016***	0.097***	-0.024***
	(0.005)	(0.005)	(0.001)	(0.007)	(0.006)	(0.004)	(0.001)	(0.003)	(0.002)	(0.005)	(0.005)	(0.002)
Large	-0.001	0.011*	-0.004***	0.010	-0.024***	-0.014***	0.000	0.026***	0.004**	-0.002	0.056***	-0.014***
	(0.005)	(0.004)	(0.001)	(0.006)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)	(0.004)	(0.004)	(0.001)
Specialist/Teaching	-0.016**	0.023***	-0.003***	0.003	-0.019**	-0.049***	0.003*	0.028***	0.025***	-0.014*	0.086***	-0.025***
	(0.006)	(0.006)	(0.001)	(0.009)	(0.007)	(0.004)	(0.001)	(0.003)	(0.003)	(0.006)	(0.005)	(0.002)
Mean Trust Performance	-0.129***	-0.180***	0.040***	0.176**	-0.216***	-0.052*	0.008	-0.270***	-0.162***	0.030	0.028	-0.090***
	(0.037)	(0.037)	(0.006)	(0.057)	(0.046)	(0.024)	(0.007)	(0.025)	(0.015)	(0.038)	(0.038)	(0.010)
Observations		**365752**			**146256**			**363260**			**561012**	
Pseudo R-squared		**0.0407**			**0.0249**			**0.0575**			**0.03555**	
Log likelihood		**-313262.52**			**-181699.3**			**-167103.46**			**-636065.73**	
Hausman test, number of outcomes failed		**1**			**0**			**2**			**0**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and patient characteristics variables such as: income score; income and time interactions; age; age squared; gender and ethnicity. Coefficients in the table of 0.01 signify 1%.

Table 19: Multinomial logit results for AMI: Aspirin at arrival; fibrinolytic therapy and smoking cessation advice

	Aspirin at arrival			Fibrinolytic therapy			Smoking cessation advice		
	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass
Income	-0.027	0.015	-0.006	-0.038	0.047	-0.004	-0.044	0.300***	-0.012
	(0.018)	(0.036)	(0.006)	(0.023)	(0.025)	(0.004)	(0.023)	(0.036)	(0.006)
Interaction between income and year									
Base category – year 1									
Year 2	-0.027	-0.007	0.017	-0.028	0.020	0.004	-0.039	0.106	-0.016
	(0.030)	(0.059)	(0.013)	(0.038)	(0.042)	(0.008)	(0.038)	(0.059)	(0.012)
Year 3	0.016	0.034	0.010	0.019	-0.026	0.007	0.019	-0.0001	-0.011
	(0.024)	(0.048)	(0.010)	(0.030)	(0.034)	(0.007)	(0.031)	(0.047)	(0.010)
Year 4	-0.005	0.082	0.002	-0.020	-0.021	0.010	-0.018	0.055	0.004
	(0.024)	(0.049)	(0.011)	(0.030)	(0.039)	(0.015)	(0.030)	(0.047)	(0.012)
Year 5	-0.023	0.005	0.005	-0.007	-0.066	0.025	-0.003	0.036	0.003
	(0.024)	(0.043)	(0.011)	(0.030)	(0.041)	(0.014)	(0.030)	(0.045)	(0.010)
Male	-0.003	-0.004	-0.005***	0.002	-0.008*	-0.000	0.001	0.014**	-0.001
	(0.002)	(0.004)	(0.001)	(0.003)	(0.003)	(0.001)	(0.003)	(0.004)	(0.001)
Age	0.029***	-0.040***	0.003***	0.051***	-0.036***	-0.001***	0.051***	0.052***	-0.002***
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)
Age squared	-0.0002	-0.004***	-0.0001	-0.001	0.003**	-0.0001	0.001	0.013***	-0.0001
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Ethnicity									
Base category – white									
Not-white	-0.003	-0.028***	-0.001	0.002	-0.004	0.004	0.001	0.041***	-0.001
	(0.005)	(0.008)	(0.002)	(0.006)	(0.007)	(0.002)	(0.006)	(0.008)	(0.002)
Missing	0.011**	0.051***	0.003	0.011*	-0.014*	0.001	0.011*	0.008	0.003
	(0.004)	(0.007)	(0.002)	(0.005)	(0.006)	(0.001)	(0.005)	(0.007)	(0.002)
Observations					**45719**				
Pseudo R-squared		**0.0896**			**0.133**			**0.1256**	
Log likelihood		**-33282.69**			**-18419.015**			**-31652.108**	
Hausman test, number of outcomes failed		**1**			**0**			**1**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 20: Multinomial logit results for AMI: Aspirin at discharge; ACEI/ARB for LVSD and beta blocker at discharge

	Aspirin at discharge			ACEI/ARB for LVSD			Beta blocker at discharge		
	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass
Income	-0.042	0.080*	-0.009*	-0.041	0.040	-0.025**	-0.043	0.120***	-0.003
	(0.023)	(0.040)	(0.004)	(0.023)	(0.041)	(0.008)	(0.023)	(0.033)	(0.003)
Interaction between income and year									
Base category – year 1									
Year 2	-0.037	-0.062	0.012	-0.038	-0.039	0.022	-0.036	-0.078	-0.004
	(0.038)	(0.064)	(0.010)	(0.038)	(0.066)	(0.016)	(0.038)	(0.054)	(0.005)
Year 3	0.020	0.032	0.013	0.019	0.009	0.021	0.022	-0.058	-0.004
	(0.031)	(0.052)	(0.008)	0.0312)	(0.053)	(0.013)	(0.031)	(0.044)	(0.004)
Year 4	-0.015	0.058	0.005	-0.018	0.103	0.026	-0.015	0.030	0.002
	(0.030)	(0.053)	(0.007)	(0.030)	(0.054)	(0.015)	(0.030)	(0.044)	(0.005)
Year 5	-0.0004	-0.020	0.015	0.001	0.005	-0.011	0.002	-0.027	-0.002
	(0.030)	(0.050)	(0.008)	(0.030)	(0.052)	(0.012)	(0.030)	(0.042)	(0.005)
Male	0.001	0.009*	-0.001	0.001	0.010*	-0.002	0.001	-0.028***	0.000
	(0.003)	(0.005)	(0.001)	(0.003)	(0.005)	(0.001)	(0.003)	(0.004)	(0.000)
Age	0.052***	-0.065***	0.002***	0.052***	-0.061***	0.006***	0.052***	-0.069***	0.001***
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.000)
Age squared	-0.0003	-0.012***	0.0005***	-0.0003	-0.010***	0.001***	-0.0005	0.002*	0.00003
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Ethnicity									
Base category – white									
Not-white	0.002	-0.038***	0.005*	0.002	-0.054***	0.001	0.002	-0.000	-0.001
	(0.006)	(0.009)	(0.003)	(0.006)	(0.009)	(0.003)	(0.006)	(0.009)	(0.001)
Missing	0.011*	0.060***	-0.001	0.011*	0.042***	-0.001	0.011*	-0.027***	0.000
	(0.005)	(0.008)	(0.001)	(0.005)	(0.008)	(0.002)	(0.005)	(0.007)	(0.001)
Observations					**45719**				
Pseudo R-squared		**0.0721**			**0.0619**			**0.067**	
Log likelihood		**-39026.215**			**-28622.608**			**-42083.008**	
Hausman test, number of outcomes failed		**1**			**1**			**1**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 21: Multinomial logit results for heart failure process measures and patient characteristics

	Evaluation of LVS function			ACEI/ARB for LVSD			Discharge instructions			Smoking cessation advice		
	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass
Income	-0.039	0.069*	-0.024	-0.035	0.102*	0.007	-0.012	0.039	0.138***	-0.039	0.166***	-0.034***
	(0.033)	(0.033)	(0.016)	(0.033)	(0.045)	(0.010)	(0.033)	(0.031)	(0.040)	(0.033)	(0.039)	(0.009)
Interaction between income and year												
Base category – year 1												
Year 2	-0.044	-0.050	0.066	-0.047	-0.104	-0.003	-0.063	-0.111*	0.023	-0.048	0.030	-0.007
	(0.055)	(0.058)	0.034)	(0.055)	(0.076)	(0.017)	(0.055)	(0.051)	(0.068)	(0.056)	(0.065)	(0.017)
Year 3	0.032	0.017	-0.036	0.027	-0.048	0.003	0.005	-0.017	-0.200***	0.028	-0.009	-0.027
	(0.045)	(0.045)	(0.025)	(0.045)	(0.060)	(0.015)	(0.045)	(0.041)	(0.053)	0.045)	(0.052)	(0.014)
Year 4	0.015	-0.016	-0.038	0.006	-0.113	-0.024	-0.011	-0.064	-0.193***	0.007	-0.020	-0.010
	(0.045)	(0.042)	(0.025)	(0.045)	(0.061)	(0.016)	(0.045)	(0.042)	(0.055)	(0.045)	(0.052)	(0.015)
Year 5	-0.087*	-0.044	0.014	-0.091	-0.021	-0.002	-0.118	-0.008	-0.070	-0.089*	0.076	0.006
	(0.044)	(0.041)	(0.026)	(0.044)	(0.060)	(0.017)	(0.044)	(0.041)	0.057)	(0.044)	(0.052)	(0.017)
Male	0.015***	0.005	-0.012***	0.015***	-0.094***	0.004**	0.015***	-0.007	-0.043***	0.015***	-0.020***	0.002
	(0.004)	(0.004)	(0.002)	(0.004)	(0.005)	(0.001)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.002)
Age	0.056***	-0.012***	0.008***	0.055***	0.009***	0.000	0.056***	0.009***	-0.017***	0.054***	0.005	-0.013***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)
Age squared	0.005***	0.003***	0.004***	0.005***	0.003*	0.00005	0.005***	0.009***	-0.003**	0.006***	0.004**	-0.003***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Ethnicity												
Base category - white												
Not-white	-0.036***	-0.019**	0.016**	-0.035***	0.049***	0.004	-0.035***	-0.034***	0.112***	-0.036***	0.068***	-0.007*
	(0.008)	(0.007)	(0.006)	(0.008)	(0.011)	(0.003)	(0.008)	(0.007)	(0.011)	(0.008)	(0.009)	(0.003)
Missing	0.019*	0.010	0.014**	0.018*	0.010	0.002	0.024**	0.020*	0.030**	0.019*	-0.032**	0.004
	(0.009)	(0.008)	(0.005)	(0.009)	(0.012)	(0.003)	(0.009)	(0.008)	(0.011)	(0.009)	(0.011)	(0.004)
Observations	**36564**											
Pseudo R-squared	**0.0441**			**0.0402**			**0.0539**			**0.069**		
Log likelihood	**-34537.14**			**-37430.709**			**-44638.001**			**-25948.406**		
Hausman test, number of outcomes failed	**1**			**0**			**1**			**0**		

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 22: Multinomial logit results for hip and knee replacement process measures and patient characteristics

	Antibiotics received			Antibiotic selection			Antibiotics discontinued			VTE prophylaxis ordered			VTE prophylaxis received		
	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass
Income	-0.007	0.025	-0.011	-0.008*	0.037**	0.025	-0.014**	0.031	-0.007	-0.003	0.058***	-0.009	-0.003	0.057***	-0.040***
	(0.004)	(0.013)	(0.014)	(0.004)	(0.013)	(0.014)	(0.005)	(0.016)	(0.008)	(0.004)	(0.015)	(0.007)	(0.004)	(0.015)	(0.012)
Interaction between income and year															
Base category - year 1															
Year 2	0.002	-0.010***	-0.010	0.004	-0.100***	-0.021	-0.007	-0.104***	0.029	-0.004	-0.033	0.032*	-0.005	-0.036	0.015
	(0.007)	(0.025)	(0.025)	(0.007)	(0.025)	(0.025)	(0.008)	(0.030)	(0.017)	(0.007)	(0.029)	(0.013)	(0.007)	(0.029)	(0.019)
Year 3	-0.006	-0.089***	0.047*	-0.005	-0.099***	-0.032	-0.002	-0.129***	-0.007	-0.006	-0.086***	0.002	-0.005	-0.085***	0.046**
	(0.005)	(0.020)	(0.022)	(0.005)	(0.020)	(0.020)	(0.006)	(0.024)	(0.012)	(0.006)	(0.023)	(0.013)	(0.006)	(0.023)	(0.018)
Year 4	-0.003	-0.114***	0.053*	0.0008	-0.115***	0.018	0.001	-0.135***	0.014	-0.007	-0.077***	-0.021	-0.007	-0.076***	0.012
	(0.005)	(0.021)	(0.024)	(0.005)	(0.021)	(0.025)	(0.006)	(0.025)	(0.014)	(0.006)	(0.023)	(0.015)	(0.006)	(0.023)	(0.020)
Year 5	0.0003	-0.044	0.011	0.002	-0.050*	-0.077**	0.0002	-0.045	-0.003	-0.003	-0.058*	0.038	-0.003	-0.053*	0.039
	(0.006)	(0.023)	(0.025)	(0.006)	(0.023)	(0.024)	(0.006)	(0.026)	(0.015)	(0.006)	(0.023)	(0.021)	(0.006)	(0.023)	(0.024)
Male	0.002**	-0.001	-0.003	0.002***	-0.001	-0.003	0.003***	0.003	-0.001	0.002**	0.020***	0.0003	0.002**	0.020***	0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Age	0.007***	0.014***	0.008***	0.007***	0.014***	0.006***	0.010***	0.022***	0.002***	0.008***	0.014***	0.001	0.008***	0.014***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	-0.0002	0.004***	0.003***	-0.0002	0.004***	0.001*	0.00002	0.006***	0.001***	-0.0004	0.001	0.001*	-0.0004	0.001	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Ethnicity															
Base category - white															
Not-white	-0.0002	-0.022***	-0.008	-0.001	-0.021***	-0.008	0.0001	-0.019**	-0.002	-0.0002	-0.027***	-0.008**	-0.0002	-0.027***	-0.009
	(0.002)	(0.005)	(0.006)	(0.002)	(0.005)	(0.006)	(0.002)	(0.006)	(0.003)	(0.002)	(0.005)	(0.003)	(0.002)	(0.005)	(0.005)
Missing	0.001	-0.007	0.001	0.002	-0.002	-0.002	0.003	-0.001	0.008**	0.002	-0.010**	0.013***	0.002	-0.010**	0.009*
	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)	(0.004)	(0.004)
Observations								72652							
Pseudo R-squared		**0.0609**			**0.085**			**0.0587**			**0.0788**			**0.0862**	
Log likelihood		**-36076.918**			**-33463.315**			**-31387.521**			**-27786.174**			**-33714.822**	
Hausman test, number of outcomes failed		**1**			**2**			**1**			**1**			**1**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 23: Multinomial logit results for pneumonia process measures and patient characteristics

	Oxygenation assessment			Antibiotic selection			Blood cultures			Antibiotics received			Smoking cessation		
	Excluded died vs Pass	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**	**Excluded died vs Pass**	**Excluded alive vs Pass**	**Fail vs Pass**
Income	0.033	0.046	-0.001	0.005	-0.042	-0.020*	0.027	0.019	-0.005	0.028	0.084**	0.008	0.011	0.247***	-0.117***
	(0.021)	(0.024)	(0.003)	(0.023)	(0.029)	(0.009)	(0.023)	(0.029)	(0.010)	(0.022)	(0.028)	(0.016)	(0.024)	(0.028)	(0.010)
Interaction between income and year															
Base category – Year 1															
Year 2	-0.038	0.023	-0.003	0.003	0.022	0.031	-0.013	-0.070	0.034*	-0.031	-0.058	0.001	0.019	-0.003	-0.0005
	(0.035)	(0.041)	(0.005)	(0.038)	(0.048)	(0.016)	(0.039)	(0.049)	(0.017)	(0.037)	0.047)	(0.027)	(0.041)	(0.047)	(0.018)
Year 3	-0.031	-0.070*	-0.007	0.004	0.0004	0.047***	-0.003	-0.064	0.016	-0.020	-0.093*	-0.007	0.011	-0.051	0.034*
	(0.027)	(0.032)	(0.005)	(0.030)	(0.038)	(0.014)	(0.031)	(0.038)	(0.014)	(0.029)	(0.037)	(0.022)	(0.032)	(0.037)	(0.014)
Year 4	-0.054*	-0.078*	-0.005	-0.005	0.034	0.016	-0.042	-0.026	-0.001	-0.037	-0.116**	0.008	-0.031	0.023	0.031*
	(0.027)	(0.032)	(0.005)	(0.030)	(0.038)	(0.014)	(0.031)	(0.038)	(0.014)	(0.029)	(0.037)	(0.022)	(0.032)	(0.037)	(0.015)
Year 5	-0.046	-0.061*	0.002	-0.022	0.036	0.032*	-0.075*	-0.032	0.005	-0.035	-0.034	-0.014	-0.043	0.016	0.040*
	(0.027)	(0.031)	(0.007)	(0.029)	(0.036)	(0.013)	(0.037)	(0.045)	(0.020)	(0.028)	(0.036)	(0.022)	(0.031)	(0.036)	(0.016)
Male	0.006**	-0.014***	-0.000	0.013***	-0.014***	-0.002	0.012***	-0.031***	0.003*	0.008**	-0.024***	-0.006**	0.018***	-0.005	-0.002
	(0.002)	(0.003)	(0.000)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)
Age	0.054***	-0.027***	-0.000	0.069***	-0.034***	-0.007***	0.076***	-0.056***	-0.005***	0.063***	-0.044***	-0.002*	0.081***	0.000	-0.029***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	-0.003***	-0.002***	-0.00003	-0.002***	-0.001*	-0.0002	-0.002***	0.001	0.0002	-0.003***	-0.001**	0.002***	0.001	0.011***	-0.004***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Ethnicity															
Base category - white															
Not-white	-0.023***	-0.028***	0.002	-0.024***	-0.020*	0.006	-0.027***	-0.010	0.005	-0.027***	-0.014	0.009	-0.033***	0.089***	-0.020***
	(0.006)	(0.007)	(0.001)	(0.007)	(0.008)	(0.003)	(0.007)	(0.009)	(0.004)	(0.006)	(0.008)	(0.005)	(0.007)	(0.008)	(0.003)
Missing	0.022***	-0.000	-0.000	0.018**	-0.042***	-0.003	0.021***	-0.018*	-0.005	0.016**	-0.022**	0.009*	0.022***	-0.023***	-0.001
	(0.005)	(0.006)	(0.001)	(0.006)	(0.007)	(0.002)	(0.006)	(0.007)	(0.003)	(0.005)	(0.007)	(0.004)	(0.006)	(0.007)	(0.003)
Observations			**93502**					**78607**				**93502**			
Pseudo R-squared		**0.0397**			**0.0394**			**0.0484**			**0.0313**			**0.0884**	
Log likelihood		**-86932.353**			**-105233.17**			**-79903.534**			**-115742.36**			**-86613.382**	
Hausman test, number of outcomes failed		**1**			**0**			**0**			**0**			**0**	

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 24: Hausman and McFadden test of the IIA

		Excluded and died		Excluded and alive		Failed		Achieved		Overall
		Chi squared	Probability > Chi squared	Chi squared	Probability > Chi squared	Chi squared	Probability > Chi squared	Chi squared	Probability > Chi squared	Number of outcomes failed
Condition specific tests										
	AMI	-69257	>0.999	206	<0.001	-80001	>0.999	-79936	>0.999	**1**
	Heart failure	-6819	>0.999	3.27	>0.999	-9242	>0.999	-7275	>0.999	**0**
	Hip and knee	-1.98	>0.999	206	<0.001	-25.5	>0.999	151440	<0.001	**2**
	Pneumonia	-91014	>0.999	-934	>0.999	-88205	>0.999	-88065	>0.999	**0**
Process measure specific tests										
AMI	**Aspirin at arrival**	5.44	>0.999	-7.17	>0.999	-3.23	>0.999	9875	<0.001	**1**
	Fibrinolytic therapy	-8119	>0.999	-2634	>0.999	-7678	>0.999	-7776	>0.999	**0**
	Smoking cessation advice	-15046	>0.999	95.9	0.018	-13636	>0.999	-13824	>0.999	**1**
	Aspirin at discharge	38.2	0.999	-38.8	>0.999	-0.540	>0.999	7381	<0.001	**1**
	ACEI/ARB for LVSD	-3908	>0.999	110	0.001	-12420	>0.999	-12313	>0.999	**1**
	Beta blocker at discharge	-7.49	>0.999	-19.3	>0.999	-0.630	>0.999	5091	<0.001	**1**
Heart Failure	**Evaluation of LVS function**	-5.69	>0.999	9.55	>0.999	-7.06	>0.999	9090	<0.001	**1**
	ACEI/ARB for LVSD	-2807	>0.999	13.89	>0.999	-7959	>0.999	-8064	>0.999	**0**
	Discharge instructions	-4198	>0.999	66.59	0.001	-3.03	>0.999	-275	>0.999	**1**
	Smoking cessation advice	-10514	>0.999	-1629	>0.999	-10178	>0.999	-10257	>0.999	**0**
Hip and knee	**Antibiotics received**	-2.40	>0.999	7.46	>0.999	-7.40	>0.999	22322	<0.001	**1**
	Antibiotic selection	0.61	>0.999	240	<0.001	17.84	>0.999	20925	<0.001	**2**
	Antibiotics discontinued	-1.56	>0.999	-56.6	>0.999	-0.930	>0.999	24887	<0.001	**1**
	VTE prophylaxis ordered	-1.66	>0.999	61.2	0.766	-4.03	>0.999	3482	<0.001	**1**
	VTE prophylaxis received	-1.15	>0.999	32.9	>0.999	-7.96	>0.999	29747	<0.001	**1**
Pneumonia	**Oxygenation assessment**	-59.05	>0.999	-4.640	>0.999	-1.230	>0.999	18778	<0.001	**1**
	Antibiotic selection	-12092	>0.999	-60.85	>0.999	-12098	>0.999	-9917	>0.999	**0**
	Blood cultures	<0.001	>0.999	<0.001	>0.999	-109	>0.999	<0.001	>0.999	**0**
	Antibiotics received	-5231	>0.999	-85.37	>0.999	-4856	>0.999	-3685	>0.999	**0**
	Smoking cessation	-23405	>0.999	-17984	>0.999	-23138	>0.999	-24530	>0.999	**0**
Number of times failed		**0**		**6**		**0**		**11**		

Table 25: Sequential logit results for condition specific conditions

	AMI	Heart failure	Hip and knee replacement	Pneumonia
			Achieved vs not achieved	
Income	-0.041**	-0.102***	-0.032***	-0.053***
	(0.015)	(0.025)	(0.008)	(0.013)
Interaction between income and year				
Base category – year 1				
Year 2	0.002	0.079	0.073***	0.014
	(0.024)	(0.041)	(0.015)	(0.022)
Year 3	-0.030	0.055	0.089***	0.039*
	(0.019)	(0.032)	(0.012)	(0.017)
Year 4	-0.061**	0.119***	0.087***	0.030
	(0.019)	(0.032)	(0.013)	(0.017)
Year 5	-0.019	0.059	0.049***	0.011
	(0.018)	(0.032)	(0.013)	(0.016)
Male	0.004*	0.037***	-0.008***	0.012***
	(0.002)	(0.003)	(0.001)	(0.001)
Age	0.007***	-0.035***	-0.022***	-0.001*
	(0.001)	(0.001)	(0.000)	(0.001)
Age squared	0.003***	-0.007***	-0.005***	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)
Ethnicity				
Base category - white				
Not-white	0.007*	-0.033***	0.030***	0.011**
	(0.003)	(0.006)	(0.003)	(0.003)
Missing	-0.022***	-0.030***	-0.000	0.011***
	(0.003)	(0.006)	(0.002)	(0.003)
Foundation Trust	-0.028***	0.022***	-0.076***	-0.056***
	(0.003)	(0.005)	(0.002)	(0.003)
Care Quality Commission - quality				
Base category - excellent				
Good	-0.018***	0.026***	-0.040***	-0.042***
	(0.002)	(0.004)	(0.002)	(0.002)
Poor	0.067***	0.043***	0.027***	0.007*
	(0.003)	(0.005)	(0.002)	(0.003)
Care Quality Commission - financial				
Base category - excellent				
Good	-0.010***	-0.019***	-0.063***	-0.018***
	(0.002)	(0.004)	(0.001)	(0.002)
Poor	0.003	-0.019***	-0.072***	-0.039***
	(0.003)	(0.006)	(0.002)	(0.003)
Trust type				
Base category – small				
Medium	-0.062***	0.064***	-0.064***	-0.078***
	(0.003)	(0.005)	(0.002)	(0.003)
Large	-0.006	0.038***	-0.030***	-0.050***
	(0.003)	(0.005)	(0.002)	(0.003)
Specialist/Teaching	-0.010*	0.077***	-0.055***	-0.065***
	(0.004)	(0.007)	(0.003)	(0.003)
Mean Trust performance	0.252***	0.198***	0.433***	0.051*
	(0.025)	(0.043)	(0.017)	(0.022)
Observations	**365752**	**146256**	**363260**	**561012**
Log likelihood	**-313257.64**	**-181699.84**	**-167103.33**	**-636073.99**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 26: Sequential logit results for AMI process measures

	Aspirin at arrival	Fibrinolytic therapy	Smoking cessation advice	Aspirin at discharge	ACEI/ARB for LVSD	Beta blocker at discharge
			Achieved vs not achieved			
Income	-0.003	-0.006	-0.266***	-0.063	-0.091**	0.003
	(0.037)	(0.011)	(0.030)	(0.043)	(0.030)	(0.044)
Interaction between income and year Base category – year 1						
Year 2	-0.005	0.003	-0.061	0.064	0.127**	0.032
	(0.062)	(0.019)	(0.050)	(0.069)	(0.049)	(0.071)
Year 3	-0.050	0.000	-0.005	-0.055	0.050	-0.041
	(0.050)	(0.016)	(0.039)	(0.056)	(0.040)	(0.058)
Year 4	-0.087	0.031	-0.045	-0.056	-0.018	-0.129*
	(0.051)	(0.025)	(0.040)	(0.056)	(0.040)	(0.059)
Year 5	-0.005	0.049	-0.037	0.009	0.034	0.008
	(0.045)	(0.032)	(0.037)	(0.053)	(0.037)	(0.055)
Male	0.010*	0.007***	-0.015***	-0.009	0.029***	-0.008
	(0.004)	(0.002)	(0.004)	(0.005)	(0.003)	(0.005)
Age	0.031***	-0.013***	-0.100***	0.051***	0.027***	0.038***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Age squared	0.004***	-0.002***	-0.014***	0.012***	-0.002**	0.010***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ethnicity Base category - white						
Not-white	0.031***	-0.002	-0.044***	0.034***	-0.001	0.055***
	(0.008)	(0.004)	(0.006)	(0.010)	(0.007)	(0.010)
Missing	-0.058***	0.003	-0.023***	-0.070***	0.020**	-0.051***
	(0.008)	(0.003)	(0.006)	(0.008)	(0.007)	(0.009)
Observations			45719			
Log likelihood	**-33283.383**	**-18422.126**	**-31648.385**	**-39026.215**	**-28622.455**	**-42079.149**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Results in light grey are process measures which are given during hospital stay. Results in dark grey are process measures given on discharge. Coefficients in the table of 0.01 signify 1%.

Table 27: Sequential logit results for heart failure process measures

	Evaluation of LVS function	ACEI/ARB for LVSD	Discharge instructions	Smoking cessation advice
		Achieved vs not achieved		
Income	-0.043	-0.105*	-0.201***	-0.106***
	(0.041)	(0.046)	(0.052)	(0.024)
Interaction between income and year Base category – year 1				
Year 2	0.004	0.176*	0.165*	0.029
	(0.074)	(0.078)	(0.084)	(0.039)
Year 3	0.010	0.029	0.247***	0.008
	(0.056)	(0.061)	(0.065)	(0.031)
Year 4	0.062	0.162**	0.315***	0.026
	(0.054)	(0.062)	(0.065)	(0.030)
Year 5	0.056	0.099	0.170**	-0.000
	(0.053)	(0.061)	(0.065)	(0.030)
Male	0.004	0.094***	0.047***	0.004
	(0.005)	(0.005)	(0.006)	(0.003)
Age	-0.010***	-0.054***	-0.029***	-0.048***
	(0.002)	(0.002)	(0.002)	(0.002)
Age squared	-0.009***	-0.007***	-0.010***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Ethnicity Base category - white				
Not-white	0.013	-0.034**	-0.066***	-0.030***
	(0.010)	(0.011)	(0.011)	(0.004)
Missing	-0.035**	-0.030*	-0.079***	0.011
	(0.011)	(0.012)	(0.013)	(0.007)
Observations		36564		
Log likelihood	**-34537.573**	**-37432.518**	**44628.57**	**-25948.78**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 28: Sequential logit results for hip and knee process measures

	Antibiotics received	Antibiotic selection	Antibiotics discontinued	VTE prophylaxis ordered	VTE prophylaxis received
			Achieved vs not achieved		
Income	-0.013	-0.061***	-0.022	-0.050**	-0.017
	(0.018)	(0.018)	(0.018)	(0.016)	(0.018)
Interaction between income and year Base category – year 1					
Year 2	0.110**	0.121***	0.076*	0.002	0.022
	(0.034)	(0.034)	(0.034)	(0.032)	(0.034)
Year 3	0.043	0.132***	0.137***	0.086***	0.041
	(0.028)	(0.027)	(0.026)	(0.025)	(0.028)
Year 4	0.061*	0.098**	0.121***	0.100***	0.067*
	(0.030)	(0.031)	(0.028)	(0.026)	(0.029)
Year 5	0.033	0.127***	0.048	0.020	0.016
	(0.031)	(0.031)	(0.029)	(0.030)	(0.031)
Male	0.004	0.004	-0.003	-0.021***	-0.022***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Age	-0.023***	-0.021***	-0.027***	-0.016***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	-0.007***	-0.005***	-0.008***	-0.001*	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ethnicity Base category - white					
Not-white	0.029***	0.029***	0.021**	0.035***	0.036***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Missing	0.005	0.003	-0.008	-0.004	0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations			**72652**		
Log likelihood	**-36076.842**	**-33462.607**	**-31388.262**	**-27786.313**	**-33715.65**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Coefficients in the table of 0.01 signify 1%.

Table 29: Sequential logit results for pneumonia process measures

	Oxygenation assessment	Antibiotic selection	Blood cultures	Antibiotics received	Smoking cessation
			Achieved vs not achieved		
Income	-0.066*	0.070*	-0.044	-0.132***	-0.170***
	(0.028)	(0.033)	(-1.45)	(0.033)	(0.021)
Interaction between income and year Base category – year 1					
Year 2	0.047	-0.067	0.058	0.090	-0.021
	(0.047)	(0.053)	(1.22)	(0.053)	(0.031)
Year 3	0.105**	-0.058	0.061	0.136**	0.006
	(0.037)	(0.042)	(1.68)	(0.042)	(0.025)
Year 4	0.118**	-0.054	0.076*	0.158***	-0.027
	(0.037)	(0.041)	(2.10)	(0.042)	(0.025)
Year 5	0.087*	-0.064	0.110**	0.084*	-0.018
	(0.036)	(0.040)	(2.65)	(0.040)	(0.024)
Male	0.015***	0.009*	0.023***	0.031***	-0.011***
	(0.003)	(0.004)	(7.31)	(0.004)	(0.002)
Age	0.013***	-0.002	-0.001	0.009***	-0.056***
	(0.001)	(0.001)	(-1.10)	(0.001)	(0.001)
Age squared	0.004***	0.003***	0.001*	0.002***	-0.010***
	(0.000)	(0.000)	(2.15)	(0.000)	(0.000)
Ethnicity Base category - white					
Not-white	0.039***	0.036***	0.032***	0.026**	-0.045***
	(0.008)	(0.009)	(4.07)	(0.009)	(0.003)
Missing	-0.006	0.043***	0.006	0.004	0.002
	(0.007)	(0.008)	(0.90)	(0.008)	(0.005)
Observations	**93502**		**78607**	**93502**	
Log likelihood	**-86956.51**	**-105239.06**	**-79903.346**	**-115740.49**	**-86569.789**

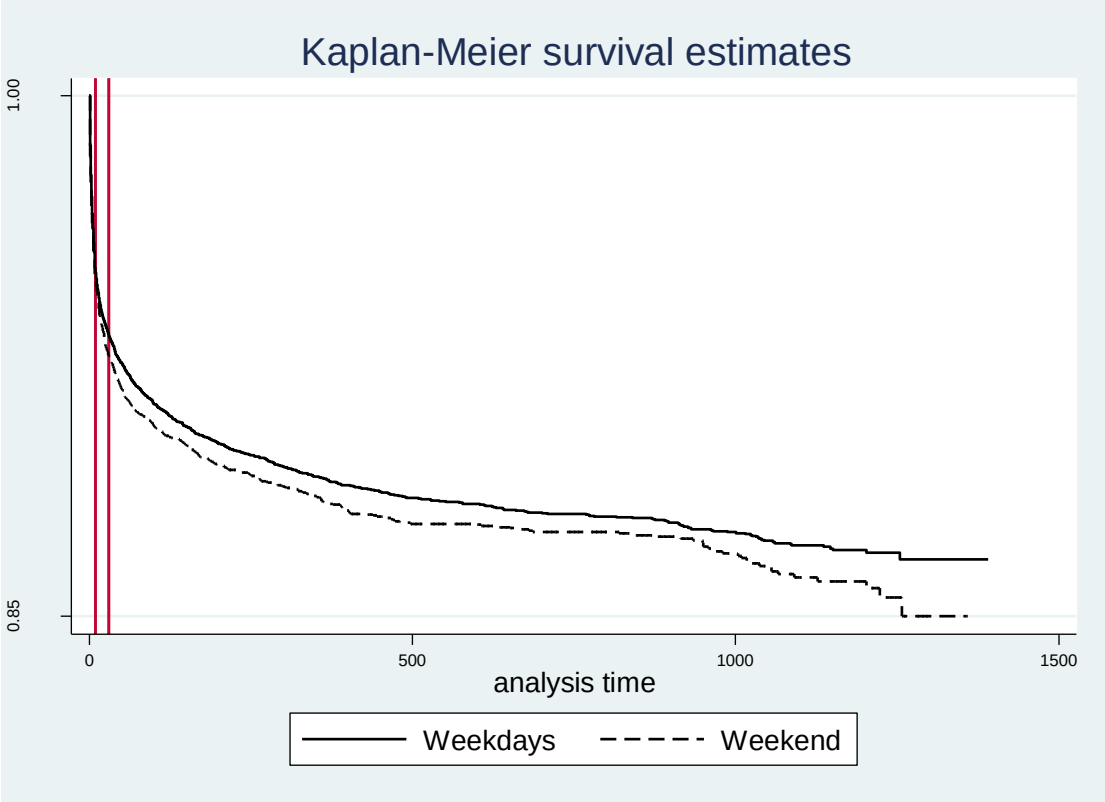
Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Models also include quarter dummies and Trust variables such as: foundation Trust status; CQC quality; CQC financial score; Trust type and Trust mean performance. Results in light grey are process measures which are given during hospital stay. Coefficients in the table of 0.01 signify 1%.

Table 30: Summary of effects of income score area on process measures of care

				Sequential logit	Multinomial Logit			Hausman Test
			Year	Probability of pass	Excluded died vs Pass	Excluded alive vs Pass	Fail vs Pass	Number of failed outcomes
		AMI	one	-0.041**	0.038	-0.069**	0.006**	1
			five	-0.019	0.005	-0.022	0.0002	
		Heart failure	one	-0.102***	0.034	-0.098***	-0.009	0
			five	0.059	0.091*	-0.002	-0.0.03	
		Pneumonia	one	-0.053***	-0.023	-0.054**	0.027***	2
			five	0.011	0.040	0.004	-0.023**	
		Hip and knee	one	-0.032***	0.007*	-0.042***	0.010	0
			five	0.049***	0.001	0.049**	-0.001	
	Process measures under emergency	Aspirin at arrival	one	-0.003	0.027	-0.015	0.006	1
		(AMI)	five	-0.005	0.023	-0.005	-0.005	
		Fibrinolytic therapy	one	-0.006	0.038	-0.047	0.004	0
		(AMI)	five	0.049	0.007	0.066	-0.025	
		Oxygenation assessment	one	-0.066*	-0.033	-0.046	0.001	1
		(Pneumonia)	five	0.087*	0.046	0.061*	-0.002	
		Blood cultures	one	-0.044	-0.027	-0.019	0.005	0
		(Pneumonia)	five	0.110**	0.075*	0.032	-0.005	
		Antibiotics received	one	-0.132***	-0.028	-0.084**	-0.008	0
		(Pneumonia)	five	0.084*	0.035	0.034	0.014	
		Antibiotic selection	one	0.070*	-0.005	0.042	0.020*	0
		(Pneumonia)	five	-0.064	0.022	-0.036	-0.032*	
Emergency surgery	Advice	Smoking cessation advice	one	-0.266***	0.044	-0.300***	0.012	1
		(AMI)	five	-0.037	0.003	-0.036	-0.003	
		Smoking cessation advice	one	-0.106***	0.039	-0.166***	0.034***	0
		(Heart failure)	five	-0.000	0.089*	-0.076	-0.006	
		Smoking cessation advice	one	-0.170***	-0.011	-0.247***	0.117***	0
		(Pneumonia)	five	-0.018	0.043	-0.016	-0.040*	
		Discharge instructions	one	-0.201***	0.012	-0.039	-0.138***	1
		(Heart failure)	five	0.170**	0.118	0.008	0.070	
	Drug and tests	ACEI/ARB for LVSD	one	-0.091**	0.041	-0.040	0.025**	1
		(AMI)	five	0.034	-0.001	-0.005	0.011	
		ACEI/ARB for LVSD	one	-0.105*	0.035	-0.102*	-0.007	0
		(Heart failure)	five	0.099	0.091	0.021	0.002	
		Aspirin at discharge	one	-0.063	0.042	-0.080*	0.009*	1
		(AMI)	five	0.009	0.0004	0.020	-0.015	
		Beta blocker at discharge	one	0.003	0.043	-0.120***	0.003	1
		(AMI)	five	0.008	-0.002	0.027	0.002	
		Evaluation of LVS function	one	-0.043	0.039	-0.069*	0.024	1
		(Heart failure)	five	0.056	-0.087*	0.044	-0.014	
Elective surgery	Antibiotics	Antibiotics received	one	-0.013	0.007	-0.025	0.011	1
		(Hip and knee)	five	0.033	-0.0003	0.044	-0.011	
		Antibiotics discontinued	one	-0.022	0.014**	-0.031	0.007	2
		(Hip and knee)	five	0.048	-0.0002	0.045	0.003	
		VTE prophylaxis received	one	-0.017	0.003	-0.057***	0.040***	1
		(Hip and knee)	five	0.016	0.003	0.053*	-0.039	
	Ordering	Antibiotic selection	one	-0.061***	0.008*	-0.037**	-0.025	1
		(Hip and knee)	five	0.127***	-0.002	0.050*	0.077**	
		VTE prophylaxis ordered	one	-0.050**	0.003	-0.058***	0.009	1
		(Hip and knee)	five	0.020	0.003	0.058*	-0.038	

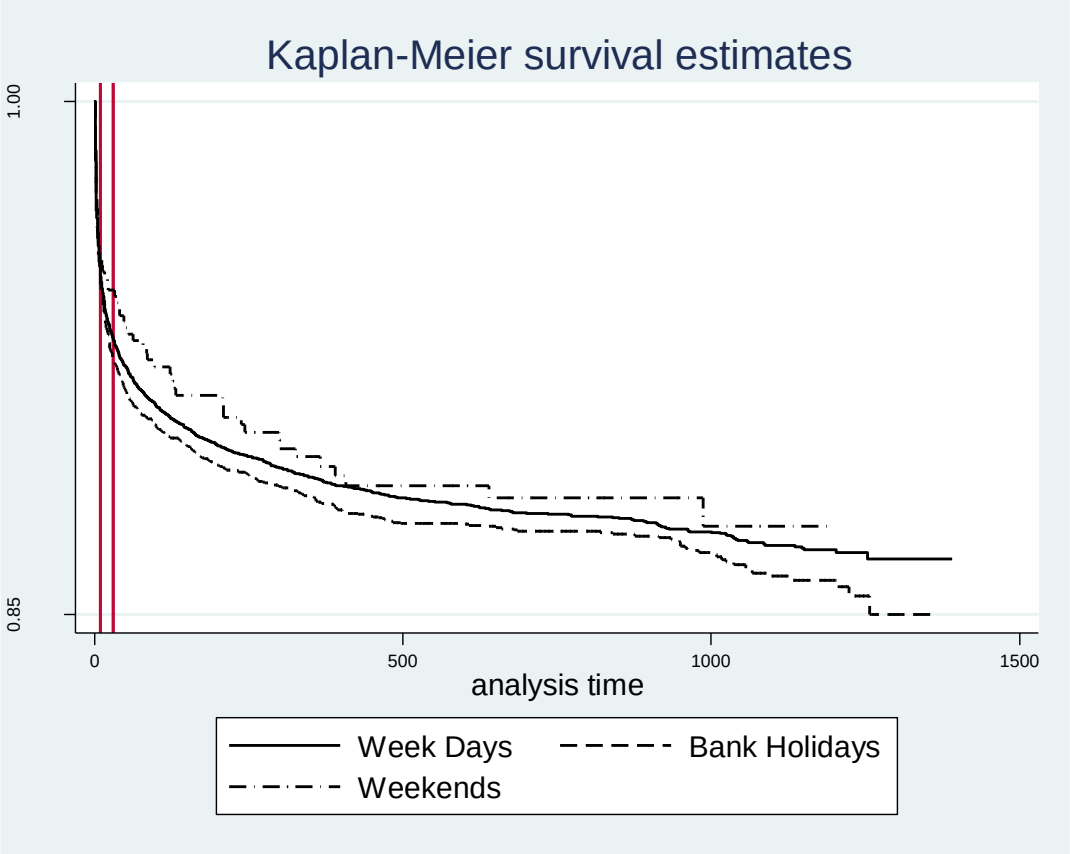
Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects displayed for all regression results. Results shown are from the previous tables. Coefficients in the table of 0.01 signify 1%.

Figure 8: AMI survival curves – Weekdays and weekends



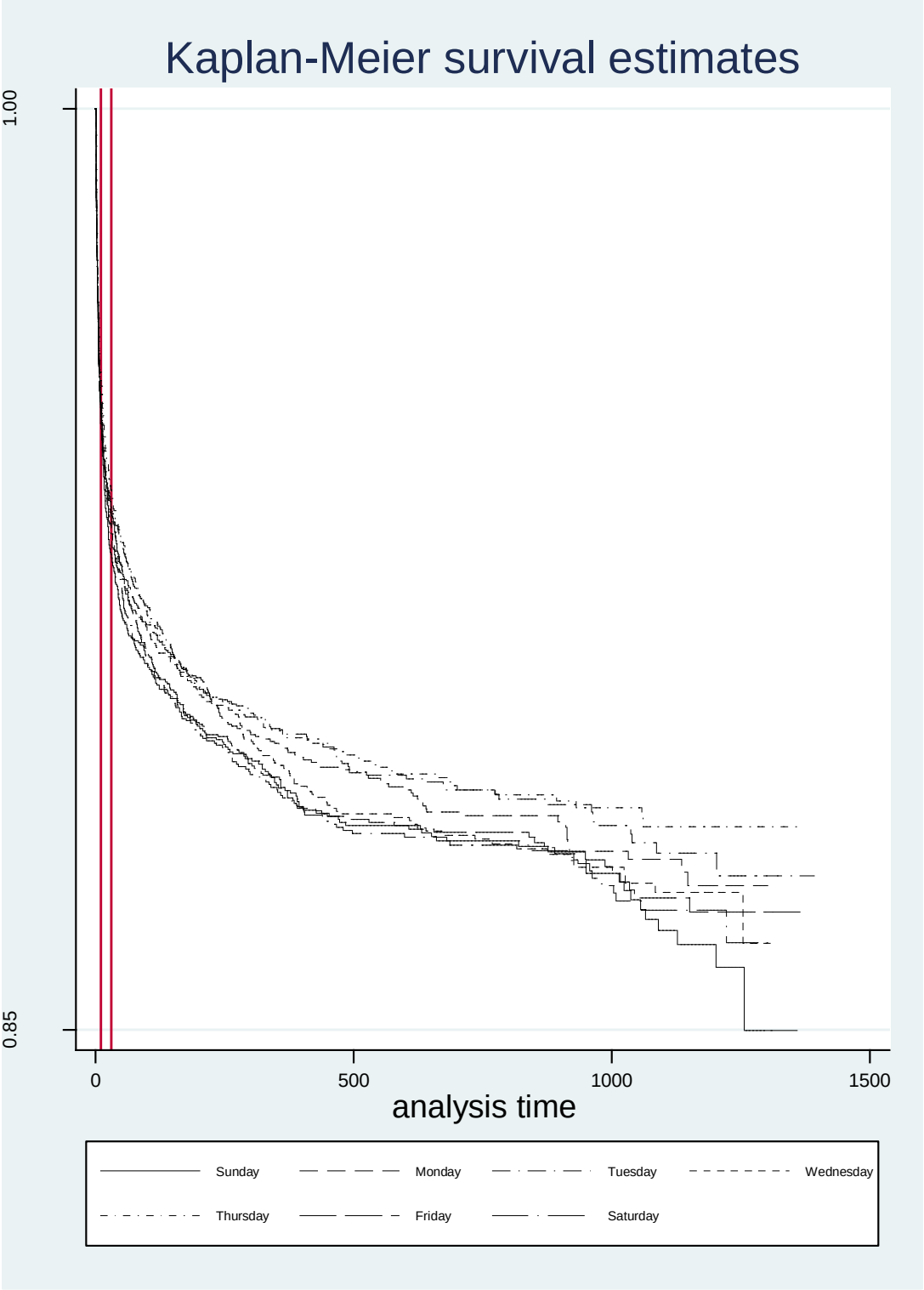
The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Figure 9: AMI survival curves – Weekdays, weekends and bank holidays



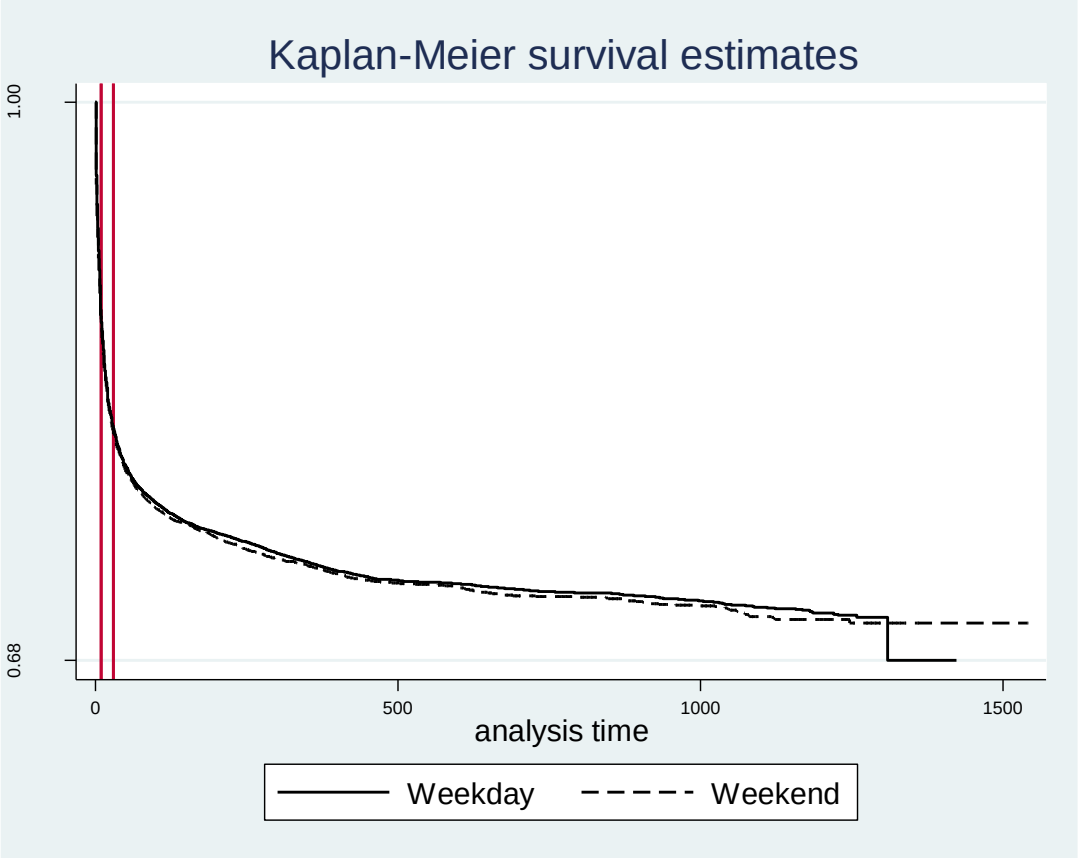
The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Figure 10: AMI survival curves – Days of the week



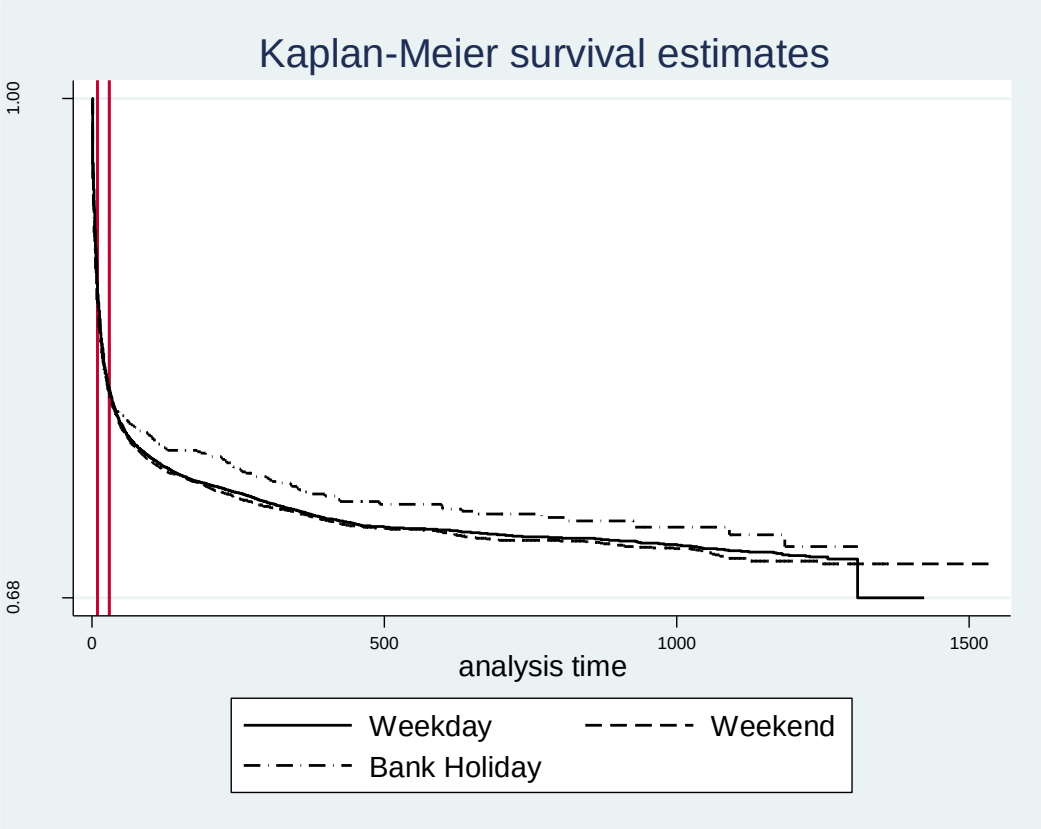
The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Figure 11: Pneumonia survival curves – Weekday and weekend



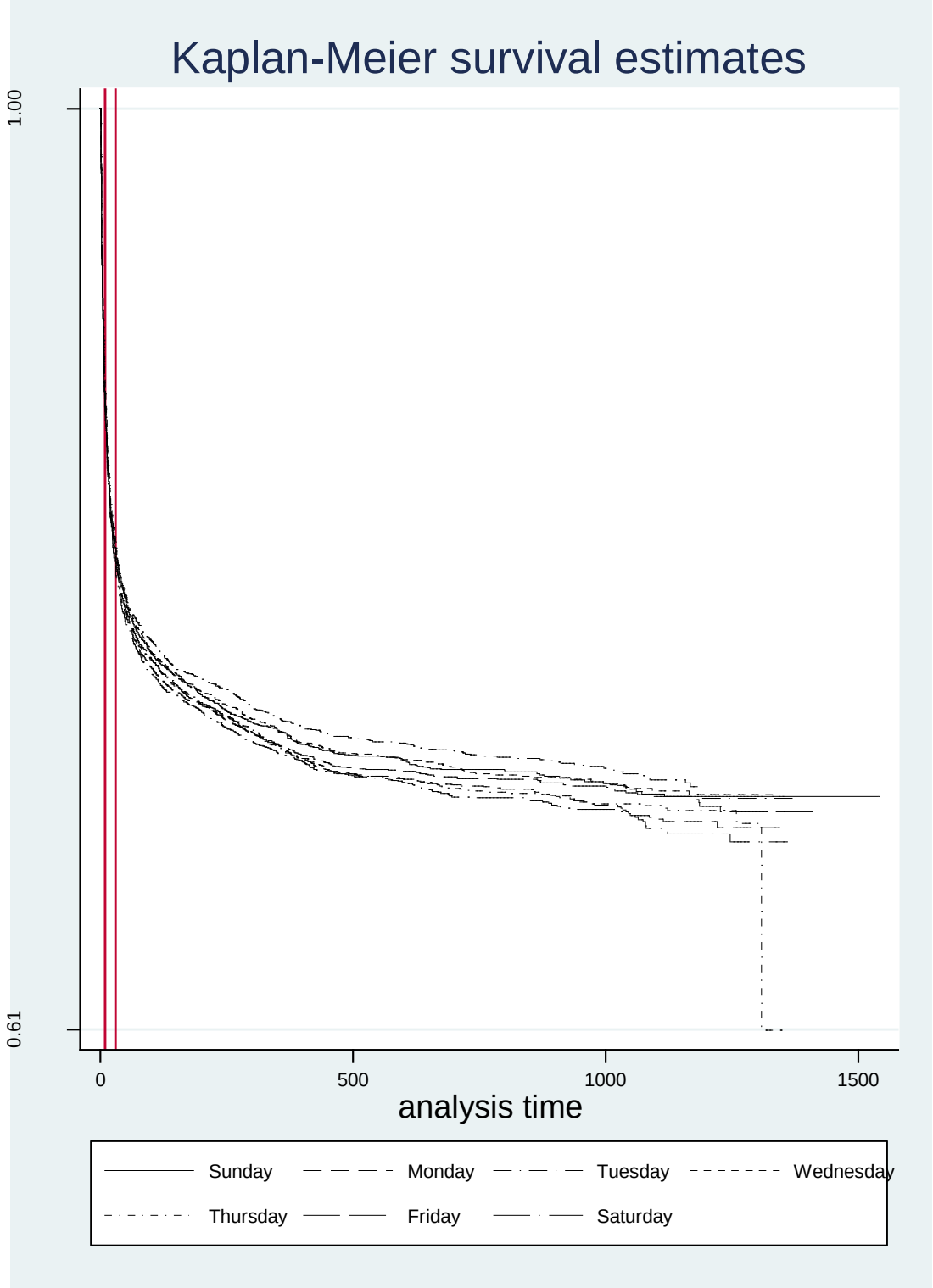
The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Figure 12: Pneumonia survival curves – Weekday, weekend and bank holiday



The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Figure 13: Pneumonia survival curves – Days of the week



The two vertical lines represent 10 and 30 days after admission respectively. Y axis is the survival proportion. Patients who died out of hospital were assumed alive as I did not observe out of hospital deaths. Volatility above 900 days increased as fewer patients were observed as analysis time increases.

Table 31: Descriptive statistics – sample means for AMI

	Admission Day of the Week						
	Mon	**Tue**	**Wed**	**Thur**	**Fri**	**Sat**	**Sun**
Process measures of care							
Aspirin Arrival	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Aspirin Discharge	0.99	0.99	0.99	0.999	0.999	0.99	0.99
ACEI for LVSD	0.99	0.98	0.98	0.98	0.99	0.99	0.98
Smoking Cessation	0.94	0.95	0.93	0.93	0.93	0.95	0.94
Beta-Blocker Discharge	0.98	0.98	0.98	0.97	0.98	0.98	0.98
Fibrinolytic Therapy	0.84	0.86	0.84	0.82	0.86	0.87	0.84
Primary PCI	0.95	0.96	0.97	0.97	0.95	0.96	0.94
Patient mortality							
Died	0.08	0.08	0.08	0.08	0.08	0.09	0.09
Patient characteristics							
Male	0.63	0.62	0.62	0.61	0.62	0.61	0.63
Age in years	69.14	69.57	69.48	69.6	69.75	70.19	70.05
Ethnicity							
White	0.87	0.87	0.86	0.87	0.86	0.86	0.87
Mixed	0.01	0.01	0.01	0.01	0.01	0.02	0.01
Asian	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Black	0.01	<0.01	0.01	0.01	0.01	<0.01	<0.01
Other	0.01	0.01	0.01	0.01	0.01	0.02	0.01
Missing	0.07	0.08	0.08	0.08	0.08	0.07	0.07
Distance to hospital bands							
0 - 2 miles	0.2	0.19	0.21	0.21	0.21	0.21	0.2
2 - 4 miles	0.26	0.27	0.26	0.26	0.26	0.27	0.26
4 - 6 miles	0.15	0.15	0.15	0.14	0.15	0.13	0.15
6 - 10 miles	0.14	0.14	0.14	0.14	0.14	0.14	0.15
10 - 14 miles	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Over 14 miles	0.12	0.12	0.12	0.13	0.12	0.12	0.13
Missing LSOA data	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Elixhauser Comorbidities (see footnotes)							
(1)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(2)	0.08	0.08	0.08	0.08	0.07	0.09	0.09
(3)	0.05	0.06	0.06	0.06	0.06	0.06	0.06
(4)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
(5)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(6)	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(7)	0.07	0.07	0.07	0.07	0.06	0.07	0.06
(8)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(9)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(10)	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(11)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
(12)	0.01	0.01	0.01	0.01	0.01	0.01	0.01
(13)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(14)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(15)	0.01	0.01	<0.01	<0.01	0.01	0.01	0.01
(16)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

(17)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(18)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(19)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(20)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(21)	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	0.01
(22)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(23)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(24)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(25)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(26)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(27)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(28)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(29)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(30)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(31)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(32)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
ICD 10 sub chapters							
I210	0.09	0.09	0.09	0.1	0.09	0.1	0.1
I211	0.12	0.12	0.12	0.11	0.12	0.13	0.13
I212	0.01	0.01	0.01	0.02	0.02	0.02	0.02
I214	0.04	0.04	0.04	0.04	0.05	0.04	0.04
I219	0.48	0.48	0.47	0.46	0.47	0.46	0.46
I220	0.01	0.01	0.01	0.01	0.01	0.01	0.01
I221	0.01	0.01	0.01	0.01	0.01	0.01	0.01
I228	0.01	0.01	0.01	0.01	0.01	0.01	0.01
I229	0.12	0.13	0.13	0.13	0.12	0.14	0.13
I251	0.05	0.05	0.05	0.07	0.06	0.05	0.05
Trust characteristics							
Small	0.11	0.11	0.1	0.11	0.1	0.1	0.11
Medium	0.32	0.32	0.32	0.32	0.33	0.32	0.31
Large	0.34	0.33	0.33	0.33	0.34	0.35	0.33
Specialist / Teaching	0.24	0.24	0.25	0.25	0.23	0.23	0.25
FT	0.61	0.62	0.61	0.62	0.61	0.61	0.61
Day volume of admissions (00's)	35.16	33.34	33.69	34.48	34.79	29.01	29.4
CQC quality score							
Excellent	0.31	0.3	0.3	0.3	0.31	0.31	0.31
Good	0.48	0.48	0.48	0.47	0.47	0.47	0.47
Fair/Weak	0.22	0.21	0.22	0.23	0.22	0.21	0.22
Observations	**6,134**	**5,861**	**5,795**	**5,894**	**5,988**	**5,090**	**5,155**
Bed utilisation (%)	0.84	0.84	0.84	0.83	0.79	0.78	0.8
Emergency volume (0000's)	0.25	0.24	0.23	0.23	0.24	0.19	0.19
Distance to hospital (km)	11.41	11.56	11.34	11.22	11.18	11.41	11.06
Observations	**3,455**	**4,061**	**3,806**	**3,680**	**3,738**	**3,819**	**3,429**

Process measures in dark grey box are discharge measures, whereas other process measures are for admission. Values in the table of 0.01 signify 1%. Achievement proportions displayed are based on non-excluded patients. Elixhauser Comorbidities: (1) Angina; (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (17) Peptic ulcer disease; (18) AIDS/HIV; (19) Lymphoma; (20) Metastatic cancer; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (27) Blood loss anemia; (28) Deficiency anemia; (29) Alcohol abuse; (30) Drug abuse; (31) Psychoses and (32) Depression.

Table 32: Descriptive statistics – sample means for pneumonia

	Arrival Day of the Week						
	Mon	**Tue**	**Wed**	**Thur**	**Fri**	**Sat**	**Sun**
Process measures of care							
Oxygenation assessment	0.999	0.999	0.99	0.999	0.99	0.99	0.99
Antibiotic Selection	0.9	0.9	0.9	0.91	0.89	0.9	0.9
Blood Cultures	0.82	0.82	0.84	0.82	0.81	0.8	0.81
Antibiotics Received	0.75	0.76	0.77	0.76	0.76	0.78	0.79
Smoking cessation	0.65	0.63	0.65	0.64	0.62	0.62	0.65
Patient mortality							
Died	0.23	0.23	0.23	0.23	0.23	0.24	0.23
Patient characteristics							
Male	0.5	0.5	0.51	0.49	0.5	0.5	0.49
Age in years	72.38	72.72	72.47	72.56	72.53	73.32	73.11
Ethnicity							
White	0.91	0.91	0.9	0.9	0.91	0.91	0.91
Mixed	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Asian	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Black	<0.01	<0.01	0.01	<0.01	0.01	<0.01	0.01
Other	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Missing	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Distance to hospital bands							
0 - 2 miles	0.26	0.26	0.25	0.25	0.26	0.26	0.26
2 - 4 miles	0.32	0.32	0.32	0.32	0.32	0.31	0.32
4 - 6 miles	0.15	0.15	0.16	0.16	0.15	0.15	0.15
6 - 10 miles	0.12	0.12	0.12	0.12	0.11	0.12	0.12
10 - 14 miles	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Over 14 miles	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Missing LSOA data	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Elixhauser Comorbidities (see footnotes)							
(1)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(2)	0.03	0.03	0.03	0.03	0.03	0.03	0.03
(3)	0.04	0.03	0.04	0.03	0.04	0.04	0.04
(4)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(5)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(6)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(7)	0.02	0.02	0.02	0.02	0.03	0.02	0.02
(8)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(9)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(10)	0.01	0.01	0.01	0.01	0.01	0.02	0.01
(11)	0.31	0.31	0.31	0.3	0.3	0.3	0.31
(12)	0.01	0.01	0.01	0.01	0.01	0.02	0.01
(13)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(14)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(15)	0.01	0.01	0.01	0.01	0.01	0.01	0.01
(16)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(17)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(18)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(19)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(20)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(21)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
(22)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(23)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

(24)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(25)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(26)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
(27)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(28)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(29)	<0.01	<0.01	0.01	0.01	0.01	0.01	<0.01
(30)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(31)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
(32)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
ICD 10 main chapters							
J13	0.02	0.02	0.01	0.02	0.02	0.01	0.02
J14	0.01	0.01	0.01	0.01	0.01	<0.01	<0.01
J15	0.02	0.02	0.02	0.02	0.02	0.02	0.02
J18	0.9	0.9	0.9	0.9	0.9	0.91	0.91
Trust characteristics							
Small	0.11	0.1	0.1	0.1	0.11	0.11	0.11
Medium	0.34	0.36	0.34	0.35	0.35	0.34	0.34
Large	0.37	0.36	0.37	0.36	0.36	0.38	0.37
Specialist/Teaching	0.18	0.18	0.19	0.19	0.18	0.18	0.18
FT	0.51	0.53	0.52	0.51	0.52	0.52	0.52
Day volume of admissions (00's)	0.77	0.73	0.71	0.73	0.74	0.65	0.67
CQC quality score							
Excellent	0.31	0.31	0.3	0.31	0.3	0.3	0.31
Good	0.58	0.57	0.57	0.58	0.58	0.58	0.58
Fair/Weak	0.11	0.12	0.13	0.12	0.12	0.12	0.11
Observations	**13,225**	**12,658**	**12,088**	**12,431**	**12,781**	**11,147**	**11,325**
Bed utilisation (%)	0.83	0.84	0.83	0.82	0.79	0.78	0.8
Emergency volume (0000's)	0.25	0.24	0.23	0.23	0.24	0.19	0.19
Distance to hospital (km)	7.87	7.99	8	7.82	7.96	8.01	7.9
Observations	**7,527**	**8,866**	**8,657**	**8,083**	**8,370**	**8,713**	**7,497**

Process measures in dark grey box are discharge measures, whereas other process measures are for admission. The proportions of utilisation and emergency volumes in tables 31 and 32 are the same variables. I find slightly different proportions as the population of patients are different; this means that the proportions are a different weighted average across the days of the week leading to a small change in proportions across the two tables. Achievement proportions displayed do not account for patients who are excluded. Elixhauser Comorbidities: (1) Angina; (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (17) Peptic ulcer disease; (18) AIDS/HIV; (19) Lymphoma; (20) Metastatic cancer; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (27) Blood loss anemia; (28) Deficiency anemia; (29) Alcohol abuse; (30) Drug abuse; (31) Psychoses; and (32) Depression. Values in the table of 0.01 signify 1%. Coefficients in the table of 0.01 signify 1%.

Table 33: Logistic regression results: AMI – univariate model

	Aspirin Arrival	ACEI for LVSD	Fibrinolytic Therapy	Primary PCI	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.003	0.009	0.002	-0.016	-0.002	0.013	-0.000
	(0.002)	(0.007)	(0.040)	(0.013)	(0.002)	(0.011)	(0.004)
Tuesday	-0.002	0.005	0.022	-0.005	-0.003	0.020	-0.001
	(0.002)	(0.008)	(0.040)	(0.013)	(0.002)	(0.010)	(0.004)
Base category – Wednesday							
Thursday	0.002	0.008	-0.020	-0.003	0.001	0.001	-0.001
	(0.002)	(0.007)	(0.041)	(0.012)	(0.002)	(0.011)	(0.004)
Friday	0.001	0.012	0.022	-0.018	0.002	0.004	0.002
	(0.002)	(0.007)	(0.040)	(0.014)	(0.002)	(0.011)	(0.004)
Saturday	0.000	0.010	0.024	-0.012	-0.002	0.021	0.002
	(0.002)	(0.007)	(0.040)	(0.014)	(0.002)	(0.011)	(0.004)
Sunday	-0.001	0.006	-0.005	-0.033*	0.000	0.013	0.002
	(0.003)	(0.007)	(0.041)	(0.015)	(0.002)	(0.011)	(0.004)
Weekend	-0.000	0.002	0.004	-0.014	-0.000	0.009	0.002
	(0.001)	(0.004)	(0.023)	(0.009)	(0.001)	(0.006)	(0.002)
Weekend - BH	-0.001	0.003	0.002	-0.016	-0.000	0.009	0.002
	(0.001)	(0.004)	(0.023)	(0.009)	(0.001)	(0.006)	(0.002)
Bank holiday	-0.008	-	-0.015	-0.022	-0.001	0.003	0.001
	(0.005)	-	(0.063)	(0.032)	(0.004)	(0.018)	(0.007)
Observations	**28150**	**4844**	**1151**	**2927**	**23400**	**7137**	**21714**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects are displayed for logit regressions using margins estimated at means. Process measures in dark grey box are discharge measures, whereas other process measures are for admission. Coefficients in the table of 0.01 signify 1%.

Table 34: Logistic regression results: pneumonia – univariate model

	Oxygenation Assessment	Antibiotic Selection For CAP	Blood Cultures	Initial Antibiotics	Smoking Cessation
Monday	0.001	-0.000	-0.028*	0.019*	0.003
	(0.001)	(0.006)	(0.012)	(0.008)	(0.016)
Tuesday	0.002	0.002	-0.020	-0.017*	-0.000
	(0.001)	(0.006)	(0.012)	(0.008)	(0.016)
Base category – Wednesday					
Thursday	0.002	-0.002	-0.020	-0.013	-0.016
	(0.001)	(0.006)	(0.012)	(0.008)	(0.016)
Friday	0.002	0.006	-0.016	-0.008	-0.006
	(0.001)	(0.006)	(0.012)	(0.008)	(0.016)
Saturday	0.001	-0.012	-0.027*	-0.011	-0.029
	(0.001)	(0.006)	(0.012)	(0.008)	(0.016)
Sunday	0.002	-0.006	-0.037**	0.009	-0.030
	(0.001)	(0.007)	(0.012)	(0.008)	(0.017)
Weekend	0.000	-0.001	-0.012	0.024***	-0.006
	(0.001)	(0.004)	(0.007)	(0.005)	(0.009)
Weekend - no BH	0.000	-0.002	-0.015*	0.024***	-0.003
	(0.001)	(0.004)	(0.007)	(0.005)	(0.010)
Bank holiday	0.001	0.005	0.019	0.009	-0.023
	(0.002)	(0.010)	(0.018)	(0.013)	(0.026)
Observations	**52767**	**30847**	**14917**	**37864**	**12548**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Marginal effects are displayed for logit regressions using margins estimated at means. Process measures in dark grey box are discharge measures, whereas other process measures are for admission. Coefficients in the table of 0.01 signify 1%.

Table 35: Logistic regression results: AMI – Full model with all covariates

	Aspirin Arrival	Fibrinolytic Therapy	Primary PCI	ACEI for LVSD	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.003	-0.017	-0.014	0.011	-0.002	0.012	-0.001
	(0.003)	(0.039)	(0.015)	(0.008)	(0.002)	(0.010)	(0.004)
Tuesday	-0.002	-0.000	-0.004	0.006	-0.004	0.024*	-0.000
	(0.003)	(0.038)	(0.015)	(0.009)	(0.002)	(0.010)	(0.004)
Base category – Wednesday							
Thursday	0.002	-0.046	0.002	0.009	0.001	0.003	-0.002
	(0.002)	(0.040)	(0.013)	(0.008)	(0.002)	(0.011)	(0.004)
Friday	0.001	-0.002	-0.013	0.015	0.002	0.007	0.002
	(0.002)	(0.039)	(0.015)	(0.008)	(0.002)	(0.011)	(0.004)
Saturday	0.000	0.014	-0.011	0.012	-0.002	0.023*	0.003
	(0.002)	(0.038)	(0.015)	(0.008)	(0.002)	(0.011)	(0.004)
Sunday	-0.001	-0.027	-0.030	0.005	0.001	0.011	0.001
	(0.003)	(0.040)	(0.016)	(0.009)	(0.002)	(0.011)	(0.004)
Trust characteristics							
Small	-0.006*	-0.079	-	-0.003	-0.004	-0.019	-0.012**
	(0.002)	(0.057)	-	(0.008)	(0.002)	(0.011)	(0.004)
Medium	-0.003	-0.021	-0.002	0.004	-0.005**	-0.006	0.002
	(0.002)	(0.028)	(0.071)	(0.006)	(0.002)	(0.008)	(0.003)
Base category – large							
Specialist/Teaching	-0.008**	-0.160	0.007	0.003	0.001	-0.010	0.010**
	(0.003)	(0.129)	(0.014)	(0.007)	(0.001)	(0.009)	(0.003)
Foundation Trust	-0.004**	0.045	0.226	-0.001	0.000	0.003	-0.008***
	(0.001)	(0.028)	(0.192)	(0.005)	(0.001)	(0.007)	(0.002)
Day volume of admissions (00's)	0.000	0.002	0.001	-0.000	0.000	0.000	0.000
	(0.000)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
CQC quality							
Base category – excellent							
Good	0.002	0.034	-	0.009	-0.000	-0.010	0.014***
	(0.002)	(0.030)	-	(0.007)	(0.002)	(0.008)	(0.004)
Fair/poor	0.011***	-0.377*	-0.009	0.016*	0.005***	0.008	0.024***
	(0.002)	(0.147)	(0.011)	(0.007)	(0.001)	(0.009)	(0.004)
Male	0.007***	0.051	0.010	0.007	0.001	-0.002	0.003
	(0.001)	(0.026)	(0.010)	(0.005)	(0.001)	(0.006)	(0.002)
Age	-0.000***	-0.001	-0.000	-0.001**	-0.000**	-0.002***	-0.001***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnicity							
Base category – white							
Mixed	-0.002	-0.099	0.019	-	0.003	0.009	-0.011
	(0.006)	(0.095)	(0.023)	-	(0.004)	(0.021)	(0.010)
Asian	-0.001	-0.000	-0.080*	0.003	-0.011	-0.039	0.004

	(0.005)	(0.061)	(0.035)	(0.015)	(0.006)	(0.024)	(0.006)
Black	-	-	0.004	-	-0.009	-0.010	-0.001
	-	-	(0.027)	-	(0.014)	(0.035)	(0.017)
Other	-0.001	-0.044	-0.003	-0.058	-0.002	-0.016	0.007
	(0.009)	(0.175)	(0.025)	(0.051)	(0.008)	(0.026)	(0.012)
Missing	-0.001	0.018	-0.061**	-0.006	-0.001	-0.007	0.001
	(0.003)	(0.033)	(0.023)	(0.010)	(0.002)	(0.011)	(0.004)
Distance							
Base category – 0-2 miles							
2-4 miles	0.000	0.102*	0.035	0.006	-0.001	0.002	-0.001
	(0.002)	(0.040)	(0.025)	(0.006)	(0.002)	(0.008)	(0.003)
4-6 miles	0.003	0.049	0.092***	0.016**	0.003*	0.020*	0.003
	(0.002)	(0.047)	(0.022)	(0.006)	(0.002)	(0.008)	(0.004)
6-10 miles	0.002	0.050	0.081***	0.000	0.002	0.012	0.005
	(0.002)	(0.043)	(0.022)	(0.008)	(0.002)	(0.009)	(0.004)
10-14 miles	-0.003	0.102*	0.067**	0.014	0.000	-0.029	-0.003
	(0.004)	(0.047)	(0.023)	(0.009)	(0.003)	(0.016)	(0.005)
Over 14 miles	-0.005	0.025	0.080***	0.015	-0.003	-0.063***	-0.006
	(0.003)	(0.043)	(0.022)	(0.008)	(0.003)	(0.016)	(0.005)
Missing	0.002	0.001	0.028	0.001	0.001	-0.030*	0.004
	(0.003)	(0.055)	(0.028)	(0.011)	(0.002)	(0.012)	(0.004)
Financial quarters							
Base category – quarter 5							
Quarter 6	-0.002	-0.001	0.023	0.024*	0.002	-0.001	0.007
	(0.005)	(0.039)	(0.030)	(0.012)	(0.003)	(0.018)	(0.007)
Quarter 7	0.007	-0.050	0.028	0.004	0.004	0.009	0.009
	(0.004)	(0.041)	(0.024)	(0.015)	(0.003)	(0.018)	(0.006)
Quarter 8	0.005	0.002	-	0.026*	0.007**	0.023	0.017**
	(0.004)	(0.042)	-	(0.012)	(0.003)	(0.017)	(0.006)
Quarter 9	0.007	-0.073	0.055*	0.002	0.001	0.017	0.011
	(0.004)	(0.050)	(0.025)	(0.015)	(0.003)	(0.017)	(0.006)
Quarter 10	0.006	0.018	-0.000	-0.001	0.005	0.037*	0.020***
	(0.004)	(0.043)	(0.003)	(0.015)	(0.003)	(0.017)	(0.006)
Quarter 11	0.009*	-0.001	0.032	0.018	0.003	0.030	0.018**
	(0.004)	(0.048)	(0.016)	(0.013)	(0.003)	(0.016)	(0.006)
Quarter 12	0.008*	-0.017	0.044*	0.026*	0.003	0.050**	0.020***
	(0.004)	(0.075)	(0.021)	(0.012)	(0.003)	(0.016)	(0.006)
Quarter 13	0.013***	0.021	0.033*	0.007	0.002	0.050**	0.023***
	(0.004)	(0.066)	(0.017)	(0.014)	(0.003)	(0.016)	(0.006)
Quarter 14	0.004	0.078	0.098***	0.008	0.004	0.045**	0.021***
	(0.004)	(0.053)	(0.025)	(0.014)	(0.003)	(0.016)	(0.006)
Quarter 15	0.008	0.011	0.057**	0.023	0.001	0.030	0.011
	(0.004)	(0.069)	(0.020)	(0.012)	(0.003)	(0.017)	(0.007)
Quarter 16	0.007	-0.149	0.057*	0.022	0.005	0.041*	0.020**
	(0.004)	(0.122)	(0.025)	(0.012)	(0.003)	(0.016)	(0.006)
Quarter 17	0.008*	-0.148	0.063***	0.025*	0.002	0.029	0.022***
	(0.004)	(0.099)	(0.019)	(0.012)	(0.004)	(0.017)	(0.006)

Quarter 18	0.012**	0.039	-	0.032**	0.007**	0.025	0.023***
	(0.004)	(0.056)	-	(0.011)	(0.003)	(0.017)	(0.006)
Elixhauser comorbidities							
(2)	-0.003	-0.104*	-0.014	0.006	-0.002	-0.013	0.002
	(0.002)	(0.044)	(0.045)	(0.005)	(0.002)	(0.010)	(0.004)
(3)	-0.001	-0.055	-0.012	-0.001	-0.001	-0.018	-0.001
	(0.002)	(0.053)	(0.035)	(0.007)	(0.002)	(0.013)	(0.004)
(4)	0.015*	-0.042	-	-0.003	0.010	-0.032	-0.008
	(0.007)	(0.083)	-	(0.011)	(0.006)	(0.019)	(0.006)
(5)	-0.026***	-	-	-	-	-	-0.042**
	(0.007)	-	-	-	-	-	(0.015)
(6)	0.003	-	-	-	-	-0.014	0.011
	(0.012)	-	-	-	-	(0.040)	(0.023)
(7)	0.012**	-0.057	0.014	0.018	0.004	-0.021	0.001
	(0.004)	(0.029)	(0.046)	(0.013)	(0.003)	(0.012)	(0.004)
(8)	0.013	0.166	-	-	-	0.013	-0.040*
	(0.009)	(0.200)	-	-	-	(0.052)	(0.018)
(9)	-	.	-	-	-0.017*	-0.106	-0.035
	-	.	-	-	(0.007)	(0.070)	(0.025)
(10)	-0.019**	-0.024	-	-	-	-0.151***	0.005
	(0.006)	(0.130)	-	-	-	(0.033)	(0.023)
(11)	-0.010***	0.086	-	-	-0.002	-0.019	-0.032***
	(0.003)	(0.091)	-	-	(0.003)	(0.014)	(0.005)
(12)	0.010	0.079	-0.058	0.006	0.005	0.032	-0.014*
	(0.009)	(0.090)	(0.051)	(0.018)	(0.006)	(0.039)	(0.007)
(13)	-0.017	-	-	-	-	-	-
	(0.012)	-	-	-	-	-	-
(14)	-	-	-	-	-	-0.070*	-
	-	-	-	-	-	(0.033)	-
(15)	-0.020***	-0.205	-	-	-0.007	-0.103*	0.026
	(0.006)	(0.149)	-	-	(0.007)	(0.043)	(0.019)
(16)	-	-	-	-	-	-	-0.062*
	-	-	-	-	-	-	(0.026)
(19)	-	-	-	-	-	-0.150*	-
	-	-	-	-	-	(0.067)	-
(21)	-0.013*	-0.197	-	-0.008	-	-0.060*	-0.009
	(0.006)	(0.176)	-	(0.018)	-	(0.029)	(0.012)
(22)	-	-	-	-	-	-0.122	-0.020
	-	-	-	-	-	(0.064)	(0.024)
(23)	-0.019	-	-	-	-	-	-
	(0.013)	-	-	-	-	-	-
(24)	-	-0.117	-	-	-	-	-0.038
	-	(0.144)	-	-	-	-	(0.024)

(25)	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
(26)	-0.012	-0.151	-	-	-	-0.135***	-0.014
	(0.006)	(0.153)	-	-	-	(0.037)	(0.012)
(27)	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
(28)	-0.002	-	-	-0.035*	-	0.018	-0.010
	(0.009)	-	-	(0.015)	-	(0.040)	(0.010)
(29)	-0.010	-	-	-	-	-0.027	-
	(0.012)	-	-	-	-	(0.055)	-
(3))	-	-0.066		-	-	-	-
	-	(0.142)	-	-	-	-	-
ICD 10 main chapters							
I210	0.017***	0.085	-0.006	0.010	0.002	0.059*	0.032***
	(0.005)	(0.071)	(0.032)	(0.020)	(0.005)	(0.024)	(0.009)
I211	0.018***	0.083	0.003	0.011	-0.001	0.051*	0.030***
	(0.005)	(0.070)	(0.032)	(0.022)	(0.005)	(0.022)	(0.009)
I212	0.029*	-0.035	0.005	-	0.003	0.018	0.018
	(0.013)	(0.088)	(0.045)	-	(0.008)	(0.028)	(0.012)
I214	0.004	-0.085	-	-0.010	-0.005	0.003	0.007
	(0.004)	(0.092)	-	(0.020)	(0.005)	(0.025)	(0.008)
I219	0.012**	0.048	0.002	-0.008	0.001	0.012	0.015*
	(0.004)	(0.075)	(0.039)	(0.018)	(0.005)	(0.021)	(0.007)
I220	.	0.129	0.012	0.004	-	0.020	0.013
	.	(0.102)	(0.055)	(0.025)	-	(0.033)	(0.012)
I221	0.017*	-0.032	0.036	-0.000	-0.001	0.054	0.023
	(0.008)	(0.081)	(0.054)	(0.025)	(0.007)	(0.034)	(0.013)
I228	0.015*	-0.138	-	-0.009	-0.002	-0.013	0.003
	(0.007)	(0.113)	-	(0.023)	(0.006)	(0.032)	(0.010)
I229	0.018***	-0.045	-	-0.004	-0.000	-0.013	0.013
	(0.004)	(0.092)	-	(0.018)	(0.005)	(0.022)	(0.007)
I251	0.024*	-0.028	-0.014	-	0.001	0.017	0.037**
	(0.009)	(0.091)	(0.033)	-	(0.006)	(0.024)	(0.012)
Observations	**27744**	**1115**	**2728**	**4482**	**22716**	**7095**	**21561**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for AMI, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Elixhauser Comorbidities: (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (19) Lymphoma; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (27) Blood loss anemia; (28) Deficiency anemia; (29) Alcohol abuse; and (30) Drug abuse. Not all patients have Elixhauser comorbidities; therefore a base category is not necessary. If no patients in the sample have certain Elixhauser comorbidity, the certain Elixhauser comorbidity is not included in the results. Coefficients in the table of 0.01 signify 1%.

Table 36: Logistic regression results: pneumonia – Full model with all covariates

	Oxygenation Assessment	Blood Cultures	Antibiotic Selection For CAP	Initial Antibiotics	Smoking Cessation
Monday	0.002	-0.023*	0.003	-0.017*	-0.001
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Tuesday	0.003*	-0.021	-0.002	-0.011	-0.009
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Base category – Wednesday					
Thursday	0.002	-0.016	0.005	-0.008	-0.006
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Friday	0.001	-0.026*	-0.011	-0.008	-0.027
	(0.001)	(0.011)	(0.006)	(0.008)	(0.016)
Saturday	0.002	-0.033**	-0.005	0.010	-0.026
	(0.001)	(0.012)	(0.007)	(0.008)	(0.016)
Sunday	0.002	-0.027*	-0.001	0.019*	0.010
	(0.001)	(0.012)	(0.006)	(0.008)	(0.016)
Trust characteristics					
Small	-0.002	0.042***	-0.056***	-0.038***	-0.017
	(0.001)	(0.011)	(0.006)	(0.008)	(0.017)
Medium	0.003***	0.049***	-0.045***	0.015**	0.103***
	(0.001)	(0.008)	(0.004)	(0.005)	(0.011)
Base category - large					
Specialist/Teaching	-0.000	0.013	-0.053***	0.010	0.187***
	(0.001)	(0.011)	(0.005)	(0.007)	(0.012)
Foundation Trust	0.003***	-0.028***	-0.027***	-0.029***	-0.133***
	(0.001)	(0.007)	(0.004)	(0.005)	(0.009)
Day volume of admissions (00's)	0.001	0.033	-0.008	-0.012	-0.044
	(0.002)	(0.020)	(0.010)	(0.013)	(0.025)
CQC quality					
Base category - excellent					
Good	0.001	-0.038***	-0.051***	-0.022***	0.035**
	(0.001)	(0.008)	(0.004)	(0.006)	(0.011)
Fair/poor	-0.001	-0.122***	-0.039***	-0.031***	0.032*
	(0.002)	(0.013)	(0.006)	(0.008)	(0.015)
Male	-0.001*	0.005	0.002	0.022***	-0.007
	(0.001)	(0.006)	(0.003)	(0.004)	(0.008)
Age	-0.000	0.000	0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnicity					

Base category - white					
Mixed	0.002	-0.002	-0.007	-0.014	-0.014
	(0.002)	(0.024)	(0.015)	(0.018)	(0.034)
Asian	-0.007*	-0.018	-0.022	0.016	-0.059
	(0.003)	(0.023)	(0.013)	(0.016)	(0.039)
Black	0.002	0.030	-0.033	-0.015	0.019
	(0.003)	(0.040)	(0.024)	(0.030)	(0.066)
Other	-0.000	0.028	-0.009	-0.000	0.029
	(0.003)	(0.033)	(0.018)	(0.025)	(0.050)
Missing	0.000	0.012	0.010	0.003	0.001
	(0.001)	(0.014)	(0.007)	(0.010)	(0.018)
Distance					
Base category - 0-2 miles					
2-4 miles	-0.000	0.002	-0.001	-0.013*	0.016
	(0.001)	(0.008)	(0.004)	(0.006)	(0.011)
4-6 miles	0.001	0.015	-0.022***	-0.014*	0.029*
	(0.001)	(0.010)	(0.006)	(0.007)	(0.014)
6-10 miles	-0.000	0.018	-0.021**	-0.009	0.002
	(0.001)	(0.011)	(0.007)	(0.008)	(0.016)
10-14 miles	0.002	0.005	0.004	-0.038**	0.020
	(0.001)	(0.018)	(0.010)	(0.012)	(0.023)
Over 14 miles	-0.007*	-0.017	-0.015	0.001	-0.095***
	(0.003)	(0.023)	(0.013)	(0.014)	(0.026)
Missing	-0.001	-0.001	0.002	-0.032***	0.031*
	(0.001)	(0.013)	(0.006)	(0.009)	(0.015)
Financial quarters					
Base category - quarter 5					
Quarter 6	0.004	0.081***	0.023*	0.066***	0.027
	(0.003)	(0.020)	(0.011)	(0.013)	(0.024)
Quarter 7	0.008**	0.134***	0.020	0.098***	-0.000
	(0.003)	(0.019)	(0.011)	(0.013)	(0.024)
Quarter 8	0.007**	0.124***	0.047***	0.092***	0.051*
	(0.003)	(0.020)	(0.011)	(0.014)	(0.025)
Quarter 9	0.009**	0.142***	0.081***	0.080***	0.058*
	(0.003)	(0.019)	(0.010)	(0.013)	(0.022)
Quarter 10	0.011***	0.133***	0.083***	0.118***	0.094***
	(0.003)	(0.019)	(0.010)	(0.013)	(0.023)
Quarter 11	0.012***	0.158***	0.062***	0.118***	0.072**
	(0.002)	(0.019)	(0.010)	(0.013)	(0.023)
Quarter 12	0.009***	0.150***	0.059***	0.104***	0.131***
	(0.003)	(0.019)	(0.011)	(0.014)	(0.024)
Quarter 13	0.010***	0.168***	0.075***	0.110***	0.128***
	(0.003)	(0.018)	(0.010)	(0.013)	(0.023)
Quarter 14	0.010***	0.147***	0.070***	0.119***	0.169***
	(0.003)	(0.019)	(0.010)	(0.013)	(0.022)
Quarter 15	0.012***	0.199***	0.061***	0.138***	0.145***
	(0.002)	(0.018)	(0.010)	(0.012)	(0.023)

Quarter 16	0.014***	0.205***	0.051***	0.128***	0.126***
	(0.002)	(0.018)	(0.011)	(0.013)	(0.024)
Quarter 17	0.012***	0.255***	0.045***	0.146***	0.213***
	(0.002)	(0.020)	(0.011)	(0.012)	(0.022)
Quarter 18	0.011***	-	0.047***	0.167***	0.200***
	(0.003)	-	(0.011)	(0.012)	(0.023)
Elixhauser Comorbidities					
(2)	0.002	-0.003	0.009	-0.000	-0.081
	(0.002)	(0.021)	(0.011)	(0.013)	(0.042)
(3)	0.002	0.022	0.008	0.011	-0.014
	(0.002)	(0.016)	(0.009)	(0.011)	(0.031)
(4)	-	-0.001	-0.005	-0.052	-0.052
	-	(0.069)	(0.033)	(0.041)	(0.096)
(5)	-0.001	-0.051	0.032	-0.053	0.163
	(0.005)	(0.060)	(0.038)	(0.038)	(0.096)
(6)	-	0.191	0.003	0.032	0.008
	-	(0.151)	(0.046)	(0.060)	(0.097)
(7)	0.003	0.040	-0.014	0.002	-0.001
	(0.002)	(0.021)	(0.009)	(0.013)	(0.029)
(8)	-	0.037	-0.011	0.004	0.111
	-	(0.071)	(0.043)	(0.050)	(0.150)
(9)	-	0.008	0.000	-0.054	0.050
	-	(0.082)	(0.047)	(0.048)	(0.110)
(10)	0.008	-0.007	-0.002	-0.004	-0.049
	(0.005)	(0.028)	(0.016)	(0.020)	(0.054)
(11)	0.002**	0.007	-0.001	0.042***	0.072***
	(0.001)	(0.007)	(0.004)	(0.005)	(0.009)
(12)	0.009	0.049	0.004	-0.021	-0.056
	(0.005)	(0.026)	(0.013)	(0.017)	(0.039)
(13)	-	0.066	0.044	-0.121*	0.159
	-	(0.110)	(0.065)	(0.058)	(0.171)
(14)	-	0.059	-0.038	-0.002	0.068
	-	(0.060)	(0.025)	(0.038)	(0.101)
(15)	0.003	-0.082	-0.003	-0.001	-0.021
	(0.005)	(0.049)	(0.026)	(0.030)	(0.086)
(16)	-0.006	0.040	-0.017	-0.128**	-0.122*
	(0.004)	(0.074)	(0.031)	(0.041)	(0.059)
(17)	-0.011	-0.333	0.012	0.201	-
	(0.006)	(0.178)	(0.092)	(0.186)	-
(19)	-0.002	0.005	-0.020	-0.037	0.058
	(0.004)	(0.048)	(0.043)	(0.036)	(0.086)
(20)	0.000	-0.066	-0.051	-0.054	-0.167
	(0.005)	(0.055)	(0.040)	(0.036)	(0.092)

(21)	-0.003*	-0.027	-0.030*	-0.023	-0.084**
	(0.002)	(0.022)	(0.015)	(0.015)	(0.031)
(22)	0.001	0.137	0.025	-0.071*	-0.042
	(0.005)	(0.077)	(0.033)	(0.034)	(0.079)
(23)	-0.009	-0.253**	0.029	0.113	0.244
	(0.005)	(0.084)	(0.065)	(0.096)	(0.239)
(24)	-	-0.224	-0.024	0.004	-0.111
	-	(0.146)	(0.056)	(0.103)	(0.136)
(25)	-	-0.099	-0.015	-0.098	0.281*
	-	(0.092)	(0.036)	(0.050)	(0.119)
(26)	-0.000	0.004	0.018	-0.018	-0.039
	(0.002)	(0.023)	(0.013)	(0.014)	(0.036)
(28)	0.002	0.042	-0.001	0.091*	0.077
	(0.005)	(0.049)	(0.024)	(0.036)	(0.076)
(29)	0.005	0.006	-0.017	0.005	0.084*
	(0.006)	(0.043)	(0.020)	(0.030)	(0.037)
(30)	-	-0.001	-0.035	0.013	-0.055
	-	(0.063)	(0.029)	(0.048)	(0.048)
(31)	-	-0.037	-0.036	0.027	-0.041
	-	(0.086)	(0.040)	(0.071)	(0.091)
(32)	-0.008*	-0.048	0.012	0.054	0.021
	(0.003)	(0.051)	(0.028)	(0.046)	(0.068)
ICD 10 main chapters					
J13	0.002	0.032	0.001	0.154***	-0.044
	(0.003)	(0.025)	(0.018)	(0.023)	(0.035)
J14	0.005	0.031	0.024	0.096*	0.063
	(0.006)	(0.048)	(0.030)	(0.037)	(0.055)
J15	-0.001	-0.013	0.001	0.026	-0.028
	(0.002)	(0.026)	(0.018)	(0.021)	(0.039)
J18	0.003	0.028	-0.002	0.019	-0.040
	(0.002)	(0.018)	(0.011)	(0.013)	(0.026)
Observations	51881	14916	30844	37861	12543

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, and Main ICD 10 chapters for pneumonia, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Elixhauser Comorbidities: (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (17) Peptic ulcer disease; (19) Lymphoma; (20) Metastatic cancer; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (28) Deficiency anemia; (29) Alcohol abuse; (30) Drug abuse; (31) Psychoses; and (32) Depression. Not all patients have Elixhauser comorbidities; therefore a base category is not necessary. If no patients in the sample have certain Elixhauser comorbidity, the certain Elixhauser comorbidity is not included in the results. Coefficients in the table of 0.01 signify 1%.

Table 37: Appropriate care scores

	AMI		Pneumonia	
	ACS	ACS admission processes	ACS	ACS admission processes
Monday	-0.001	-0.002	-0.004	-0.005
	(0.002)	(0.003)	(0.003)	(0.003)
Tuesday	0.000	-0.001	-0.007*	-0.004
	(0.002)	(0.003)	(0.003)	(0.003)
Base category – Wednesday				
Thursday	0.000	0.002	-0.002	-0.002
	(0.002)	(0.003)	(0.003)	(0.003)
Friday	0.001	0.002	-0.009*	-0.007*
	(0.002)	(0.003)	(0.003)	(0.003)
Saturday	0.002	0.002	-0.005	-0.003
	(0.002)	(0.003)	(0.004)	(0.003)
Sunday	-0.002	-0.003	0.003	0.000
	(0.002)	(0.003)	(0.003)	(0.003)
Observations	**34509**	**29804**	**55302**	**52884**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: appropriate care score, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for AMI, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 38: Full model results with different ways to measure days of the week – AMI

	Aspirin Arrival	ACEI for LVSD	Fibrinolytic Therapy	Primary PCI	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.003	0.011	-0.017	-0.014	-0.002	0.012	-0.001
	(0.003)	(0.008)	(0.039)	(0.015)	(0.002)	(0.010)	(0.004)
Tuesday	-0.002	0.006	-0.000	-0.004	-0.004	0.024*	-0.000
	(0.003)	(0.009)	(0.038)	(0.015)	(0.002)	(0.010)	(0.004)
Base category – Wednesday							
Thursday	0.002	0.009	-0.046	0.002	0.001	0.003	-0.002
	(0.002)	(0.008)	(0.040)	(0.013)	(0.002)	(0.011)	(0.004)
Friday	0.001	0.015	-0.002	-0.013	0.002	0.007	0.002
	(0.002)	(0.008)	(0.039)	(0.015)	(0.002)	(0.011)	(0.004)
Saturday	0.000	0.012	0.014	-0.011	-0.002	0.023*	0.003
	(0.002)	(0.008)	(0.038)	(0.015)	(0.002)	(0.011)	(0.004)
Sunday	-0.001	0.005	-0.027	-0.030	0.001	0.011	0.001
	(0.003)	(0.009)	(0.040)	(0.016)	(0.002)	(0.011)	(0.004)
Weekend	-0.000	0.007	0.003	0.003	-0.000	0.009	0.002
	(0.002)	(0.006)	(0.024)	(0.022)	(0.001)	(0.006)	(0.002)
Weekend - BH	-0.001	0.001	-0.001	-0.017	-0.000	0.009	0.002
	(0.001)	(0.005)	(0.022)	(0.009)	(0.001)	(0.006)	(0.002)
Bank holiday	-0.008	-	-0.024	-0.004	-0.002	-0.006	-0.002
	(0.006)	-	(0.066)	(0.028)	(0.004)	(0.021)	(0.008)
Observations	**27744**	**4482**	**1115**	**2728**	**22716**	**7095**	**21561**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Weekend and bank holiday variable, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for AMI, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume, time fixed effects, interaction terms between weekend and bank holiday variable with Trust size and CQC score. Coefficients in the table of 0.01 signify 1%.

Table 39: Full model results with different ways to measure days of the week – pneumonia

	Oxygenation Assessment	Blood Cultures	Antibiotic Selection For CAP	Initial Antibiotics	Smoking Cessation
Monday	0.002	-0.023*	0.003	-0.017*	-0.001
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Tuesday	0.003*	-0.021	-0.002	-0.011	-0.009
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Base category – Wednesday					
Thursday	0.002	-0.016	0.005	-0.008	-0.006
	(0.001)	(0.011)	(0.006)	(0.008)	(0.015)
Friday	0.001	-0.026*	-0.011	-0.008	-0.027
	(0.001)	(0.011)	(0.006)	(0.008)	(0.016)
Saturday	0.002	-0.033**	-0.005	0.010	-0.026
	(0.001)	(0.012)	(0.007)	(0.008)	(0.016)
Sunday	0.002	-0.027*	-0.001	0.019*	0.010
	(0.001)	(0.012)	(0.006)	(0.008)	(0.016)
Weekend	-0.000	-0.018	-0.004	0.008	0.020
	(0.002)	(0.021)	(0.011)	(0.014)	(0.026)
Weekend -BH	0.000	-0.012	-0.001	0.024***	-0.006
	(0.001)	(0.007)	(0.004)	(0.005)	(0.009)
Bank holiday	0.000	0.017	0.004	-0.008	-0.020
	(0.002)	(0.019)	(0.011)	(0.014)	(0.027)
Observations	51881	14916	30844	37861	12543

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Weekend and bank holiday variable, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for pneumonia, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume, time fixed effects, interaction terms between weekend and bank holiday variable with Trust size and CQC score. Coefficients in the table of 0.01 signify 1%.

Table 40: Sensitivity to inclusion of continuous distance measure – AMI

	Aspirin Arrival	Fibrinolytic Therapy	Primary PCI	ACEI for LVSD	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.004	-0.017	0.010	0.011	-0.002	0.008	-0.001
	(0.003)	(0.043)	(0.014)	(0.009)	(0.002)	(0.011)	(0.004)
Tuesday	-0.004	-0.004	0.002	0.006	-0.003	0.024*	0.000
	(0.003)	(0.043)	(0.015)	(0.009)	(0.002)	(0.011)	(0.004)
Base category – Wednesday							
Thursday	0.001	-0.025	0.007	0.008	0.002	0.005	-0.000
	(0.003)	(0.043)	(0.014)	(0.009)	(0.002)	(0.012)	(0.004)
Friday	0.002	0.014	-0.013	0.017*	0.002	0.013	0.002
	(0.002)	(0.043)	(0.017)	(0.008)	(0.002)	(0.011)	(0.004)
Saturday	0.000	0.019	-0.017	0.012	-0.002	0.016	0.005
	(0.003)	(0.041)	(0.017)	(0.009)	(0.002)	(0.011)	(0.004)
Sunday	-0.002	-0.029	-0.033	0.007	0.001	0.009	0.003
	(0.003)	(0.044)	(0.018)	(0.010)	(0.002)	(0.012)	(0.004)
Distance	-0.000	-0.000	0.001*	0.000	-0.000	-0.001***	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	**24043**	**901**	**2132**	**3688**	**19091**	**5906**	**18040**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for AMI, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 41: Sensitivity to inclusion of continuous distance measure – pneumonia

	Oxygenation Assessment	Blood Cultures	Antibiotic Selection For CAP	Initial Antibiotics	Smoking Cessation
Monday	0.002	-0.024	-0.000	-0.014	0.003
	(0.001)	(0.012)	(0.006)	(0.009)	(0.016)
Tuesday	0.003*	-0.019	-0.003	-0.014	-0.002
	(0.001)	(0.012)	(0.007)	(0.009)	(0.017)
Base category – Wednesday					
Thursday	0.002	-0.013	0.004	-0.008	0.002
	(0.001)	(0.012)	(0.007)	(0.009)	(0.017)
Friday	0.001	-0.018	-0.014*	-0.006	-0.028
	(0.001)	(0.012)	(0.007)	(0.009)	(0.017)
Saturday	0.002	-0.035**	-0.006	0.013	-0.023
	(0.001)	(0.012)	(0.007)	(0.009)	(0.017)
Sunday	0.002	-0.023	-0.003	0.022*	0.011
	(0.001)	(0.012)	(0.007)	(0.009)	(0.017)
Distance	-0.000	0.000	-0.000	-0.000	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	45347	13436	27182	33215	10898

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, and Main ICD 10 chapters for pneumonia, distance from patients LSOA to Trust site, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 42: Full model with hospital utilisation – AMI

	Aspirin Arrival	Fibrinolytic Therapy	Primary PCI	ACEI for LVSD	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.004	-0.009	-0.006	0.015	-0.000	0.015	-0.003
	(0.003)	(0.040)	(0.018)	(0.012)	(0.003)	(0.014)	(0.005)
Tuesday	-0.001	0.001	-0.019	0.013	-0.003	0.026	-0.003
	(0.003)	(0.040)	(0.021)	(0.012)	(0.003)	(0.014)	(0.005)
Base category – Wednesday							
Thursday	0.002	-0.027	-0.015	0.008	0.003	0.010	-0.004
	(0.003)	(0.041)	(0.019)	(0.012)	(0.003)	(0.015)	(0.005)
Friday	0.001	-0.014	-0.019	0.023*	0.004	0.010	0.003
	(0.003)	(0.043)	(0.021)	(0.011)	(0.003)	(0.015)	(0.005)
Saturday	0.002	0.003	-0.003	0.019	-0.001	0.028	0.003
	(0.003)	(0.042)	(0.020)	(0.011)	(0.003)	(0.014)	(0.005)
Sunday	0.001	-0.025	-0.036	0.014	0.003	0.012	0.002
	(0.003)	(0.043)	(0.022)	(0.012)	(0.003)	(0.015)	(0.005)
Utilisation	0.004	-0.209	0.024	0.039	0.013	0.033	0.011
	(0.012)	(0.223)	(0.052)	(0.039)	(0.012)	(0.051)	(0.022)
Observations	**19138**	**1016**	**1406**	**2813**	**13314**	**4394**	**13635**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for AMI, Trust type, CQC quality metric, Foundation Trust status, day volume, hospital utilisation and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 43: Full model with hospital utilisation – pneumonia

	Oxygenation Assessment	Blood Cultures	Antibiotic Selection For CAP	Initial Antibiotics	Smoking Cessation
Monday	0.004*	-0.016	-0.003	-0.012	-0.003
	(0.002)	(0.013)	(0.007)	(0.010)	(0.018)
Tuesday	0.003	-0.020	-0.007	-0.010	-0.023
	(0.002)	(0.013)	(0.008)	(0.010)	(0.019)
Base category – Wednesday					
Thursday	0.003	-0.012	0.005	-0.009	0.002
	(0.002)	(0.013)	(0.007)	(0.010)	(0.019)
Friday	0.002	-0.024	-0.016*	-0.008	-0.021
	(0.002)	(0.013)	(0.008)	(0.010)	(0.019)
Saturday	0.003	-0.037**	-0.008	0.012	-0.011
	(0.002)	(0.014)	(0.008)	(0.010)	(0.020)
Sunday	0.003	-0.028*	-0.002	0.015	0.019
	(0.002)	(0.013)	(0.008)	(0.010)	(0.019)
Utilisation	0.014*	-0.009	-0.026	-0.081*	0.184**
	(0.006)	(0.051)	(0.027)	(0.037)	(0.070)
Observations	35126	12115	21298	25726	8947

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, Elixhauser comorbidities, Main ICD 10 chapters for pneumonia, Trust type, CQC quality metric, Foundation Trust status, day volume, hospital utilisation and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 44: Full model with Charlson Index – AMI

	Aspirin Arrival	Fibrinolytic Therapy	Primary PCI	ACEI for LVSD	Aspirin Discharge	Smoking Cessation	Beta-Blocker Discharge
Monday	-0.003	-0.007	-0.017	0.010	-0.002	0.008	-0.001
	(0.003)	(0.039)	(0.015)	(0.008)	(0.002)	(0.010)	(0.004)
Tuesday	-0.002	0.005	-0.004	0.005	-0.003	0.020*	-0.001
	(0.003)	(0.038)	(0.014)	(0.008)	(0.002)	(0.010)	(0.004)
Base category – Wednesday							
Thursday	0.003	-0.047	-0.002	0.008	0.001	-0.001	-0.001
	(0.002)	(0.041)	(0.013)	(0.008)	(0.002)	(0.011)	(0.004)
Friday	0.001	0.003	-0.020	0.013	0.002	0.007	0.002
	(0.002)	(0.039)	(0.015)	(0.007)	(0.002)	(0.010)	(0.004)
Saturday	0.000	0.015	-0.014	0.011	-0.001	0.018	0.004
	(0.002)	(0.038)	(0.015)	(0.008)	(0.002)	(0.010)	(0.004)
Sunday	-0.002	-0.026	-0.032*	0.005	0.000	0.007	0.002
	(0.003)	(0.040)	(0.016)	(0.008)	(0.002)	(0.011)	(0.004)
Charlson	-0.002**	-0.030*	-0.003	-0.002	-0.001*	-0.009***	-0.002
	(0.001)	(0.014)	(0.005)	(0.002)	(0.000)	(0.002)	(0.001)
Observations	**27444**	**1112**	**2606**	**4368**	**22598**	**6905**	**21196**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, weighted Charlson index, Main ICD 10 chapters for pneumonia, Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 45: Full model with Charlson Index - pneumonia

	Oxygenation Assessment	Blood Cultures	Antibiotic Selection For CAP	Initial Antibiotics	Smoking Cessation
Monday	0.002	-0.024*	0.004	-0.019*	-0.006
	(0.001)	(0.012)	(0.006)	(0.008)	(0.015)
Tuesday	0.002	-0.022	-0.002	-0.011	-0.010
	(0.001)	(0.011)	(0.006)	(0.008)	(0.016)
Base category – Wednesday					
Thursday	0.002	-0.016	0.006	-0.009	-0.010
	(0.001)	(0.011)	(0.006)	(0.008)	(0.016)
Friday	0.001	-0.025*	-0.010	-0.010	-0.028
	(0.001)	(0.011)	(0.006)	(0.008)	(0.016)
Saturday	0.002	-0.035**	-0.004	0.010	-0.030
	(0.001)	(0.012)	(0.007)	(0.008)	(0.016)
Sunday	0.001	-0.029*	-0.000	0.020*	0.005
	(0.001)	(0.012)	(0.006)	(0.008)	(0.016)
Charlson	-0.000	-0.000	-0.001	0.000	-0.004
	(0.000)	(0.002)	(0.002)	(0.002)	(0.003)
Observations	51842	14709	30257	37240	12292

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Full model specification: Individual days of the week, age, gender, ethnicity, weighted Charlson index, Main ICD 10 chapters for pneumonia Trust type, CQC quality metric, Foundation Trust status, day volume and time fixed effects. Coefficients in the table of 0.01 signify 1%.

Table 46: In-hospital mortality - AMI

	In-hospital Mortality						
	A	B	C	D	E	F	G
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5*	Model 5**
Weekend Dummy	0.008*	0.003	-0.001	0.001	0.001	-0.001	-0.006
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.008)
Day volume of admissions (00's)							
	-	-0.001***	-0.001***	-0.000	-0.000	0.000	-0.108
	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.146)
Patient Characteristics							
Male	-	-	0.002	0.002	0.002	0.003	0.003
	-	-	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Age	-	-	0.005***	0.005***	0.005***	0.005***	0.005***
	-	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnicity							
Base category - white							
Mixed	-	-	-0.013	-0.017	-0.017	-0.023	-0.023
	-	-	(0.011)	(0.010)	(0.010)	(0.013)	(0.013)
Asian	-	-	0.017	0.012	0.012	0.013	0.013
	-	-	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)
Black	-	-	-0.034	-0.035	-0.036	-0.076***	-0.076***
	-	-	(0.018)	(0.018)	(0.018)	(0.014)	(0.014)
Other	-	-	0.001	0.002	0.001	-0.007	-0.007
	-	-	(0.014)	(0.014)	(0.014)	(0.018)	(0.018)
Missing	-	-	0.007	0.008	0.008	0.011	0.011
	-	-	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)
Distance							
Base category - 0-2 miles							
2-4 miles	-	-	-0.000	0.001	0.001	0.001	0.001
	-	-	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
4-6 miles	-	-	-0.002	-0.002	-0.002	0.000	0.000
	-	-	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
6-10 miles	-	-	-0.007	-0.010*	-0.010*	-0.011	-0.011
	-	-	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
10-14 miles	-	-	-0.014*	-0.017**	-0.017**	-0.016*	-0.016*
	-	-	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)
Over 14 miles	-	-	-0.010	-0.012*	-0.013*	-0.010	-0.010
	-	-	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)
Missing	-	-	0.003	0.003	0.003	-0.002	-0.002
	-	-	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
Elixhauser comorbidities							
(2)	-	-	0.042***	0.042***	0.042***	0.047***	0.047***
	-	-	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
(3)	-	-	-0.013*	-0.014*	-0.014*	-0.015*	-0.015*
	-	-	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
(4)	-	-	-0.020*	-0.017	-0.017	-0.016	-0.016

	-	-	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)
(5)	-	-	0.062*	0.063*	0.063*	0.091*	0.091*
	-	-	(0.028)	(0.028)	(0.028)	(0.035)	(0.035)
(6)	-	-	0.030	0.030	0.030	0.037	0.037
	-	-	(0.018)	(0.018)	(0.018)	(0.023)	(0.023)
(7)	-	-	-0.076***	-0.079***	-0.079***	-0.086***	-0.086***
	-	-	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)
(8)	-	-	0.016	0.010	0.010	0.015	0.015
	-	-	(0.023)	(0.023)	(0.023)	(0.032)	(0.032)
(9)	-	-	-0.017	-0.026	-0.027	-0.007	-0.006
	-	-	(0.068)	(0.070)	(0.071)	(0.081)	(0.081)
(10)	-	-	0.160***	0.156***	0.156***	0.161***	0.162***
	-	-	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
(11)	-	-	0.009	0.006	0.006	-0.005	-0.005
	-	-	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)
(12)	-	-	-0.032*	-0.036*	-0.035*	-0.049*	-0.049*
	-	-	(0.014)	(0.014)	(0.014)	(0.019)	(0.019)
(13)	-	-	-0.048	-0.049	-0.049	0.005	0.005
	-	-	(0.072)	(0.072)	(0.072)	(0.078)	(0.078)
(14)	-	-	-0.138*	-0.140*	-0.140*	-0.134	-0.134
	-	-	(0.070)	(0.070)	(0.070)	(0.076)	(0.076)
(15)	-	-	0.004	0.003	0.002	0.001	0.001
	-	-	(0.018)	(0.019)	(0.019)	(0.029)	(0.029)
(16)	-	-	0.149*	0.141*	0.141*	0.185**	0.186**
	-	-	(0.062)	(0.059)	(0.059)	(0.064)	(0.064)
(19)	-	-	0.088*	0.093*	0.093*	0.021	0.021
	-	-	(0.037)	(0.037)	(0.036)	(0.080)	(0.080)
(21)	-	-	0.064	0.059	0.058	-0.029	-0.028
	-	-	(0.050)	(0.053)	(0.053)	(0.079)	(0.079)
(22)	-	-	0.052***	0.050***	0.050***	0.051**	0.051**
	-	-	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)
(23)	-	-	-0.071	-0.075	-0.075	-0.073	-0.073
	-	-	(0.073)	(0.073)	(0.073)	(0.077)	(0.077)
(24)	-	-	-0.028	-0.034	-0.033	-0.018	-0.018
	-	-	(0.077)	(0.077)	(0.077)	(0.080)	(0.080)
(25)	-	-	0.178**	0.173**	0.173**	0.218**	0.219**
	-	-	(0.059)	(0.059)	(0.059)	(0.079)	(0.079)
(26)	-	-	0.046**	0.046**	0.046**	0.060**	0.059**
	-	-	(0.016)	(0.016)	(0.016)	(0.019)	(0.019)
(27)	-	-	0.138*	0.128*	0.127*	0.130	0.130
	-	-	(0.063)	(0.063)	(0.063)	(0.070)	(0.070)
(28)	-	-	-0.051*	-0.051*	-0.051*	-0.054	-0.054
	-	-	(0.024)	(0.024)	(0.024)	(0.030)	(0.030)
(29)	-	-	-0.061	-0.060	-0.060	-0.055	-0.055
	-	-	(0.073)	(0.073)	(0.073)	(0.078)	(0.078)
ICD 10 sub chapters							
I210	-	-	-0.036***	-0.032**	-0.032**	-0.016	-0.016

	-	-	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)
I211	-	-	-0.050***	-0.045***	-0.045***	-0.039**	-0.039**
	-	-	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)
I212	-	-	-0.034*	-0.030*	-0.031*	-0.016	-0.016
	-	-	(0.014)	(0.014)	(0.014)	(0.018)	(0.018)
I214	-	-	-0.076***	-0.073***	-0.073***	-0.062***	-0.062***
	-	-	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)
I219	-	-	-0.050***	-0.047***	-0.048***	-0.034**	-0.034**
	-	-	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)
I220	-	-	-0.033*	-0.033*	-0.034*	-0.019	-0.020
	-	-	(0.014)	(0.015)	(0.015)	(0.017)	(0.017)
I221	-	-	-0.027	-0.026	-0.026	-0.013	-0.014
	-	-	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)
I228	-	-	-0.056***	-0.058***	-0.058***	-0.049**	-0.049**
	-	-	(0.015)	(0.015)	(0.015)	(0.018)	(0.018)
I229	-	-	-0.065***	-0.066***	-0.066***	-0.054***	-0.054***
	-	-	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)
I251	-	-	-0.124***	-0.118***	-0.118***	-0.083***	-0.083***
	-	-	(0.015)	(0.015)	(0.015)	(0.020)	(0.020)
Trust characteristics							
Small	-	-	-	0.001	0.001	-0.005	-0.005
	-	-	-	(0.005)	(0.005)	(0.006)	(0.006)
Medium	-	-	-	-0.020***	-0.020***	-0.021***	-0.021***
	-	-	-	(0.004)	(0.004)	(0.005)	(0.005)
Base category - **large**							
Specialist/Teaching	-	-	-	-0.016***	-0.016***	-0.017**	-0.017**
	-	-	-	(0.004)	(0.004)	(0.006)	(0.006)
Foundation Trust	-	-	-	-0.011***	-0.011***	-0.009*	-0.009*
	-	-	-	(0.003)	(0.003)	(0.004)	(0.004)
CQC quality							
Base category - excellent							
Good	-	-	-	-0.012**	-0.012**	-0.010*	-0.010*
	-	-	-	(0.004)	(0.004)	(0.005)	(0.005)
Fair/poor	-	-	-	-0.004	-0.004	0.002	0.002
Financial quarters	-	-	-	(0.005)	(0.005)	(0.007)	(0.007)
Base category - quarter 5							
Quarter 6	-	-	-	-0.000	-0.001	0.001	0.001
	-	-	-	(0.007)	(0.007)	(0.008)	(0.008)
Quarter 7	-	-	-	0.001	0.001	0.002	0.001
	-	-	-	(0.007)	(0.007)	(0.008)	(0.008)
Quarter 8	-	-	-	0.006	0.006	0.007	0.005
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 9	-	-	-	0.004	0.004	0.005	0.005
	-	-	-	(0.007)	(0.007)	(0.007)	(0.007)
Quarter 10	-	-	-	0.003	0.002	0.003	0.003
	-	-	-	(0.007)	(0.007)	(0.008)	(0.008)
Quarter 11	-	-	-	0.008	0.008	0.009	0.008

	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 12	-	-	-	-0.009	-0.009	-0.009	-0.010
	-	-	-	(0.007)	(0.007)	(0.007)	(0.008)
Quarter 13	-	-	-	0.005	0.005	0.004	0.003
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 14	-	-	-	-0.006	-0.006	-0.007	-0.007
	-	-	-	(0.007)	(0.007)	(0.007)	(0.007)
Quarter 15	-	-	-	-0.024***	-0.024***	-	-
	-	-	-	(0.007)	(0.007)	-	-
Quarter 16	-	-	-	-0.025***	-0.025***	-	-
	-	-	-	(0.007)	(0.007)	-	-
Quarter 17	-	-	-	-0.008	-0.008	-	-
	-	-	-	(0.007)	(0.007)	-	-
Quarter 18	-	-	-	-0.006	-0.006	-	-
	-	-	-	(0.007)	(0.007)	-	-
Interactions with weekend							
Small	-	-	-	-	0.010	0.007	0.007
	-	-	-	-	(0.011)	(0.012)	(0.012)
Medium	-	-	-	-	0.007	0.006	0.006
	-	-	-	-	(0.006)	(0.007)	(0.007)
Large	-	-	-	-	0.006	0.001	0.001
	-	-	-	-	(0.006)	(0.007)	(0.007)
Specialist/Teaching	-	-	-	-	0.011	0.004	0.004
	-	-	-	-	(0.006)	(0.008)	(0.008)
Excellent	-	-	-	-	0.001	0.005	0.005
	-	-	-	-	(0.006)	(0.007)	(0.007)
Good	-	-	-	-	0.011*	0.005	0.005
	-	-	-	-	(0.005)	(0.006)	(0.006)
Fair/Poor	-	-	-	-	0.010	-0.002	-0.002
	-	-	-	-	(0.006)	(0.008)	(0.009)
Observations	**39917**	**39917**	**39796**	**39796**	**39796**	**25885**	**25885**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Model specifications:
A Model 1: weekend dummy
B Model 2: Model 1 with day volume
C Model 3: Model 2 with gender, age, ethnicity and Elixhauser comorbidities
D Model 4: Model 3 with time fixed effects, CQC quality metrics, Foundation Trust status and Trust type
E Model 5: Model 4 with interaction terms between weekend with Trust type and CQC quality metrics
F Model 5*: Model 5 with same estimation sample as Model 5**
G Model 5**: Model 5*with Emergency department day volume instead of AMI day volume
Elixhauser Comorbidities: (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (19) Lymphoma; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (27) Blood loss anemia; (28) Deficiency anemia; and (29) Alcohol abuse. Coefficients in the table of 0.01 signify 1%.

Table 47: In-hospital mortality – pneumonia

	In-hospital Mortality						
	A	B	C	D	E	F	G
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5*	Model 5**
Weekend Dummy	0.004	0.004	-0.000	0.001	0.001	0.001	0.008
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.008)
Day volume of admissions (00's)	-	-0.007	-0.001	0.013	0.013	0.018	0.174
	-	(0.007)	(0.007)	(0.008)	(0.008)	(0.010)	(0.139)
Patient Characteristics							
Male	-	-	0.014***	0.014***	0.014***	0.015***	0.015***
	-	-	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age	-	-	0.008***	0.008***	0.008***	0.008***	0.008***
	-	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnicity							
Base category - white							
Mixed	-	-	-0.008	-0.012	-0.012	0.001	0.001
	-	-	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)
Asian	-	-	-0.047***	-0.048***	-0.048***	-0.042**	-0.042**
	-	-	(0.011)	(0.011)	(0.011)	(0.014)	(0.014)
Black	-	-	-0.070***	-0.066**	-0.066**	-0.064*	-0.064*
	-	-	(0.021)	(0.021)	(0.021)	(0.026)	(0.026)
Other	-	-	-0.042*	-0.037*	-0.037*	-0.055**	-0.055**
	-	-	(0.017)	(0.017)	(0.017)	(0.021)	(0.021)
Missing	-	-	0.022**	0.018**	0.018**	0.016*	0.017*
	-	-	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Distance							
Base category - 0-2 miles							
2-4 miles	-	-	0.003	0.004	0.004	0.004	0.004
	-	-	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
4-6 miles	-	-	-0.006	-0.008	-0.008	-0.007	-0.007
	-	-	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)
6-10 miles	-	-	0.002	-0.006	-0.006	-0.003	-0.003
	-	-	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
10-14 miles	-	-	0.007	-0.006	-0.006	-0.010	-0.010
	-	-	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
Over 14 miles	-	-	0.012	0.011	0.011	0.013	0.013
	-	-	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
Missing	-	-	0.025***	0.016*	0.016*	0.022**	0.024**
	-	-	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
Elixhauser comorbidities							
(2)	-	-	0.024**	0.022**	0.022**	0.023*	0.023**
	-	-	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
(3)	-	-	-0.103***	-0.104***	-0.104***	-0.105***	-0.105***
	-	-	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
(4)	-	-	-0.092***	-0.093***	-0.093***	-0.113**	-0.113**
	-	-	(0.028)	(0.028)	(0.028)	(0.034)	(0.034)

(5)	-	-	0.014	0.013	0.013	0.037	0.036
	-	-	(0.025)	(0.025)	(0.025)	(0.028)	(0.028)
(6)	-	-	-0.019	-0.019	-0.019	0.003	0.003
	-	-	(0.033)	(0.032)	(0.032)	(0.038)	(0.038)
(7)	-	-	-0.233***	-0.236***	-0.236***	-0.219***	-0.219***
	-	-	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)
(8)	-	-	0.043	0.032	0.032	0.017	0.018
	-	-	(0.027)	(0.027)	(0.027)	(0.032)	(0.032)
(9)	-	-	-0.115**	-0.119**	-0.120**	-0.092*	-0.092*
	-	-	(0.040)	(0.040)	(0.040)	(0.043)	(0.043)
(10)	-	-	-0.003	-0.007	-0.007	0.005	0.005
	-	-	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
(11)	-	-	-0.073***	-0.073***	-0.073***	-0.080***	-0.080***
	-	-	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
(12)	-	-	-0.169***	-0.173***	-0.173***	-0.170***	-0.170***
	-	-	(0.016)	(0.016)	(0.016)	(0.019)	(0.019)
(13)	-	-	-0.146*	-0.147*	-0.147*	-0.170*	-0.171*
	-	-	(0.058)	(0.058)	(0.058)	(0.066)	(0.066)
(14)	-	-	-0.199***	-0.203***	-0.203***	-0.198***	-0.198***
	-	-	(0.038)	(0.038)	(0.038)	(0.046)	(0.046)
(15)	-	-	-0.005	-0.004	-0.004	0.016	0.016
	-	-	(0.017)	(0.017)	(0.017)	(0.023)	(0.023)
(16)	-	-	0.232***	0.229***	0.229***	0.225***	0.225***
	-	-	(0.025)	(0.025)	(0.025)	(0.029)	(0.029)
(17)	-	-	-0.209*	-0.211	-0.211	-0.378*	-0.378*
	-	-	(0.105)	(0.108)	(0.108)	(0.169)	(0.169)
(19)	-	-	-0.003	-0.007	-0.006	0.004	0.004
	-	-	(0.022)	(0.022)	(0.022)	(0.026)	(0.026)
(20)	-	-	0.108***	0.106***	0.106***	0.114***	0.113***
	-	-	(0.022)	(0.022)	(0.022)	(0.024)	(0.024)
(21)	-	-	0.106***	0.103***	0.103***	0.113***	0.113***
	-	-	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
(22)	-	-	-0.185***	-0.186***	-0.186***	-0.235***	-0.235***
	-	-	(0.038)	(0.038)	(0.038)	(0.054)	(0.054)
(23)	-	-	0.001	-0.001	-0.001	0.036	0.036
	-	-	(0.053)	(0.053)	(0.053)	(0.065)	(0.065)
(24)	-	-	-0.150	-0.147	-0.146	-0.040	-0.040
	-	-	(0.120)	(0.121)	(0.121)	(0.051)	(0.051)
(25)	-	-	-0.067	-0.069	-0.069	-0.002	-0.002
	-	-	(0.046)	(0.046)	(0.046)	(0.011)	(0.011)
(26)	-	-	-0.003	-0.003	-0.003		
	-	-	(0.009)	(0.009)	(0.009)		
(27)	-	-	0.118	0.109	0.112	0.158	0.157
	-	-	(0.164)	(0.160)	(0.162)	(0.182)	(0.182)
(28)	-	-	-0.210***	-0.212***	-0.212***	-0.197***	-0.198***
	-	-	(0.031)	(0.031)	(0.031)	(0.035)	(0.035)
(29)	-	-	0.001	-0.002	-0.002	0.013	0.013

	-	-	(0.022)	(0.021)	(0.021)	(0.028)	(0.028)
(30)	-	-	-0.338*	-0.340*	-0.340*	-0.290	-0.290
	-	-	(0.165)	(0.164)	(0.164)	(0.163)	(0.163)
(31)	-	-	-0.210*	-0.215*	-0.215*	-0.221*	-0.220*
	-	-	(0.084)	(0.084)	(0.084)	(0.099)	(0.099)
(32)	-	-	-0.173***	-0.174***	-0.174***	-0.124*	-0.125*
	-	-	(0.051)	(0.051)	(0.051)	(0.059)	(0.059)
ICD 10 main chapters							
J13	-	-	-0.133***	-0.140***	-0.140***	-0.140***	-0.140***
	-	-	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)
J14	-	-	-0.220***	-0.226***	-0.226***	-0.260***	-0.260***
	-	-	(0.026)	(0.026)	(0.026)	(0.032)	(0.032)
J15	-	-	-0.117***	-0.122***	-0.122***	-0.125***	-0.125***
	-	-	(0.012)	(0.012)	(0.012)	(0.014)	(0.014)
J18	-	-	-0.129***	-0.135***	-0.135***	-0.147***	-0.147***
	-	-	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Trust characteristics							
Small	-	-	-	0.020***	0.020***	0.022***	0.022***
	-	-	-	(0.005)	(0.005)	(0.006)	(0.006)
Medium	-	-	-	-0.027***	-0.027***	-0.029***	-0.029***
	-	-	-	(0.004)	(0.004)	(0.004)	(0.004)
Base Category - large	-	-	-				
Specialist/Teaching	-	-	-	-0.029***	-0.029***	-0.034***	-0.035***
	-	-	-	(0.004)	(0.004)	(0.005)	(0.005)
Foundation Trust	-	-	-	-0.003	-0.003	0.002	0.002
	-	-	-	(0.003)	(0.003)	(0.004)	(0.004)
CQC quality	-	-	-				
Base category - excellent	-	-	-				
Good	-	-	-	-0.015***	-0.015***	-0.015***	-0.015***
	-	-	-	(0.004)	(0.004)	(0.004)	(0.004)
Fair/poor	-	-	-	0.000	0.001	-0.003	-0.003
	-	-	-	(0.005)	(0.005)	(0.007)	(0.007)
Financial quarters							
Base category - quarter 5	-	-	-				
Quarter 6	-	-	-	0.005	0.005	0.004	0.006
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 7	-	-	-	-0.018*	-0.018*	-0.016*	-0.016
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 8	-	-	-	-0.008	-0.008	-0.006	-0.006
	-	-	-	(0.009)	(0.009)	(0.009)	(0.009)
Quarter 9	-	-	-	0.002	0.002	0.003	0.006
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 10	-	-	-	-0.012	-0.012	-0.012	-0.007
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 11	-	-	-	-0.028***	-0.028***	-0.026***	-0.024**
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 12	-	-	-	-0.027**	-0.027**	-0.025**	-0.024**

	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 13	-	-	-	-0.020*	-0.020*	-0.019*	-0.017*
	-	-	-	(0.008)	(0.008)	(0.008)	(0.008)
Quarter 14	-	-	-	-0.033***	-0.033***	-0.033***	-0.028***
	-	-	-	(0.008)	(0.008)	(0.008)	(0.007)
Quarter 15	-	-	-	-0.033***	-0.033***	-	-
	-	-	-	(0.008)	(0.008)	-	-
Quarter 16	-	-	-	-0.043***	-0.043***	-	-
	-	-	-	(0.008)	(0.008)	-	-
Quarter 17	-	-	-	-0.033***	-0.033***	-	-
	-	-	-	(0.008)	(0.008)	-	-
Quarter 18	-	-	-	-0.036***	-0.036***	-	-
	-	-	-	(0.008)	(0.008)	-	-
Interactions with weekend							
Small	-	-	-	-	-0.001	-0.010	-0.010
	-	-	-	-	(0.010)	(0.012)	(0.011)
Medium	-	-	-	-	0.008	0.003	0.002
	-	-	-	-	(0.005)	(0.007)	(0.006)
Large	-	-	-	-	-0.001	-0.003	-0.003
	-	-	-	-	(0.005)	(0.007)	(0.006)
Specialist/Teaching	-	-	-	-	0.009	0.020*	0.019*
	-	-	-	-	(0.007)	(0.009)	(0.008)
Excellent	-	-	-	-	-0.003	-0.005	-0.006
	-	-	-	-	(0.006)	(0.007)	(0.006)
Good	-	-	-	-	0.008	0.007	0.005
	-	-	-	-	(0.004)	(0.005)	(0.005)
Fair/Poor	-	-	-	-	0.004	0.007	0.007
	-	-	-	-	(0.009)	(0.011)	(0.010)
Observations	**85655**	**85655**	**85651**	**85651**	**85651**	**57589**	**57589**

Significance: * p<0.05, ** p<0.01, *** p<0.001. Robust Standard errors displayed in parentheses. Model specifications:
A Model 1: weekend dummy
B Model 2: Model 1 with day volume
C Model 3: Model 2 with gender, age, ethnicity and Elixhauser comorbidities
D Model 4: Model 3 with time fixed effects, CQC quality metrics, Foundation Trust status and Trust type
E Model 5: Model 4 with interaction terms between weekend with Trust type and CQC quality metrics
F Model 5*: Model 5 with same estimation sample as Model 5**
G Model 5**: Model 5*with Emergency department day volume instead of pneumonia day volume
Elixhauser Comorbidities: (2) Congestive heart failure; (3) Cardiac arrhythm; (4) Valvular disease; (5) Pulmonary circulation disorders; (6) Peripheral vascular disorder; (7) Hypertension uncomplicated; (8) Hypertension complicated; (9) Paralysis; (10) Other neurological disorders; (11) Chronic pulmonary disease; (12) Diabetes uncomplicated; (13) Diabetes complicated; (14) Hypothyroidism; (15) Renal failure; (16) Liver disease; (17) Peptic ulcer disease; (19) Lymphoma; (20) Metastatic cancer; (21) Solid tumour without metastasis; (22) Rheumatoid arthritis; (23) Coagulopathy; (24) Obesity; (25) Weight loss; (26) Fluid and electrolyte disorders; (28) Deficiency anemia; (29) Alcohol abuse; (30) Drug abuse; (31) Psychoses; and (32) Depression. Coefficients in the table of 0.01 signify 1%.

Appendix 1

Advancing Quality data recording for Quality Measures Reporter

During my 2014, I thought it would be advantageous to my research if I visited a Trust which was involved in Advancing Quality. My aim was to speak to clinicians regarding how the decision process was made regarding the provision of care to patients and how the data was collected and stored. I was granted access to an Emergency Assessment Unit (EAU) where three junior doctors and one specialty registrar were available from the cardiology specialty for advice and shadowing.

Data entered into the systems were from hand written medical records from both paramedics for emergency conditions and clinicians. Most of the medical records regarding Advancing Quality were mostly in a tick box format, where clinicians tick which process measures of care have been performed. However, any exclusion reasons were hand written in medical records. The data entry from the medical records was conducted by a data team. The clinicians did not know who the data team are. The data entry workers were occasionally students on work experience or people with no medical training. Entering the correct data from hand written records will be difficult for people of no experience to understand due to high usage of abbreviations. Therefore, there were possibilities that coding for exclusion reasons may be ignored or entered incorrectly.

The Audit Commission rated the documentation of process measures of care as generally excellent, especially when the process measures of care were drugs which have been prescribed. Occasionally, clinicians may not document process measures of care such as smoking cessation advice as this process measure is not collected through a check box, and was hand written. Not documenting process measures may be due to three reasons. Firstly, documentation of smoking cessation advice will not directly impact a patient's survival in the emergency unit; therefore the importance of documenting the advice may be of lower importance than drugs the patients received. The second reason was occasionally, urgent attention was needed for other patients, where a clinician did not have the time to complete the medical record. The third reason was that the clinicians may not know the reason for documenting smoking cessation advice.

All three junior clinicians I spoke to had limited knowledge regarding Advancing Quality. The junior clinicians had no knowledge on details of Advancing Quality such as which process measure of care is included in the quality incentive scheme and which were not. Instead, clinicians followed clinical guidelines from National Institute for Health and Care Excellence, of which included the process measures of care from Advancing Quality. The reasoning behind the lack of knowledge on Advancing Quality is that the junior clinicians have rotations between specialties within a hospital (McDonald et al. 2014). A typical placement in cardiology lasts four months. The junior clinicians were not educated about Advancing Quality directly.

The specialty registrar was very knowledgeable on Advancing Quality. This included knowledge on which process measures of care which were incentivised and in some cases, the knowledge of the research on the process measures of care. There may therefore be trickle-down effect of knowledge on Advancing Quality, not regarding which process measures of care were in Advancing Quality, but what should be included in the patient's care pathway for junior doctors to follow, and what information to include on the medical records.

Appendix 2

Stata commands for marginal effects

To calculate marginal effects in Stata, I used the *margins* command. To calculate marginal effects after a multinomial and sequential logit regression, different option specifications were required.

For marginal effects after a multinomial logit, I used the following specification:

margins, dydx(*) pr(outcome(#))

The option *dydx(*)* specifies that I would like to calculate marginal effects from all explanatory variables in my regressions. *Pr(outcome(#))* specifies which regression output I would like to calculate marginal effects for. The output from multinomial logit for my specification has three regression results: excluded and died vs pass; excluded alive vs pass; and fail vs pass. *Pr(outcome(1))* calculates marginal effects for excluded and died vs pass. *Pr(outcome(2))* calculates marginal effects for excluded and alive vs pass. *Pr(outcome(3))* calculates marginal effects fail vs pass. For marginal effects after a sequential logit, I use the following specification:

margins, dydx(*) predict(trpr transition(2) choice(2))

Within this command I specified *trpr transition(2)*. My specification of the sequential logit has two stages: in the first transition the choice is excluded and died and not excluded and died and the second transition compares three choices. The three choices are: excluded and alive (0); fail (1) and pass (2). Therefore *predict(trpr transition(2) choice(2))* calculates the probability of being in transition two and being given a process measure of care. In other words it is the probability of being given a process measure conditional on not being excluded through death.