Substance misuse related hospital admissions, costs and treatment outcomes: econometric analysis of administrative data for England

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2019

Thomas Mason

School of Health Sciences

Contents

List of Tables

List of Figures

List of Abbreviations

A&E – Accident and Emergency
AIC – Akaike Information Criteria
AIDS - Acquired immune deficiency syndrome
APC – Admitted patient care
APE – Average partial effect
AQ – Advancing Quality
BIC – Bayesian Information Criteria
CCG – Clinical Commissioning Group
CDU – Chronic drug users
CI – Confidence interval
CJS – Criminal Justice System
COI – Cost of illness
CQUIN - Commissioning for Quality and Innovation
DALY – Disability adjusted life year
DAT – Drug Action Team
DH – Department of Health
DiD – Difference-in-differences
DIP – Drug interventions programme
DTORS - Drug Treatment Outcomes Research Study
FE – Fixed effects
GOR – Government Office Region
HES – Hospital Episode Statistics
HIV - Human immunodeficiency viruses
HRG – Healthcare resource group
ICD-10 - International Classification of Diseases, 10th revision
ICD-9 - International Classification of Diseases, 9th revision
IDU – Injecting drug users
IMD – Indices of multiple deprivation
IRR – Incidence rate ratio
LDV – Lagged dependent variable
LOS – Length of stay
LQ – Lower quartile
ME – Marginal effect
MFF – Market forces factor
MLE – Maximum likelihood estimation
MSDS – Maternity Services Data Set
NB – Negative binomial
NDEC – National Drug Evidence Centre
NDTMS – National Drug Treatment Monitoring System
NHS – National Health Service

NICE – National Institute for Clinical Excellence
NTA – National Treatment Agency
NTPS – National Tariff Payment System
OAP – Opioid agonist pharmacotherapy
OLS – Ordinary least squares
ONS- Office for National Statistics
OR – Odds ratio
P4P – Pay for performance
PAF – Population attributable fraction
PbR – Payment by Results
PCT – Primary Care Trust
PHE – Public Health England
PSSRU – Personal Social Services Research Unit
QOF – Quality and Outcomes Framework
RE – Random effects
S.D. – Standard deviation
SUS – Secondary uses service
TOP – Treatment outcomes profile
UK – United Kingdom
UQ – Upper quartile

Abstract

Substance misuse represents a global public health concern. The harms caused by substance use remain considerable internationally: 0.6 percent of the global adult population were estimated to have a substance use disorder in 2015, with 28 million disability-adjusted life years lost as a result of substance misuse and 17 million health years of life lost as a result of substance use disorders. However, there is relatively limited evidence on the economic impact of substance misuse on health care for England. An important pillar of recent substance misuse policy focusing on 'recovery' was a pay-for-performance (P4P) scheme which linked the income of providers of addiction treatment services to treatment outcomes. There is limited evidence on the effects of this scheme on in-treatment outcomes and wider related health care use and costs.

This thesis applies econometric methods to two large administrative datasets for England between 2009-10 and 2015-16 in four main chapters. The first explores how the costs of substance misuse related hospital admissions have changed over the period 2009-10 to 2015-16 using Hospital Episode Statistics (HES), and how different costing approaches yield different estimates. Costs are shown to be overestimated by top-down approaches that apply average costs to specific diseases, and that estimates are highly sensitive to the use of primary and/or secondary diagnosis codes of substance misuse. The second uses an event study design to examine how hospital activity and related costs change around the time of the first occurrence. Annual costs per person are on average £2,953.65 higher in the event year than six years before, and remain £687.02 higher five years after.

The third uses a difference-in-differences design to evaluate the impact of the introduction of P4P for providers of addiction treatment services on treatment outcomes using data from the National Drug Treatment Monitoring System (NDTMS) for 2010-11 to 2012-13. P4P reduced the probability of individuals completing substance misuse treatment successfully by 1.3 percentage points and increased the probability of individuals declining to continue with treatment by 0.9 percentage points.

The final chapter considers the schemes impact on hospital admissions and costs, finding that individuals' hospital admissions were 1.507 times higher in P4P areas after its introduction compared with non-intervention areas. The number and cost of

substance misuse related hospital admissions increased substantially between 2009-10 and 2015-16. Furthermore, individuals' hospital activity and costs increases prior to an admission diagnosed as being related to substance misuse - and is/are persistently higher after the event. Introduction of P4P for providers of addiction treatment services had negative effects on outcomes in the addiction treatment population and led to increased hospital admissions in P4P areas for a wider population with a record of substance misuse.

Declaration

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

Chapter 4 of this thesis was published as a peer-reviewed research report in the journal 'Addiction' on 7th June 2015, first-authored by the author of this PhD thesis and co-authored by several collaborators. The analysis and interpretation of the data and the design of the study in this research report was undertaken by Thomas Mason, Andrew Jones, Matthew Sutton and William Whittaker. The paper was drafted in full by Thomas Mason. All other authors contributed to and approved the final version of the paper. The paper is available online at: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/add.12920`

The above research report was written as part of a wider study funded by the Department of Health Policy Research Programme. The study published its results in a wider report which is publicly available. This thesis cites aspects of this report but does not duplicate its content.

Aside from Chapter 4 which was published in a peer-reviewed journal, the analysis and writing of all content herein was undertaken by Thomas Mason. Andrew Jones, Matthew Sutton and William Whittaker provided supervision on all chapters.

Copyright Statement

Acknowledgements

This PhD has damaged me irreparably and it is only fair that others take their share of the blame.

I would first like to thank Andy, Matt and Will for their supervision over the one thousand, four hundred and twenty-seven years of this part-time PhD. Their academic support and advice have been essential throughout, they have supported me with good humour and genuine compassion, and I regard them each as a mentor and friend.

I am grateful to the University of Manchester for its support for this PhD. My appreciation also extends first to colleagues with whom I have worked as a researcher at the University - who have become close personal friends; and second to those who have helped my career and personal development – especially Matt and Will. Thank you both, I couldn't and wouldn't have done it without you. Having said that: you have each taken years off my life and I will never forget it.

I would like to grudgingly thank my 'so-called' friends who have at different stages supported me: housemates, teammates, bandmates and cellmates.

A big thank you to Charlie for her patience, support and insults over the period when the stress has been at its most concentrated.

And, on the subject of stress: I want to thank my family, who prepared me for the last few years as if by design. Thanks to my brother, aunts and uncles, cousins, grandparents; thank you to those who have supported and encouraged me in my studies – especially Rick, Margaret and Tony, and Helen and Tony. Finally, thank you to Mum and Dad for everything – for the sacrifices along the way, for supporting me unconditionally in everything I have done, and for being there whenever I have needed it.

1 Introduction

1.1 Background

Drugs misuse represents an international public health concern with a considerable burden of economic (Godfrey et al. 2004; French et al. 1996; Healey et al 1998; Parthasarathy & Weisner 2005; Healey et al. 2003), physical (Rippeth 2008; Reece 2007; Crowe et al. 2000) and social consequences (Galea & Vlahov 2002; March et al. 2006; Garcia-Altes et al. 2002; Roxburgh et al. 2005).

The harms caused by drug use remain significant: 0.6 percent of the global adult population were estimated to have a drug use disorder in 2015, with 28 million disability-adjusted life years (DALYs) lost as a result of drug-use and 17 million health years of life lost as a result of drug use disorders (UNODC 2017). Healthcare costs, along with crime costs and productivity losses, are consistently identified in international research as one of the principal contributors to the costs of substance misuse (ONDCP 2004; Harwood et al. 1999; Baumberg 2006; Wickizer 2013; McCollister et al. 2003; Olsson & Fridell 2015).

These recognised global developments apply to the context of the United Kingdom (UK). There were 2,593 registered drug related deaths in 2016, eighty percent of which were due to accidental poisoning (ONS 2017). This represents a "dramatic and tragic increase" in drug misuse deaths since 2012 (HMG 2017: 4). In 2015-16 in England, there were also 82,135 hospital admissions with a primary or secondary diagnosis of drug-related mental health and behavioural disorders in 2015-16; over double the level in 2006-07 (NHS Digital 2018). This represents 0.5% of all hospital admissions. However, the costs of these admissions are unknown – the most recent efforts to estimate these costs were for 2011-12 (Home Office 2013). Further complication exists insofar as there are various approaches that can be used both in defining activity that is related to substance misuse, and in then costing this activity once it is defined (Tarricone 2006; Jo 2014).

A range of approaches have been adopted internationally over recent decades by policymakers in an attempt to reduce the use of internationally controlled substances and to treat addiction to drugs (IDPC 2018). Despite this, the evidence suggests that

the production and consumption of illegal substances continues to grow apace – as does the consequent harms associated with drugs misuse including overdoses and transmission of diseases such as HIV and hepatitis (IDPC 2018: i). Harm reduction has increasingly represented the primary basis of structured treatment for individuals with dependence on illicit drugs (IDPC 2018). Harm reduction is defined as (IDPC 2018: 37):

> "policies, programmes and practices that aim primarily to reduce the negative health, social and economic risks and harms associated with drug use without necessarily reducing drug consumption"

Examples of harm reduction policies include drug consumption rooms, needle exchanges, housing and employment initiatives that do not require abstinence, and substitute prescribing. The effectiveness of harm reduction policies in protecting the health of substance misusing individuals is well recognised (WHO 2013). For England, the National Institute for Clinical Excellence (NICE) recommends harm reduction treatments such as opioid agonist pharmacotherapy (OAP) as the primary maintenance treatment for opioid dependence (NICE 2007).

1.2 Substance misuse and its sequalae

The misuse of illicit drugs has a range of physical, social, psychological and economic impacts that are well established across a range of settings and over time. These impacts vary across substances, longevity of use and have interactions with the specific characteristics of individual users.

In the UK setting, the most commonly used illicit drugs are estimated to be cannabis, cocaine and ecstasy (Roe & Mann 2006). Opioid misuse occurs on a smaller scale but is associated with much higher rates of harm than other substances. Opioids are defined as a class of psychoactive substances that are derived from poppy plants (including opium, morphine and codeine), as well as semi-synthetic forms (including heroin) and synthetic compounds (such as methadone) which have similar properties (WHO 2006). Drugs such as cocaine are classified as 'stimulants', which refer broadly to any substance that activates, increases or enhances neural activity (WHO 2006). Whilst cocaine is the most commonly misused illicit stimulant in the UK (Roe & Mann 2006), stimulants also refer to crack cocaine and amphetamines which are a

group of synthetic substances with unique chemical structures but broadly similar effects to cocaine. Cocaine itself is extracted from the leaf of coca plants and is typically sniffed in powder form. Cannabis is an umbrella term denoting the various preparations of the cannabis sativa plant, including cannabis leaves (the most common), hashish resin and cannabis oils (NICE 2008). It has psychoactive properties and represents the most commonly used illicit drug in the UK (NICE 2008).

Substance misuse can be defined in different ways. The WHO (2006) considers it as being the use of a substance for a purpose that is not consistent with legal of medical guidelines. It typically has a negative impact on health and/or functioning and may take the form of drug dependence – or be part of a wider spectrum of harmful behaviours. Dependence in this context is defined as (WHO 2006):

"a strong desire or sense of compulsion to take a substance, a difficulty in controlling its use, the presence of a physiological withdrawal state, tolerance of the use of the drug, neglect of alternative pleasures and interests and persistent use of the drug, despite harm to oneself and others".

Diagnosis of drug dependence is the clearest in the case of opioids. The WHO (2006) provides the following summary:

"opioid dependence develops after a period of regular use of opioids, with the time required varying according to the quantity, frequency and route of administration, as well as factors of individual vulnerability and the context in which drug use occurs. Opioid dependence is not just a heavy use of the drug but a complex health connotation that has social, psychological and biological determinants and consequences, including changes in the brain. It is not a weakness of character or will."

The health consequences of problematic substance misuse vary across different types of substances and on the specifics of each individual case, but there are some general patterns that are associated with types of substances, and then with substance misuse more broadly. Across all illicit drugs, use is most prevalent for younger adults aged between 20 and 29 (Li & Gunja 2013) and related hospitalisation rates are highest for those aged between 20 and 39. Smoking is very common in substance misusers, and its sequelae in an ageing cohort is often compounded by poor living conditions and in some cases respiratory problems are exacerbated by frequent inhalation of drugs such as cannabis and crack cocaine (Cornford & Close 2016). Mental health care needs in individuals who use illicit drugs are pronounced, with high rates of prevalence of psychotic illness, personality disorders and more common mental health problems such as anxiety and depression (Cornford & Close 2016).

Opiate use is recognised as having the highest mortality rate amongst illicit drugs, despite having a relatively lower prevalence in terms of use compared with cannabis and cocaine. Deaths and hospitalisations are typically due to overdoses, diseases associated with use (for example HIV, infection, viral hepatitis) and suicide (Darke et al. 2011).

Some of the health problems associated with the opiate heroin are essentially the health effects of injecting drug use more generally. Injecting drug users have a high prevalence of skin and groin infections, and of soft-tissue infections more generally (Cornford & Close 2016). These include common but serious infections such as subacute bacterial endocarditis (SABE) and osteomyelitis; in addition to rarer infections such as tetanus and necrotising fasciitis (Cornford & Close 2016). Deep vein thrombosis is more common than in the general population, and injecting drug users have high and complex pain management needs.

In the case of the amphetamines and cocaine, excessive use of these substances is typically characterised by cardiovascular and autonomic hyperstimulation; in addition to neuropsychiatric disturbance such as agitation, delirium and hallucinations (Li & Gunja 2013). Cocaine and amphetamine use complications include acute myocardial infarction; and chronic use of these stimulants is a strong risk factor for early onset coronary syndromes in younger adults and teenagers (Vasica & Tennant 2002).

For cannabis, acute hospital admissions typically occur in cannabis naïve patients who report symptoms such as anxiety, panic, depersonalisation, altered mood, impaired memory and difficulty in concentrating (Ashton 2001). Despite its relatively wide prevalence in terms of use, there is limited conclusive evidence relating to the other health impacts of cannabis: there has been no robust causal relationship proven in multiple settings between psychological (such as psychosis risk) and social (such as educational) outcomes (Macleod et al. 2004). Some studies have tentatively demonstrated the negative impacts of cannabis use in the developing brains of young adults and teenagers (Fergusson et al. 2015). Links between cannabis smoking and lung cancer are also not well established. Cannabis is typically smoked with tobacco, which is the main risk factor for lung cancer – and lung cancer remains the leading cause of cancer death worldwide (Zhang et al. 2015). Cannabis smoke has some carcinogens in common with tobacco smoke, and previous studies have demonstrated precancerous histological and molecular abnormalities in the respiratory tracts of cannabis smokers (Fligiel et al. 1997). However, epidemiological studies have been limited, conducted for smaller sample sizes and reported conflicting results (Aldington et al. 2008; Hashibe et al. 2006; Berthiller et al. 2008). Pooled analysis across

studies provides little evidence for an increased risk of lung cancer among habitual or long-term cannabis smokers (Zhang et al. 2015).

The effects of ecstasy use depend partly on whether ecstasy tablets (which are mixed or 'cut' with other substances such as amphetamines) or the purer form "MDMA" (methylenedioxymethamphetamine) are used. Individuals who are admitted to hospital as a result of using MDMA may present with hypothermia, either through the direct effects of its use or often due to the environment in which it is taken (for example at nightclubs and 'rave' parties where dehydration can be a problem) (Hamid et al. 2005). Admission to hospital as a result of MDMA poisoning has a range of presenting health effects such as bruxism, palpitations, hypertension and diaphoresis (in milder cases); to an amphetamine-like sympathomimetic toxidrome (in more severe cases) (Li & Gunja 2013).

Drugs misuse more generally is often characterised using multiple drugs; and is also associated with a range of economic and social consequences including higher unemployment, higher rates of homelessness, unstable family lives and social situations, higher rates of acquisitive crime, and recidivism. As such, the consequences of drugs misuse involve complex interactions between a range of direct and indirect effects; and physical and environmental effects. As such, the policy response spans across a range of government services including health services (including primary care, secondary care, health protection); social services (including education, housing and community services) and the criminal justice system.

1.3 Substance Misuse Treatment in England: an overview

In England, individuals who engage in structured treatment for substance misuse can enter via a number of routes. These are recorded in the National Drug Treatment Monitoring System (NDTMS), which collects data on those receiving treatment from roughly 900 sites covering all local authority areas in England (PHE 2016). In NDTMS data, individuals' route into treatment can be understood via the 'referral source'. In 2015-16, 138,081 new presentations to treatment were referred from the following sources (PHE 2016)

i. "Self, family and friends" (51%);

ii. "Health services and social care" (20%);

iii. “Criminal justice” (16%);

iv. “Substance misuse service” (8%).

After referral and entry into treatment, individuals are assessed in terms of need and risk at the onset of treatment in line with national clinical guidelines (DH 2017). This initial ‘drugs misuse assessment’ is intensive and is illustrative of the broad, wide ranging focus of treatment. It includes:

i. identification of and emergency or acute problems;

ii. confirming whether the individual is using any psychoactive substances (based on history, testing and clinical records);

iii. identifying the degree and nature of problematic use and dependence;

iv. identifying physical and mental health problems, including: assessment of injecting-related disease; sexual health; personality disorders; self-harm; history of abuse and/or trauma; depression; anxiety and severe psychiatric comorbidity;

v. identifying social problems, including: problems in personal relationships and social integration including domestic violence, housing and living arrangements, benefits and financial matters; childcare; and criminal matters;

vi. determination of the need for substitute medication or other prescribing for dependence.

Once an individual’s needs are assessed, treatment options are discussed, and a treatment plan is agreed in liaison between users and providers. These treatments are classified in monitoring data as pharmacological (substitute prescribing) or psychosocial (such as support groups; detoxification; cognitive behavioural therapy). These treatments/interventions are provided in a range of settings – in 2015-16, 265,490 individuals were treated in the following settings (note that individuals can be treated in multiple settings during their treatment) (PHE 2016):

i. community based and specialist outpatient drug services (97%);

ii. inpatient units (6%);

iii. primary care (11%);

iv. residential rehabilitation centres (1%).

Once in treatment, information on individuals are recorded in NDTMS. These details include: patients' demographic and socioeconomic characteristics; personal circumstances; details of the treatments delivered to them; and associated outcomes. In the case of the latter, the outcomes recorded include:

i. successful completion of treatment abstinent from drugs;

ii. change in use of drugs at six-monthly follow-up;

iii. receipt of smoking cessation interventions amongst smokers;

iv. changes in employment status at six-monthly follow-up;

v. changes in education level at six-monthly follow-up;

vi. changes in housing need at six-monthly follow up;

vii. withdrawing from treatment (including reason).

1.4 Substance Misuse Policy for England: 'recovery' and P4P

Developments in government policy for the UK and, more specifically, England have recognised the importance of harm reduction in recent decades. The UK government's 2004 Drug Strategy aimed primarily to reduce the harms caused by substance misuse by increasing the numbers of individuals entering structured treatment for addiction to illicit drugs through the criminal justice system and more broadly - in addition to focusing efforts on prevention of use of 'Class A' drugs in persons aged under 25 (Monaghan 2012). More recently however, there has been a prominent change in rhetoric emphasising the goal of abstinence, a term which has been replaced with a less precise term 'recovery'. This has been interpreted by some as a signal of the desire for structured treatment to move away from maintenance policies such as substitute prescribing (Monaghan 2012).

The 2010 Drug Strategy set out to change the delivery of structured treatment. It claimed that the emphasis of previous Drug Strategies had been on harm reduction, i.e. promoting substance misusers to enter and then remain in treatment. For example (HMG 2010: 18):

> "The investment made in the drug treatment system over the last decade has built capacity and enabled people to access treatment for a sufficient period of time to bring about substantial health gains. We now need to make the same progress in treating those with more severe alcohol dependence and to become much more ambitious for individuals to leave treatment free of their drug or alcohol dependence so they can recover fully.
>
> We will create a recovery system that focuses not only on getting people into treatment and meeting process-driven targets, but getting them into full recovery and off drugs and alcohol for good. It is only through this permanent change that individuals will cease offending, stop harming themselves and their communities and successfully contribute to society."

As such, the new direction of policy was to accept the role of substitute prescribing in the treatment of opioid dependence, but change its priority relative to the goal of drug 'recovery' (Disley et al. 2016; NTA 2012). For example (HMG 2010: 18):

> "However, for too many people [sic] currently on a substitute prescription, what should be the first step on the journey to recovery risks ending there. This must change. We will ensure that all those on a substitute prescription engage in recovery activities and build upon the 15,000 heroin and crack cocaine users who successfully leave treatment every year free of their drug(s) of dependence."

A key pillar of this strategy was a pay-for-performance (P4P) pilot policy for providers of substance misuse treatment services, the so-called 'Payment by results for drugs recovery pilot'. The purpose of this pilot was the explore how reform of payment mechanisms could be used to refocus treatment towards the 'recovery outcomes' defined by policymakers (HMG 2010: 24):

> "We are keen to explore how we can incentivise the system to deliver on recovery outcomes. We will therefore introduce six [sic] pilots to explore how Payment by Results (PBR) can work for drugs recovery for adults, which will also provide evidence on affordability and value for money as part of the evaluation of these pilots."

This approach is described in the 2010 Strategy being geared to "test approaches where money will follow success" (HMG 2010: 24).

The P4P scheme was introduced in eight geographical areas in England in April 2012, and differentially emphasised the desirability of service users recovering from their dependency on drugs and successfully completing treatment abstinent from drugs by linkage of providers' incomes to 'recovery-focused' outcomes (HMG 2010). As such, this scheme represented a 'hard' mechanism which promoted providers' increased emphasis on recovery and abstinence insofar as it refocused treatment towards defined outcomes, such as abstinent completion of treatment, by making providers' financially dependent on their achievement.

Previously, providers were partly funded through allocations from the public health budget to Primary Care Trusts (PCTs) from central government. These allocations were based on formulae which reflected the sizes of areas' treatment populations and factors such as deprivation. The remainder of funding was distributed from the Department of Health (DH) to local commissioners, 'Drug Action Teams' (DATs), who then granted funding to local providers. DATs are the government agencies responsible for the commissioning of drugs misuse treatment services. DATs were established in 1995 under the government's national drug strategy "Tackling Drugs Together: a strategy for England 1995-1998" (MacGregor & Thickett 2011). The principles underpinning their creation were based on (MacGregor & Thickett 2011: 480):

> "importance of multi-agency working; individual members of a co-ordinating group to be sufficiently senior in their own organisation to deliver policy and resource decisions; clear accountability by one local agency to central Government for the formation of co-ordination arrangements; all parts of England to be covered by local co-ordination arrangements with some flexibility on the selection of suitable geographical boundaries; and minimum disruption to what is already working well."

More generally, the DAT's were setup to perform a strategic role: to assess the nature and scale of local drug problems; and to ensure that the relevant local agencies were co-ordinating to ensure the principles of the drug strategy were acted upon (Higate et al. 2006). At this time, DATs included representation from a range of local government services, including: health, police, probation, education, housing,

prison, and customs and excise (Higate et al. 2006). In 1995, their geography was coterminous with health authorities. However, changes in the organisation of the NHS in the late 1990's led to changes in responsibilities for planning and commissioning drug services and DATs became coterminous with local authorities – an arrangement that remains the case. The NHS Act 1999 then led to the pooling of budgets from across several government agencies, notably the Home Office and the Department of Health (Higate et al. 2006). DATs then allocate these budgets to local providers. Generally, providers' budgets were previously calculated based on the predicted size of their treatment populations and the average price per individual in treatment (Maynard et al. 2011).

Under the P4P pilot scheme, provider reimbursement was instead linked to performance indicators by DATs. Commissioners were given limited scope for local design provided their payment scheme adhered to a 'national outcomes framework'. Local commissioners also had some flexibility regarding the payment weights attached to indicators in the national framework (although 'successful completion' and 'non-representation' outcomes were weighted to always comprise a large share of P4P income). Providers in the P4P pilot areas were paid for outcomes in three nationally agreed domains: (i) progression towards abstinence from presenting drug(s) of dependence; (ii) reduction in offending; and (iii) improved health and wellbeing (Appendix Tables A1.1-A1.4.) There were two overarching objectives of this P4P scheme. The first was to focus providers on achieving 'recovery' for individuals and not simply retention in treatment. The second objective was to deliver greater efficiency and value for money (HMG 2010: 24):

> "Fundamental to this approach will be the reduction in bureaucracy, moving away from unnecessary repeated assessments towards enhanced continuity of support. All areas will be encouraged to set up a single assessment and referral system. This single, unified system would redirect available funding away from bureaucratic processing and into the recovery support that individuals need. This approach, which will be a mandatory element of the pilots, will ensure that local partners provide more cost effective services delivering greater efficiency."

Impacts on both the effectiveness and efficiency of structured treatment require not only assessment of the treatment and health outcomes of individuals in addiction treatment, but also in the wider substance misusing population. This wider population are observed across a range of government services such as welfare services, the

criminal justice system and hospitals. The latter uniquely allows for consideration of not only the economic impacts of policy on government expenditure, but also on the health outcomes of the substance misusing population.

Evaluation of this scheme is aided by its design insofar as it was piloted in some areas and not others, allowing for robust assessment of its impacts on the effectiveness and efficiency of treatment. However, consideration of treatment efficiency requires some assessment of the economic impacts in the wider substance misusing population – for example consideration of the impacts on costs to the criminal justice system, or on hospitals. Valuing the costs of government services is a complex and technical exercise and can entail a range of approaches that result in different findings. Complementing assessment of the impacts of a policy such as this scheme on outcomes with the economic impacts is therefore a challenging undertaking, but it is nonetheless a desirable feature of any policy evaluation (Sutton et al. 2018). Understanding the impact of this particular scheme can contribute to the general evidence base on P4P in the setting of provision of public services and especially in health care.

1.5 Reform of payment mechanisms for providers of public services and the evidence on P4P

The use of P4P in health care and other public service settings reflects longer term developments in the approaches used for paying providers - policymakers have adopted more 'active' forms of reimbursement. These developments are exemplified by repeated reforms to the way that hospitals are paid in England.

The history of payment reforms for National Health Service (NHS) hospitals dates back to the establishment of the 'purchaser provider split' in the NHS, which led the relationship between these distinct entities to be governed by contracts. The first were block contracts, for which providers received a given amount of money to provide a given service irrespective of the amount of activity provided. Block contracts were subsequently replaced with 'activity-based financing' which was intended to create a link between volume and revenue. This approach to reimbursement was complemented with other policies aimed at linking payment with the quality of care provided. For example, from April 2011 until April 2012, hospitals in England were not reimbursed for emergency readmissions within thirty days of an elective admission. Finally, schemes such as the Advancing Quality (AQ) initiative were introduced which paid providers based on whether they performed particular activities for in-

dividuals admitted to hospital with health conditions for which there were clinically established best practice pathways that applied to a majority of patients – such as giving Stroke patients aspirin upon their arrival in hospital (Mason et al. 2017).

The evidence base on P4P schemes adopted in health care has grown considerably over the past 15 years, but there is no consensus as to their effects (Scott et al. 2011; Van Herck et al. 2011; Alhamsan et al. 2010; Eijkenaar et al. 2013; Mendelson et al. 2017; Scott et al. 2016). Studies evaluating these schemes generally consider their impact on indicators of providers' activities: typically, whether a provider performed a particular task which is considered to be clearly beneficial in achieving the objective for the service in question. Schemes have tended to link payment to these processes, because linking to outcomes can be more problematic (particularly for health outcomes which can be rare) (see for example Gravelle et al. 2010; Crawley et al. 2009; Karve et al. 2008; Sutton et al. 2012). Overall, there is no evidence that P4P schemes have a consistent effect on intermediate health outcomes and there is insufficient evidence to characterise the impact on the health outcomes of patients (Mendelson et al. 2017).

Most studies have examined P4P schemes in the setting of either primary or secondary health care. There are relatively few examples of the use of P4P for the specific case of providers of structured treatment for substance misusers. Studies evaluating two schemes which incorporated an element of financial incentives in the United States tentatively suggest that, at least by some measurements, their introduction led to improvements in treatment outcomes (McLellan et al. 2008; Shen 2003). The studies are however limited in their inference by design problems, which includes the lack of a control group, missing baseline measurements and reliance of self-reports of abstinence. More generally, the P4P evidence base has been characterised by studies that have been unable to adopt a robust design in evaluating their effects – there are few methodologically rigorous studies (Mendelson et al. 2017). The various reviews of this literature have consistently noted weaknesses to this effect, and have consequently endorsed the use of patient-level data that enable confounding for patient-level characteristics; and of quasi-experimental methods for assessment of P4P schemes that have been piloted in some areas and not others (Flodgren et al 2011; Alhamsan et al 2010; Mendelson et al 2017). Other recommendations for assessing P4P schemes includes consideration of the economic impacts of schemes; examination of possible adverse or unintended consequences; consideration of schemes' impacts on outcomes rather than processes; and use of econometric techniques to account for possible unobserved heterogeneity and the potential confounding of changes in time

trends (Flodgren et al 2010; Scott et al. 2011).

1.6 Research objectives

This thesis comprises four empirical chapters. The final two chapters comprise an assessment of the impacts of the aforementioned P4P scheme on both the outcomes of the population in treatment for their substance misuse, and of hospital admissions and costs for a wider substance misusing population admitted to hospital. Assessment of the latter allows for analysis of a less restricted population, and for measurement of both health and economic impacts. However, a detailed assessment of economic impact requires a detailed exploration of approaches that can be used to estimate the costs of substance misuse related hospital admissions. The scale of these costs has not been estimated since 2011-12, and there are different approaches that can be used to identify activity that is considered as being related to substance misuse. Additionally, there are different approaches that can be used to cost this activity once it has been identified.

Chapter 3 of this thesis therefore explores the approaches that can be used for defining and costing relevant activity. Chapter 4 then explores a relatively novel application of an econometric method which uses the timing of admission to hospital in order to identify the impact of substance misuse on patterns of activity and related costs. The fifth chapter uses difference-in-differences methodology to understand how the P4P scheme impacted on outcomes for the population in structured treatment for substance misuse. Chapter 6 complements this by extending assessment to the scheme's impact on hospital activity and costs in a population with a diagnosis related to substance misuse.

This thesis makes several contributions. First, it provides more recent evidence on the costs of substance misuse related hospital admissions than are currently available for England. In doing so, it explores different approaches for identifying and costing substance misuse related activity. It then considers the limitations of approaches for identifying related activity and uses a relatively novel application of an econometric method for identifying patterns of cost and activity related to substance misuse. This application also explores a range of approaches that can be used to estimate regression models in the case of overdispersed outcomes (such as data on health care activity and costs).

These contributions are followed by two empirical chapters which contribute to the

P4P evidence base. These chapters represent relatively novel, wide-ranging and robust evaluation of an atypical P4P scheme: where payments have been linked to genuine treatment outcomes and not to processes; where piloting of the scheme allows for a quasi-experimental design to be adopted; and in the setting of providers of substance misuse treatment. These chapters complement robust assessment of the policy's impact on treatment outcomes with examination of possible unintended consequences and economic impacts on the hospitals' costs. Both chapters use large, administrative patient-level data to allow for assessment of different subsets of the wider substance misusing population.

The remainder of this thesis is structured as follows: first, a data and methods chapter explains the analysis approach for this thesis in detail. Then, each of the four empirical chapters are presented in turn. Each empirical chapter introduces the background to and motivation for the specific research questions, defines them, and results are followed by a discussion of the findings. These empirical chapters are then followed by a discussion of the overall thesis findings.

2 Data sources, cleaning, structuring and methodology

This chapter details the data sources, the approach taken to data cleaning, the procedures used in structuring the data and the methodology applied in the subsequent empirical chapters. There is duplication and commonality across the empirical chapters of this thesis in some aspects of the data and methods. Therefore, where particular aspects have previously been explained, reference is made to this fact so as to avoid repetition throughout this Chapter.

The subsequent empirical chapters (Chapters 3 to 6) then outline the introduction in terms of the individual study background and motivation, report the results of the analysis and provide a discussion of the results. The remainder of this chapter briefly introduces the research objective for each empirical study, and then describes in detail the data and methods for each empirical chapter in chronological order.

2.1 Data and methods for Chapter 3: "The costs of substance misuse related hospital admissions in England 2009-10 to 2015-16: an exploration using Hospital Episode Statistics"

This study estimates substance misuse related hospital costs for England for 2009-10 to 2015-16 and explores how different approaches to applying costs and defining related activity impact on these estimates.

2.1.1 Data

I use activity data from the 'admitted patient care' (APC) data files in Hospital Episode Statistics (HES) for the financial years 2009-10 to 2015-16. HES are episode-level data of hospital care covering all NHS Clinical Commissioning Groups (CCGs) in England. This includes: privately paid patients treated in NHS hospitals; patients resident outside of England; and care delivered by non-NHS providers that is funded by the NHS (NHS Digital 2016). HES data provide clinical information (such as

diagnoses and operations), patient information (such as age and sex), administrative information (such as admission date and discharge method) and geographical information such as patient area of residence (NHS Digital 2016).

The APC data are a subset of the HES data, and they contain what are known as 'Finished Consultant Episodes'. Each episode relates to a period of care for a patient under a specified consultant delivered in a single hospital (NHS Digital 2016). The number of rows in APC data files therefore gives the number of 'finished' consultant episodes of care for admitted patients rather than the number of patients treated (or the number of admissions). Episodes are labelled as finished according to the financial year in which they end.

Patients can have multiple episodes during a given stay in hospital. Episodes are uniquely identified within 'spells' – defined as the period from admission to discharge (DH 2012). Spells are the units for which hospitals are paid for their activities. The term spell is used interchangeably with the (better understood) term admission herein. Admissions are defined by some as "any secondary care-based activity that requires a hospital bed" (Herbert et al. 2017). Throughout this thesis, its definition is: an inpatient hospital admission delivered in one hospital covering the period from admission to discharge.

In order to retain only hospital activity related to substance misuse, I define substance misuse related hospital admissions using the International Classification of Diseases, Tenth Revision (ICD-10) diagnosis codes summarised in Table 2.1.

The codes prefixed with the letters F, X and Y included are used by the ONS (2018) and others (Pierce et al. 2016) to define drugs misuse related mortality. These are the codes F11-F16 and F18-F19 (mental and behavioural disorders due to psychoactive substance use, excluding alcohol and tobacco); X40-X44 (accidental poisoning by drugs, medicaments and biological substances); X60-X64 (intentional self-poisoning by drugs, medicaments and biological substances); X85 (assault by drugs, medicaments and biological substances); and Y10-Y14 (poisoning by drugs, medicaments and biological substances, undetermined intent). However, I include a set of additional codes: T401-407; T409; T436 (poisoning by drugs, medicaments and biological substances [including overdose and taken in error]). These codes are used in government reporting of substance misuse related hospital admissions (see for example Public Health Wales 2016: 92; NHS England 2014: 28; Home Office 2013); and have been shown to be used in clinical coding of discharge summaries in which recreational drug toxicity is reported (Shah et al. 2011).

Table 2.1: ICD-10 codes used to define admissions related to substance misuse

Description	ICD-10 codes
Mental and behavioural disorders due to psychoactive substance use, excluding alcohol and tobacco	F11-16; F18-19
Accidental poisoning by drugs, medicaments and biological substances	X40-X44
Intentional self-poisoning by drugs, medicaments and biological substances	X60-X64
Assault by drugs, medicaments and biological substances	X85
Poisoning by drugs, medicaments and biological substances, undetermined intent	Y10-Y14
Poisoning by drugs, medicaments and biological substances (including overdose and taken in error)	T40.0-40.7; T40.9; T43.6

HES APC data contain twenty diagnosis code fields for each episode (i.e. row) of data. The first field contains the primary ICD-10 diagnosis, and the other fields contain (in no particular order) secondary ICD-10 diagnoses. I identified episodes with any of the above codes in the primary or secondary diagnosis fields in HES and retained all episodes within the spell containing these episodes using both a unique spell identifier and an episode ordering variable.

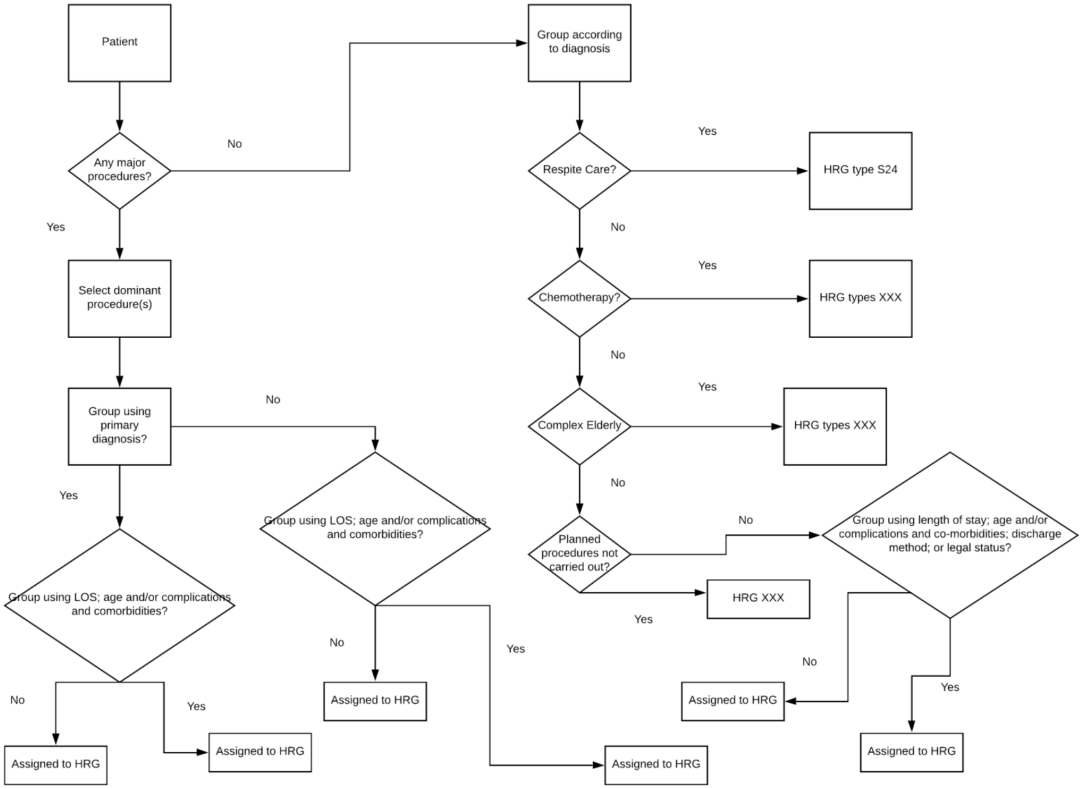
There are then two main samples of interest for this study that are contained in the above subsample of HES:

i. hospital spells in the financial years 2009-10 to 2015-16 containing at least one episode with at least one primary drug-related diagnosis;

ii. hospital spells in the financial years 2009-10 to 2015-16 containing at least one episode with at least one primary or secondary drug-related diagnosis.

Hospitals in the English NHS are paid under the activity-based financing model called 'Payment by Results' (PbR). PbR is the payment system in England under which commissioners pay providers of health care for each patient seen or treated, accounting for the complexity of the patient's health care needs (DH 2012). PbR covers the majority of acute health care in hospitals, with national tariffs for admitted patient care. In order to create this payment system, a set of clinically meaningful

Figure 2a: Determination of HRGs



groups of diagnoses and interventions that consume similar amounts of NHS resources known as healthcare resource group(s) (HRGs) are defined, and each HRG covers a spell of care. HRG codes are used to assign patients to groups for which the treatments, and therefore the associated costs, are similar (Jacobs et al. 2018). HRGs are constructed from electronic patient records and depend on an algorithm that considers elements such as procedures, diagnoses and age (Street 2006). This is outlined in Figure 2a.

It is necessary to restructure the dataset from episode to spell-level in order to cost hospital activity. This allows for the linkage of payment data but entails a range of tasks, and at various stages data are not retained for several reasons. This procedure is therefore outlined in detail.

2.1.1.1 Cleaning the episode-level HES data

There were N=1,491,535 episodes of care containing either a primary or secondary diagnosis of substance misuse for 2009-10 and 2015-16. I generated a spell identifier for these episodes using a grouping variable based on a combination of: person identifier; admission date; provider code; provider spell number; and the regular attender identifier. There were N=14,395 episodes that had missing data for the provider spell number meaning that a spell identifier could not be created – these episodes were dropped, leaving N=1,477,140 episodes. The final episode for each spell should contain information regarding a patient's discharge (such as the time, date and discharge method). This information is required in order to calculate length of stay (LOS) (which is in turn required to calculate what providers are paid by commissioners). I copied this information across all episodes within a spell.

There are several fields in HES that provide information regarding the HRG and how it is determined. The HRG field of interest for this study is the "suscorehrg" (SUS core HRG) field, which gives the Secondary Uses Service (SUS) PbR derived main health care resource group (HRG) at spell-level. Within this field, I set values of either "UZ01Z" or "N/A" equal to missing, as these values are indicative of invalid HRGs.

2.1.1.2 Cleaning the data after aggregation to spell-level

After performing the above tasks in the episode-level data (N=1,477,140), I aggregated to spell-level by taking an index episode (N=1,263,422). I then cleaned the data prior to merging in the payment data, which required addressing several issues.

2.1.1.2.1 Addressing spells containing 'unfinished' episodes of care

The first of these issues relates to the fact that information regarding whether the period of care being delivered was finished or unfinished is allocated at episode-level. This means that a spell, for which the identifier is based on the admission date, can include unfinished episode(s). A range of approaches are possible for addressing this – including removing unfinished episodes from the data prior to aggregation to spell-level. In this study, I created the spell identifier as set out and created an indicator for spells containing unfinished episodes prior to aggregation of the data to

spell-level. I then examined spells that included unfinished episodes to examine the proportion of spells affected by this issue.

First, I dropped spells that had missing data in the HRG field (the first step outlined in Figure 2b. This reduced the sample by from N=1,263,422 to N=1,133,712.

Of N=1,133,712 spells including a primary or secondary diagnosis of substance misuse that had non-missing HRG data, there were N=2,993 (0. 26%) spells that included unfinished episodes.

It was not possible to ascertain whether hospitals are or are not paid in such cases for the spell in the financial year in which the admission started. There is an ambiguity as: the spell-level core HRG code which is used for charging for the activity is non-missing, but the spell includes one or more episodes flagged as being 'unfinished'. Guidance for acute costing by hospitals (NHS Improvement 2019) indicates that hospitals are required to apportion activity to 'patient events regardless of whether they are complete or incomplete at the end of the costing period' (NHS Improvement 2019: 19). There would however appear to be flexibility given to spells for which the assignment of activity to financial years is complicated by ambiguity (NHS Improvement 2019: 19):

> "While incomplete patient events may not be material for some providers, for those providing specialist and/or long-term physical or mental healthcare such as spinal units or high secure units, they can be significant. We know that 'work in progress' is included in financial accounts. Organisations are required to follow the principles of IAS18 in relation to revenue recognition; for example, income relating to partially completed episodes at financial year-end should be apportioned across the financial years on a pro-rata basis. Costs of treatment are then accumulated as they are incurred. Given the timing of the completion of the final accounts and cost data, the values for work in progress and for incomplete patient events will be different. There is no requirement to reconcile them, though the incomplete patient events cost data may be helpful in future assessments of income due for annual accounts purposes."

Of these N=2,993 spells with valid HRG data including unfinished episodes, the vast majority (91.91%) of cases occur in the period 2013-14 to 2015-16 (Table 2.2).

Figure 2b: Cleaning spell-level data prior to merging payment data

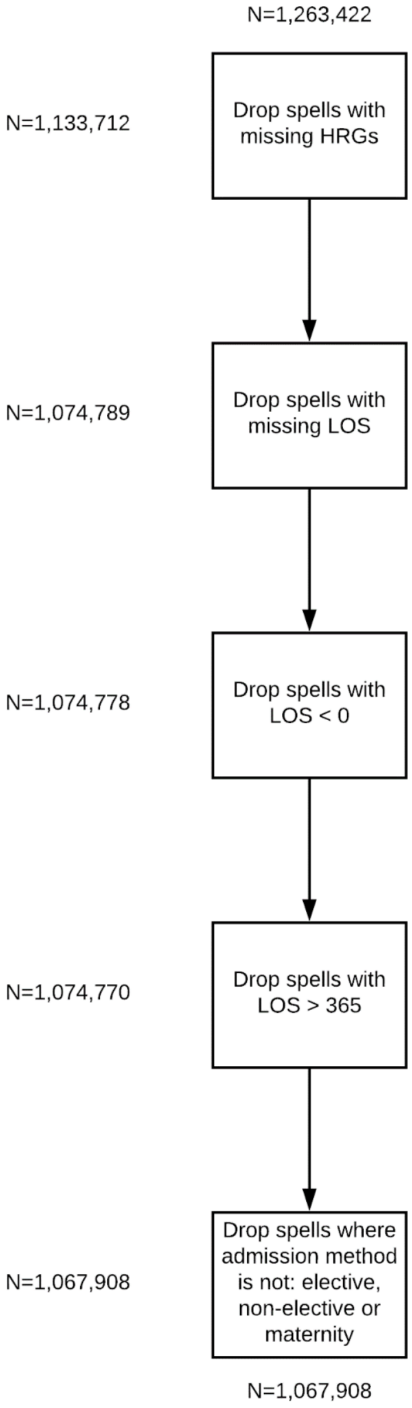


Table 2.2: Spells affected by unfinished episodes, by year

Year	Including unfinished episode(s)	Finished	Total
2009-10	67	141,903	141,970
2010-11	50	152,293	152,343
2011-12	78	159,102	159,180
2012-13	47	157,396	157,443
2013-14	548	170,909	171,457
2014-15	1,158	169,215	170,373
2015-16	1,045	179,901	180,946
Total	2,993	1,130,719	1,133,712

Note: spells including a primary or secondary diagnosis of substance misuse; spells shown are those with non-missing HRG data; prior to data cleaning on LOS and admission method (see Figure 2b for matching sample size N=1,133,712)

2.1.1.2.2 Further exclusions from the spell-level data

After removing spells with missing HRG data, I dropped spells with missing (N=58,923) or negative values (N=1) for length of stay (LOS), removed spells with values of LOS greater than one year (N=19), and finally dropped spells without a value of admission method that could be used in determining appropriate tariff prices (N=6,862) (Figure 2b).

2.1.1.3 Merging in payment data from the NTPS

Once I had aggregated the data to spell-level, I sorted by HRG and financial year of admission. This matches the stratification of the hospital payment data – the National Tariff Payment System (NTPS). The NTPS is published annually and contains the prices and payment rules for commissioners and providers to use for one or more financial years (NHS England 2017). Each row of the NTPS provides information required for pricing each HRG. Generally, in each year of the NTPS, each HRG has a code, a name describing the treatments/conditions covered, and pricing information recorded in the following fields:

i. a combined tariff for elective day cases and longer stays;

ii. a tariff for non-elective admissions;

iii. trimpoints for elective and non-elective admissions;

iv. per-diem rates for the component of each stay after the LOS exceeds the trim-point;

v. reduced short-stay tariffs;

vi. information on best-practice tariffs and top-ups for specialised services.

I linked the data between HES and the NTPS by HRG and financial year. Not all admissions in HES, however, have a matching record in the NTPS.

Table 2.3 shows the number of admissions in HES with any substance misuse related diagnoses by financial year, and numbers of these admissions that have an HRG code for either 'mental health' or maternity care. These two types of admission warrant further discussion.

Table 2.3: Number of spells, maternity admission and mental health HRGs (N=1,067,908)

	Total	Maternity		Mental Health	
	N	N	%	N	%
2009-10	134,896	2,784	2.06%	2,962	2.20%
2010-11	143,947	2,696	1.87%	3,493	2.43%
2011-12	150,318	3,590	2.39%	3,767	2.51%
2012-13	148,113	3,594	2.43%	4,196	2.83%
2013-14	161,171	3,395	2.11%	5,212	3.23%
2014-15	159,939	2,882	1.80%	5,999	3.75%
2015-16	169,524	3,057	1.80%	7,095	4.19%
% Change since 0910	25.67%	9.81%		139.53%	

Maternity admissions (as identified by the admission method[1]) comprise just under 2.06% of all substance misuse related admissions in 2009-10 – falling to 1.8% in 2015-16 (Table 2.3). Having a baby is the most common reason for any admission to hospital in England (NAO 2013). Hospital reimbursement for maternity services is complex – changing from episodic payment to a pathway-based payment system in April 2013, and there is limited certainty about how resources are assigned to activity in the maternity pathway (NAO 2013: 5). HES data on maternity admissions have significant limitations: mothers and babies cannot be linked, many fields have significant problems with missing data, and important information about the background of the mother and her care during labour is not collected (IFS 2017; NHS Digital 2016). The NHS Maternity Services Data Set (MSDS) was introduced

[1]Identification by the admission method is more reliable due to changes in HRG assignment in April 2013

in June 2015 in recognition of these limitations. Maternity admissions are removed from the sample(s) for this study for a combination of these reasons.

Table 2.4 then shows, after the removal of maternity admissions, how many of the remaining admissions have no linked record in the NTPS by financial year. These are shown separately for admissions containing 'mental health' HRGs; and then 'all other' HRG codes in HES that have no matched equivalent in the NTPS.

Table 2.4: Number and % of admissions (excluding maternity admissions) with HRGs in HES without a match in the NTPS

	Total	Unmatched HRGs		Mental Health	
	N	N	%	N	%
2009-10	132,112	508	0.38%	2,957	2.24%
2010-11	141,251	727	0.51%	3,490	2.47%
2011-12	146,728	887	0.60%	3,766	2.57%
2012-13	144,519	1,066	0.74%	4,194	2.90%
2013-14	157,776	1,327	0.84%	5,211	3.30%
2014-15	157,057	1,554	0.99%	5,998	3.82%
2015-16	166,467	1,782	1.07%	7,094	4.26%
% Change since 910	26.00%	250.79%		139.91%	

Admissions with mental health HRGs are shown separately in Table 2.4 because there are no national prices in the NTPS for three specific mental health HRG codes, but they comprise a significant proportion of the HRGs in this study. Each of the three codes describe the same type of treatment delivered to different age groups:

i. WD11Z: All patients 70 years and older with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider;

ii. WD22Z: All patients between 19 and 69 years with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider;

iii. WD33Z: All patients 18 years and younger with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider.

For these three HRGs, prices are negotiated locally between providers and commissioners. Historically, this constituted block contracts between the commissioner and the hospital to provide broadly-defined mental health services in a geographical area (Jacobs et al. 2018). More recently, an episodic payment approach has been 'introduced' for mental health services provided by hospitals – but evidence suggests

that (despite guidance to the contrary) providers and commissioners have not implemented this in practice. There are alternative costs which could be used for these codes in the form of the NHS reference costs; but from 2012-13 onwards only – meaning that this activity could not be costed prior to this. Activity assigned to these HRG codes was therefore removed from the sample.

Once maternity admissions and admissions with mental health HRGs were removed, there were still a small proportion of admissions that could not be costed using the NTPS because there was no HRG in the payment data that matched the HRG in HES. The full definitions of the four most frequently occurring unmatched codes are:

i. PA52A: Behavioural disorders with LOS 1 day or less (HRG has no national price);

ii. WA14A: Planned procedures not carried out for medical or patient reasons (HRG has no national price) (introduced in 2015-16);

iii. WA14B: Planned procedures not carried out for other or unspecified reasons (HRG has no national price) (introduced in 2015-16);

iv. WA14Z: Planned procedure not carried out (HRG has no national price) (replaced in 2015-16 by WA14A and WA14B).

The latter three codes represent activity for which hospitals are not paid. I removed admissions assigned to HRGs with no matching code in the NTPS data from the study sample.

2.1.2 Methods

2.1.2.1 Estimating top-down costs

Top-down estimates of the costs for the same sets of hospital admissions are produced using unit costs for hospital stays produced by Personal Social Services Research Unit (PSSRU). These costs are produced annually and provide the estimated costs for different aspects of hospital care (see for example PSSRU 2016): elective inpatient stays; non-elective inpatient long stays; non-elective inpatient short stays; and day cases (Table 2.5).

Table 2.5: Unit costs used in top-down costing 2009-10 to 2015-16 (£)

	09-10	10-11	11-12	12-13	13-14	14-15	15-16	% Ch.
Elective inpatient stays	2,749	2,931	3,191	3,283	3,403	3,405	3,653	32.88%
Non-elective inpatient stays (long)	2,197	2,334	2,461	2,581	2,716	2,863	2,900	32.00%
Non-elective inpatient stays (short)	523	549	586	598	611	608	616	17.78%
Day cases	637	686	680	697	157	844	713	11.93%

They are based on the 'NHS Trust and Primary Care Trust combined' dataset and include (PSSRU 2016): costs which can be easily identified with a particular activity (such as consultants and nurses); costs which cannot be directly attributed to an activity but can usually be shared (such as laundry); and overheads which relate to the running of each organisation.

The identical samples used to estimate bottom-up costs are divided into these four categories, and then unit costs are applied in each year:

$$C_t^t = \sum_{z=1}^{Z} c_{zt}^t \tag{2.1}$$

where c_{zt}^t represents top-down costs (denoted by superscript t) for each of the four aspects of hospital care z in year t; C_t^t represents aggregated costs in year t.

2.1.2.2 Estimating bottom-up costs

Estimating bottom-up costs requires the use of more complicated algorithms. For the period 2009-10 to 2015-16, separate tariffs exist for different types of admissions: non-elective admissions; and elective admissions – with separate tariffs within elective care for day cases, and short stays for 2009-10 only. In general, each HRG in each year has a lump sum amount that providers are paid depending on whether the activity is elective or non-elective. If a patient stays in hospital beyond a given threshold in terms of the number of days (known as the 'trimpoint'), then providers are paid an

additional amount equal to the product of: an HRG-specific per diem rate; and the LOS minus the trimpoint.

From 2009-10 to 2015-16, admissions are costed using the following general approach:

$$c^b_{ijkt} = \pi_{jkt} \qquad \text{if} \quad x_{ijkt} > s_{ijkt} \qquad 2.2$$

$$c^b_{ijkt} = \pi_{jkt} + [p_{jkt}\,(s_{ijkt} - x_{ijkt})] \qquad \text{if} \quad x_{ijkt} \leq s_{ijkt} \qquad 2.3$$

where c^b_{ijkt} denotes bottom-up (denoted by superscript b) cost for spell i for admission type j, HRG k, in financial year t; π denotes the lump sum paid; x denotes the 'trimpoint' beyond which providers are paid an additional amount for each additional day of care; p denotes the day rate; and s denotes the LOS. For HRGs that have a specific tariff for day cases and short stays, this activity is costed using day case and short stay tariffs depending on tariff rules in each financial year.

These costs are aggregated to give national costs by financial years as follows:

$$C^b_t = \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{N} c^b_{ijkt} \qquad 2.4$$

where C^b_t represents bottom-up aggregate/national costs for financial year t.

This process is performed for admissions containing any (i.e. primary or secondary) substance misuse related diagnosis; and is then repeated for admissions containing primary diagnoses only. All costs are adjusted for the market forces factor (MFF) specific to the provider using the provider codes in HES. The MFF adjustment accounts for unavoidable variation in input prices across the country (Street et al. 2014). MFF adjustments are contained in the NTPS and give index values between 1 and around 1.3 by provider code which are applied to costs after calculation using the above algorithm. This is consistent with guidance which instructs that 'PbR income' is equal to the product of (activity*tariff price) and the value the MFF index (DH PbR 2010). Inclusion of MFF adjustments allows annual changes to reflect the distribution of activity across England (for example from low cost to high cost areas).

2.1.2.3 Comparing approaches to costing, inclusion of related activity and comparing across individual characteristics

Estimated costs are compared using top-down and bottom-up approaches, and depending on whether primary only or primary and secondary (i.e. any) diagnoses of substance misuse are used for identifying relevant activity. These different approaches to costing are then compared across frequently occurring HRG codes and a number of individual characteristics: age, sex, ethnic origin, IMD score.

2.2 Data and methods for Chapter 4: "Exploring patterns of hospital activity and costs relative to admission with a diagnosis of substance misuse: an event study"

The aim of this study is to characterise patterns of use of hospital inpatient services in England for individuals between 2009-10 and 2015-16 who are observed to have at least one substance misuse related hospital admission during the analysis period using an 'event study approach', in which event (sometimes described as index) time is defined for each individual as being relative to their first (observed) substance misuse related hospital admission.

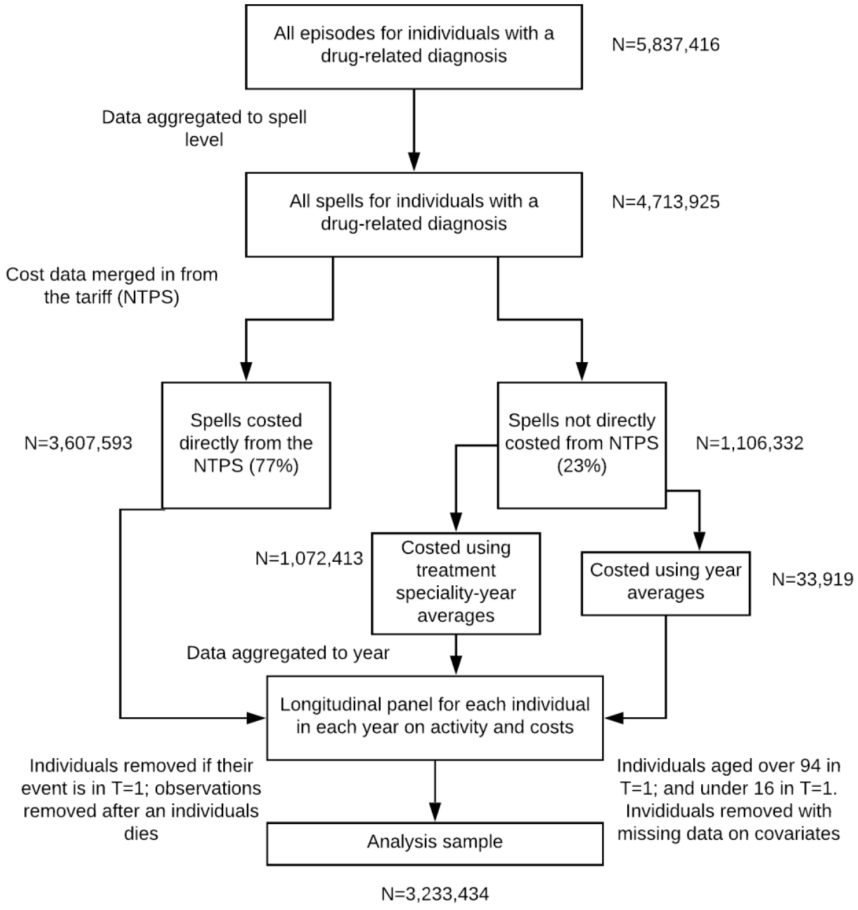
2.2.1 Data

This section sets out in detail how the analysis sample is created using activity data from HES, linked to payment data from the NTPS and mortality information from the Office for National Statistics (ONS). The lengthy procedure entails transforming an episode-level dataset into a longitudinal panel of hospital activity and costs for each person in the study population. This procedure is also illustrated broadly in Figure 2c.

2.2.1.1 Identifying the study population

I used activity data from the 'admitted patient care' (APC) data files in Hospital Episode Statistics (HES) for the financial years 2009-10 to 2015-16. I retained a subset of the APC data based on the same definition of activity related to substance

Figure 2c: Construction of the analysis sample



misuse that was set out previously. I then generated an identifier for individuals who had any episode(s) containing at least one of the (previously listed) substance misuse related diagnosis codes in either the primary or secondary diagnosis field(s) in HES. All of the inpatient hospital activity was retained for individuals with this identifier for the period 2009-10 to 2015-16. As such this led to the retention of an episode-level activity dataset on a population who had at least one drug-related hospital episode in the period 2009-10 to 2015-16.

Use of secondary codes as well as primary codes to define activity related to substance misuse again reflects evidence which suggests that primary codes are often related to a specific symptom rather than the drug use which is the underlying cause (Shah et al. 2011). For example, Shah et al. (2011: 782) found that, for cocaine-related

chest pain the primary code assigned was more likely to relate to the chest pain symptom than the use of cocaine. Use of both primary and secondary codes allows for examination of patterns of activity before and after admissions with a diagnosis of substance misuse where the 'event' does not depend on being defined narrowly for diagnoses that typically reflect the presenting condition. Chapter 4 of this thesis and other studies (see for example Ward et al 2000) indicate the restrictiveness of an approach based on primary codes only.

2.2.1.2 Attaching costs to hospital spells for the study population

Costs were attached to hospital spells using the bottom-up costing methodology outlined previously. However, a specific approach is needed for the treatment of missing costs in this study, as it is necessary to include hospital activity where costs are missing as such activity can contain information relating to the defined 'event' (admission to hospital diagnosed as being related to substance misuse).

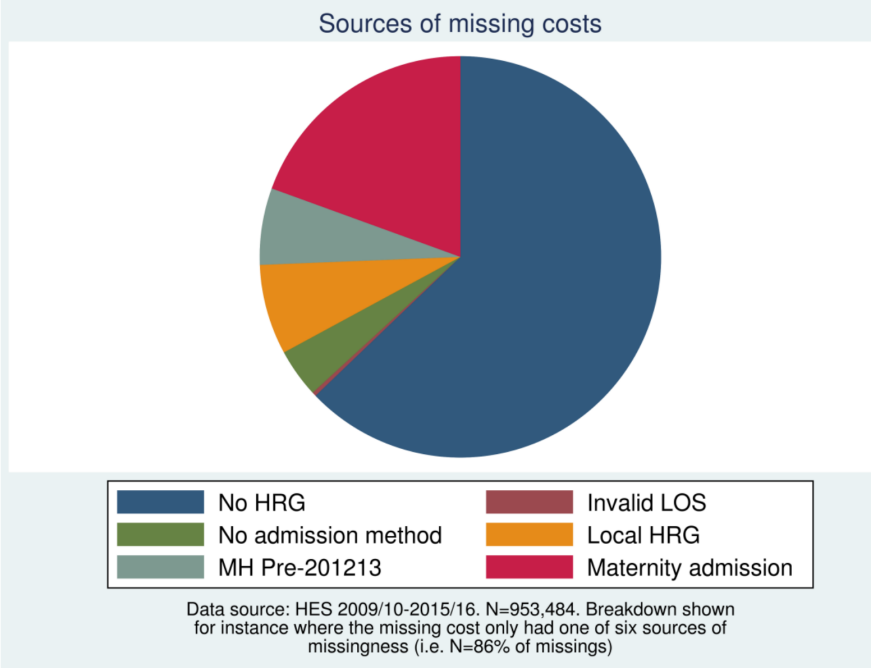
2.2.1.3 Missing Costs

Application of bottom-up costing algorithms to the hospital activity in the study dataset led to costing of 76.46% of the activity retained from HES (see Figure 2c). The dataset for this study is larger and more varied in its scope than that described in the previous section (2.1) insofar as it does not restrict its focus to admissions containing a diagnosis of substance misuse only.

There is therefore a higher proportion of activity that is not costed subsequent to the procedure outlined above. However, whilst Chapter 3 is primarily concerned with comparing approaches to costing; this study has a different purpose. It seeks to characterise the yearly pattern of costs and activity relative to a year in which an individual has hospital activity containing a diagnosis of substance misuse. This means that, in some cases, there is hospital activity that includes a diagnosis related to substance misuse – that for some reason cannot be costed directly using data from the NTPS (the national tariff). It is therefore necessary to seek an alternative approach to enable the widest possible retention of activity. This sub-section sets out in detail how missing costs are addressed to that end.

In this instance, there are six possible reasons as to why a spell is not costed. These are outlined in full for transparency and completeness:

Figure 2d: Reasons for missing costs



(i) the spell contained episodes where the HRG field was missing;

(ii) the information on length of stay was either missing or invalid;

(iii) the information on admission method was missing;

(iv) the HRG code in HES was not contained in the NTPS (locally agreed HRGs);

(v) the HRG code was for one of three mental health HRGs (WD11Z; WD22Z; WD33Z) prior to 2012-13;

(vi) the admission method indicated a maternity admission.

Additionally, 14% of spells with missing costs had two or more of the above reasons as to why costs were missing. For the remaining 86% (N=953,484) of missing costs, the breakdown by reason is illustrated visually in Figure 2d. In addition, Table 2.6 sets out, for each year: the number of spells with missing costs and a breakdown of the numbers missing by reason.

One reason for costs not being attached after merging in the NTPS data is that the dominant HRG for the spell relates to three codes for mental health activity that do not have national prices – prices are set locally between commissioners and hospitals.

Table 2.6: Reasons for spells not having a matched cost in the NTPS (tariff) data

Year	Spells not costed	Reason for missing costs						
		Invalid LOS	Missing HRG	Maternity admission	Mental health HRG (pre-1213)	Invalid admission method	Multiple reasons	Locally priced
2009-10	160,566	283	91,936	33,515	17,959	5,353	9,116	2,404
2010-11	186,059	261	106,089	35,275	19,446	6,035	10,818	8,135
2011-12	187,869	237	102,913	38,884	21,526	6,072	10,103	8,134
2012-13	164,507	218	100,063	40,567	0	5,915	8,798	8,946
2013-14	136,367	252	68,401	17,664	0	5,335	31,522	13,193
2014-15	135,002	265	66,743	17,756	0	4,700	31,426	14,112
2015-16	135,962	1,570	63,022	1,842	0	4,119	51,065	14,344
Total	1,106,332	3,086	599,167	185,503	58,931	37,529	152,848	69,268

Each of these HRG codes refers to the same type of treatment delivered to different age groups:

i. WD11Z: All patients 70 years and older with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider

ii. WD22Z: All patients between 19 and 69 years with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider;

iii. WD33Z: All patients 18 years and younger with a Mental Health Primary Diagnosis, treated by a Non-Specialist Mental Health Service Provider.

These codes are for activities covered under local pricing arrangements between hospitals and commissioners. In the absence of tariff data, I use data for the above three mental health HRGs from the NHS reference costs for the years in which they are available - which is 2012-13 onwards only.

The NHS reference costs provide a comprehensive picture of how NHS providers in England spend allocated funds delivering health care to patients. They represent the average unit cost to the NHS of providing defined services to NHS patients in England and have been collected annually since 1997. They differ from the tariff insofar as the tariff represents a price that is equal to the amount that providers are reimbursed for particular activities; whereas the reference costs are unit costs that represent the average cost of providing particular activities. Both the NTPS and the NHS reference costs are publicly available data.

The costs for the above three codes remain missing for years prior to 2012-13, meaning that another approach was required. For these and the majority of the remaining missing costs I adopted an alternative approach to allow a cost to be attached to the activity. I followed the methodology used by NHS England (2016) and calculated average costs by year and treatment specialty. Treatment specialty is a field in HES that defines the specialty under which the consultant was working during the period of care (NHS Digital 2017). There are 144 treatment specialty codes in HES for the population of interest; and I calculated the within-sample average cost for each of these specialties in each financial year. I then merged these data into the spell-level data and replaced missing costs with the treatment speciality-year averages. This led to costs being attached to 1,072,413 spells (of 4,713,925). The remaining 33,919 cases were records where the treatment specialty field was blank. In these cases, year averages were used.[2]

[2]I also repeat the analysis of costs excluding spells not costed directly from the tariff

2.2.1.4 Transformation of spell level data to a longitudinal panel

I then transformed the spell-level dataset into a panel dataset by retaining an index spell. This first required copying particular information across spells within years. Aspects of this procedure required the following assumptions to be made:

i. I generated an indicator for in-hospital death so that it applied to all spells in year t;

ii. I generated a mean value for the 'Indices of Multiple Deprivation' (IMD) score and copied it across all spells in year t.

iii. I generated a variable to show whether the year contained a spell in which care was provided where a drug-related diagnosis was made.

In addition to this I created a range of variables that calculated a particular summation in order to produce the dependent variables in the subsequent analysis:

i. the total number of spells (herein admissions) for each individual i in year t;

ii. the total number of emergency admissions for each i in year t;

iii. the total number of bed days (total LOS) as a hospital inpatient for each i in year t;

iv. the total cost of the admissions for each i in year t.

Characteristics that have been shown to be related to a range of health outcomes and the use of services were retained including some that are fixed over time such as sex and ethnicity. No manipulation of the data was required for these time invariant characteristics.

Conversion of the spell-level data into a panel format required the generation of records where the dependent variable(s) are set to be equal to zero (i.e. there was no hospital admission for that i in year t). As such, an observation was created for each i in year t where there was no hospital admission that contains only: the individual identifier and year.

Initially, the remaining variables for such an observation were set equal to missing. For variables that represent counts, such as the number of hospital admissions an

individual had in a given year, these missing values were then set equal to zero as is appropriate. For the time invariant characteristics such as age and sex, the information was simply copied across for each year t for the individual i.

In the case of age, information was copied across as follows: the data were sorted by individual and year in chronological order; and age was set to equal its value in $t-1$ plus one:

$$a_{it} = a_{it-1} + 1 \tag{2.5}$$

$$(\text{where } a_{it} = \omega;\ \text{and } a_{it-1} \in R \mid a_{it-1} > 0)$$

where a_{it} represents the age of individual i in year t; and ω represents 'missing'. The data were then sorted in reverse-chronological order, and age was set to equal its value in $t-1$.

$$a_{it} = a_{it-1} - 1 \tag{2.6}$$

$$(\text{where } a_{it} = \omega;\ \text{and } a_{it-1} \in R \mid a_{it-1} > 0)$$

This led to the formation of a balanced panel for the period 2009-10 to 2015-16. I generated a binary variable equal to one if the year contained an admission in which care was provided where a drug-related diagnosis was recorded – this defines the event in this study. Finally, an event time index was created.

The transformation of the data from spell-level to a panel structure is also illustrated in Figures 2e and 2f. These two figures illustrate how 'missing years' are set to zero for values of the dependent variables, and how the event variables are generated using a combination of whether an individual i in year t had hospital admission(s) containing a record of a drug-related diagnosis. The figures also illustrate how time varying characteristics (such as the IMD score) are dealt with.

Figure 2e: Data structure at spell level

Hesid	Spell information									individual characteristics			
	Admission year	Spell	Spell cost	LOS	Admission method	Hrg code	.	Drug diagnosis	Treatment specialty	Age	Sex	IMD	ethnicity
1	201112	1	420	1	elect	WA15F	.	0	110	38	F	26.94	white
1	201415	2	190	1	non-elec	BX12X	.	0	140	41	F	41.3	white
1	201415	3	2597	5	non-elec	WT13D	.	1	311	41	F	54.42	white
1	201415	4	7811	10	non-elec	BX12X	.	1	311	41	F	54.4	white
2	201011	1	489	1	non-elec	ME12P	.	1	.	25	M	15.55	mixed
3	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
N	T	S	.	.	.	.	.	.	.	.	.	.	.

Figure 2f: Data structure after transformation to panel

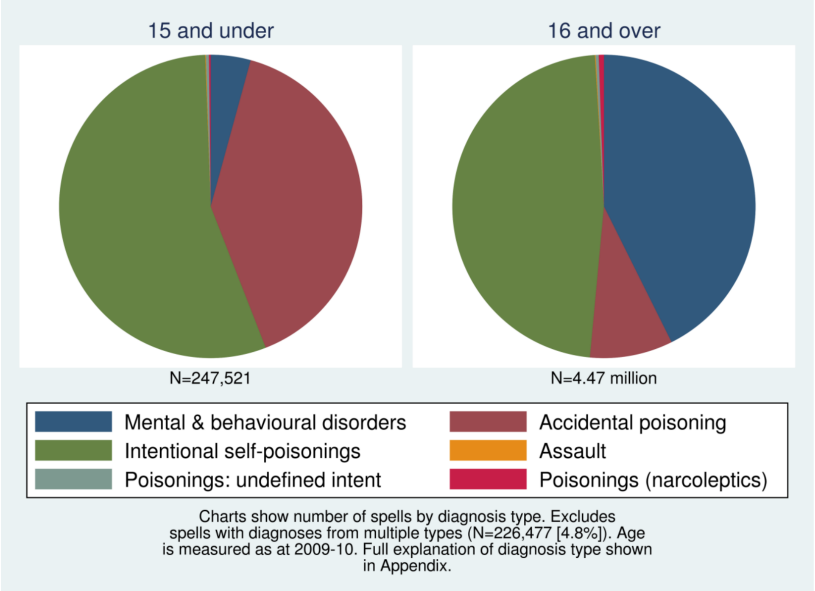
Panel variables		Event variables			Use of services				Individual characteristics			
Hesid	Year	Event	Pre / Post	Event time	Year cost	Total admissions	Emergency admissions	Total LOS	Age	IMD	.	Sex
1	200910	0	0	-5	0	0	0	0	36	38.49	.	F
1	201011	0	0	-4	0	0	0	0	37	38.49	.	F
1	201112	0	0	-3	420	1	0	1	38	26.94	.	F
1	201213	0	0	-2	0	0	0	0	39	38.49	.	F
1	201314	0	0	-1	0	0	0	0	40	38.49	.	F
1	201415	1	1	0	10598	3	3	16	41	50.04	.	F
1	201516	0	1	1	0	0	0	0	42	38.49	.	F
2	200910	0	0	-1	0	0	0	0	24	15.55	.	M
2	201011	1	1	0	489	1	1	1	25	15.55	.	M
2	201112	0	1	1	0	0	0	0	26	15.55	.	M
2	201213	0	1	2	0	0	0	0	27	15.55	.	M
2	201314	0	1	3	0	0	0	0	28	15.55	.	M
2	201415	0	1	4	0	0	0	0	29	15.55	.	M
2	201516	0	1	5	0	0	0	0	30	15.55	.	M
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
N	T	(0,1)	(0,1)		.	.	.	.	.	.	.	.

2.2.1.5 Exclusions from the longitudinal panel

Examination of the data on admissions by age showed that there were individuals being admitted at very young ages (infants and children). This is potentially related to polypharmacy and drug interactions. This empirical study does not limit its scope to activity excluding such a diagnosis – all activity is included for individuals who have such a diagnosis during the study period.

However, the activity at younger ages warranted further consideration. Further exploration showed that these admissions had a markedly different profile in terms of the types of diagnoses. Figure 2g shows activity for persons aged 15 and under, excluding activity containing multiple ICD codes of interest. The same comparisons were performed for other age groups (including older age/elderly). At older ages (70 to 95), there was more limited evidence of higher numbers of accidental poisonings – potentially indicative of polypharmacy, but for persons aged 95 and over the breakdown closely reflected that shown in Figure 2g for younger persons. This examination led to exclusion of 123,804 individuals aged 15 and under in 2015-16 (N=866,056); and 142 individuals aged 95 and over in 2009-10 (N=1,089). Further exclusions from the sample after its aggregation to a panel included individuals with missing data on covariates (age, sex, IMD, ethnicity); and individuals whose event occurred in 2009-10 (N=851,144), who are excluded as the effect of the event cannot

Figure 2g: Types of drug-related diagnosis for over and under-16s



be identified for individuals who have no time variation in the event dummy.

Finally, those who were observed to have died during the analysis period were removed from the dataset for the years after their death. This information (financial year of death) was merged in using an encrypted identifier in HES and in ONS mortality data for the same period. This led to the removal of N=95,852 observations from the dataset. Those who were observed to have died in hospital during the analysis period represent 1.55% of all individuals. The attrition from the sample occurring as a result of death is shown in the results.

2.2.2 Methods

2.2.2.1 Modelling counts

The dependent variables in this study are counts (numbers of events for a fixed period of time), and count health care data are known to exhibit marked skewness (Basu & Manning 2009; Manning & Mullahy 2001), which violates the assumption of normally distributed outcomes in OLS regression.

Specifically, the outcome (Y_{it}) variables considered are:

(i) the number of hospital admissions for each i in year t;

(ii) the number of emergency hospital admissions for each i in year t;

(iii) The number of drug related hospital admissions for each i in year t;

(iv) The number of non-drug related hospital admissions for each i in year t;

(v) the total length of stay for each i in year t; and

(vi) the cost of the hospital admissions for each i in year t.

Counts are non-negative integers which represent the number of events for a fixed period (Cameron & Trivedi 2013). The Poisson distribution is often used to model count data. Poisson models of count data assume that events occur randomly and uniformly in time, and that events occur with a known average (Greene 2003). To demonstrate this formally, let x equal the number of events per time period:

$$Pr(x) = \frac{\lambda^x e^{-\lambda}}{x!}; \quad x = 0, 1, 2, 3 \ldots \tag{2.7}$$

where λ is an intensity parameter which gives the expected number of occurrences in a fixed period of time. It is also the variance of the count, that is:

$$\lambda = E[X]; \quad \lambda = Var[X] \tag{2.8}$$

The Poisson distribution is most appropriate for use when the number of time periods is large and the probability of the event is small (Cameron & Trivedi 1990). Alternatives to OLS are required for count data because the dependent variable has restricted support (Greene 2003). OLS can predict values that are either negative and or non-integer. Counts can therefore be modelled using Poisson regression using the generalised model:

$$Pr(Y_i = j \mid X_i) = \frac{\lambda_i^j e^{-\lambda_i}}{j!}; \quad j = 0, 1, 2, 3 \ldots \tag{2.9}$$

$$\lambda_i = E[Y_i \mid X_i] = Var[Y_i \mid X_i] = e^{(X_i'\beta)} \tag{2.10}$$

The coefficients calculated in Poisson regression do not have a direct interpretation due to exponentiation. To interpret the coefficients, partial effects can be calculated (using either the delta method or bootstrapping for standard errors (Brostrom & Holmberg 2011)):

$$\frac{\delta\{\lambda_i = E[Y_i \mid X_i]\}}{\delta X_{i,k}} = \lambda_i \beta_k \tag{2.11}$$

The parameters do not indicate their marginal impact – the relative sizes indicate the relative strength of the variable:

$$\frac{\dfrac{\delta\{\lambda_i = E[Y_i \mid X_i]\}}{\delta X_{i,k}}}{\dfrac{\delta\{\lambda_i = E[Y_i \mid X_i]\}}{\delta X_{i,j}}} = \frac{\lambda_i \beta_k}{\lambda_i \beta_j} = \frac{\beta_k}{\beta_j} \tag{2.12}$$

The Poisson model has limitations – for example it assumes that all events are independent. What is referred to in the literature as 'herding behavior' violates the assumption of independence (Cameron & Trivedi 1990). Herding behavior is formally referred to as positive contagion and it increases the variance of the count. If the mean and variance are not equal then this causes overdispersion:

$$E[Y_i \mid X_i] < Var[Y_i \mid X_i] \tag{2.13}$$

Under overdispersion, the standard errors and p-values estimated by Poisson regression are too small. Overdispersion can also be caused by either omitted variables bias or unobserved heterogeneity. This requires the estimation of a model without equidispersion (equal mean and variance); and one of these is a negative binomial regression, which uses the negative binomial distribution (Cameron & Trivedi 2001).

Negative binomial regression takes the form:

$$Pr(Y_i = j \mid X_i) = \frac{\Gamma(\theta + y_i)}{\Gamma(y_i + 1)\Gamma(\theta)}\, r_i^{y_i}(1 - r_i)^{\theta}; \quad j = 0, 1, 2, 3 \ldots \tag{2.14}$$

where $\lambda_i = e^{(X_i'\beta)}$; $r_i = \dfrac{\lambda_i}{\lambda_i + \theta}$

The conditional mean function is the same as for Poisson, but the variance function is different:

$$\lambda_i = E[Y_i \mid X_i] = e^{(X_i'\beta)}; \quad Var[Y_i \mid X_i] = \lambda_i(1 + \alpha\lambda_i) > \lambda_i \tag{2.15}$$

As such, negative binomial regression incorporates over (and under) dispersion at the cost of an additional parameter $\alpha = \frac{1}{\theta}$. Where $\alpha = 0$, the model is equivalent to the Poisson model. Negative binomial regression is considered to be appropriate in cases where this parameter is significantly different to zero as this is indicative of overdispersion; and that the conditional variance exceeds the conditional mean.

Another issue to consider in the analysis of count data is that of 'excess zeros': if there are genuine zeros and excess zeros generated by two separate processes, then a zero-inflated model or equivalent is required to incorporate this (Atkins & Gallop 2007). The typical example is models that include two subgroups of individuals: 'always zero' and 'not always zeros'. The assumption contained in this study is that the

data generating process reflects that individuals have at least one hospital admission during the study period, and therefore where they are observed to have had zero admissions in a given year this represents a true measurement.

Additional methodological concerns that arise in this setting are that the values of the dependent variable are likely to be serially correlated for each unit in part due to state dependence and serial correlation in shocks. Different approaches can be used such as a population averaged model, two-way effects model, individual-specific effects model, and mixed (random coefficients) model. The general approaches are similar to those for linear panel regression models. However, the use of random and fixed effects models has complications in this setting: the use of random effects is often not tractable); and fixed effects models are complicated by the incidental parameters problem (Lancaster 2000). This can be shown for a generalized FE model for count data which takes the form:

$$y_{it} \mid x_{it} \sim f[\alpha_i \lambda_{it}] = f[\alpha_i e^{\boldsymbol{x}'_{it}\boldsymbol{\beta}}] \tag{2.16}$$

In general, estimation of the above model is not possible for shorter panels. This is due to what is referred to as the incidental parameters problem: where there are N fixed effects α_i in addition to K regressors, this means that there are $(N + K)$ parameters. $(N+K)$ tends to ∞ as N tends to ∞. This is a particular issue for large samples with relatively fewer time periods, such as this study. In general, FE count models contain considerable computational issues despite solutions to the incidental parameters problem such as FE Poisson estimated using MLE on a differenced transformation (Cameron & Trivedi 2013). In the case of the negative binomial model, a true FE model cannot be fitted if there are time-invariant variables as the model uses a regression decomposition of the overdispersion parameter and not a regression decomposition of the mean:

$$\lambda_i = e^{(c_i + \boldsymbol{x}'_{it}\boldsymbol{\beta})} \tag{2.17}$$

As an alternative to use of FE, a generalized RE model for count data takes the form:

$$y_{it} \mid x_{it} \sim f\left[\alpha_i e^{\boldsymbol{x}'_{it}\boldsymbol{\beta}}\right] = f[\alpha_i e^{(ln\alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta})}] \tag{2.18}$$

where α_i is unobserved and assumed to be uncorrelated with x_i. Negative Bino-

mial regression with random effects is equivalent to two effects where one varies over time and the other is time-invariant (see Greene 1997), and it is often the case that the model is over-specified. Solutions to this are available including conditionally corrected approaches, such as Mundlak's (1978) conditionally correlated random effects. However, such approaches are generally attractive only for balanced panels (Wooldridge 2010), and this study does not use a balanced panel as a result of attrition out of the study population due to death.

For this study, the preferred count data approach will depend on whether the outcomes exhibit overdispersion. In the case of overdispersion, negative binomial regression is preferred. However, alternative approaches will be compared including adjustment for unobserved heterogeneity.

Whilst the appropriate estimation technique for dependent variables exhibiting these count-data properties is generally assumed to be either Poisson or negative binomial regression, there is also evidence to suggest that simpler methods are preferred in large samples where near-normality of the sample means is assured (Mihaylova et al. 2011; Li et al. 2012). Other studies have shown how OLS type approaches can be used for health care resource use or cost data that are transformed in order to overcome problems of skewness (Manning & Mullahy 2001; Veazie et al. 2003; Jones et al. 2015). As Basu & Manning (2009: S109) noted: "no method is optimal or dominant for all cost applications. Many of the diagnostics used in choosing among alternatives have limitations that need more careful study". I therefore make a final comparison – OLS regression applied to transformation of the data using the inverse hyperbolic sine function (Zhang et al. 2000):

$$asinh\, y = ln\left(y + \sqrt{y^2 + 1}\right) \tag{2.19}$$

This transformation is comparable to taking the log transformation of a variable but is preferable in the case of a variable with zero values.

The distributions of the outcomes are illustrated in Figures 2h to 2j and Table 2.7. The figures show, in quadrants, the distribution of four of the dependent variables listed earlier. The four admission variables are closely related and have similar distributional properties. Two example cases (all admissions and emergency admissions) of these four are therefore shown alongside LOS and cost which have some difference in terms of their distributions. The effect of the inverse hyperbolic sine transformation can be seen on the distributions of these variables when comparing across Figures.

Table 2.7: Distributional statistics for dependent variables and inverse hyperbolic sine transformations

Statistic	Annual cost of admissions		Total LOS (per year)		Number of hospital admissions (per year)		Number of emergency admissions (per year)	
	Y	$asinh(Y)$	Y	$asinh(Y)$	Y	$asinh(Y)$	Y	$asinh(Y)$
Mean	1,504.59	3.27	4.62	0.65	0.94	0.53	0.57	0.37
Variance	3.00E+07	16.52	460.50	1.66	12.23	0.57	2.01	0.39
Standard Deviation	5,477.23	4.06	21.46	1.29	3.50	0.75	1.42	0.62
N	3,233,434							

Figure 2h: Distribution of dependent variables

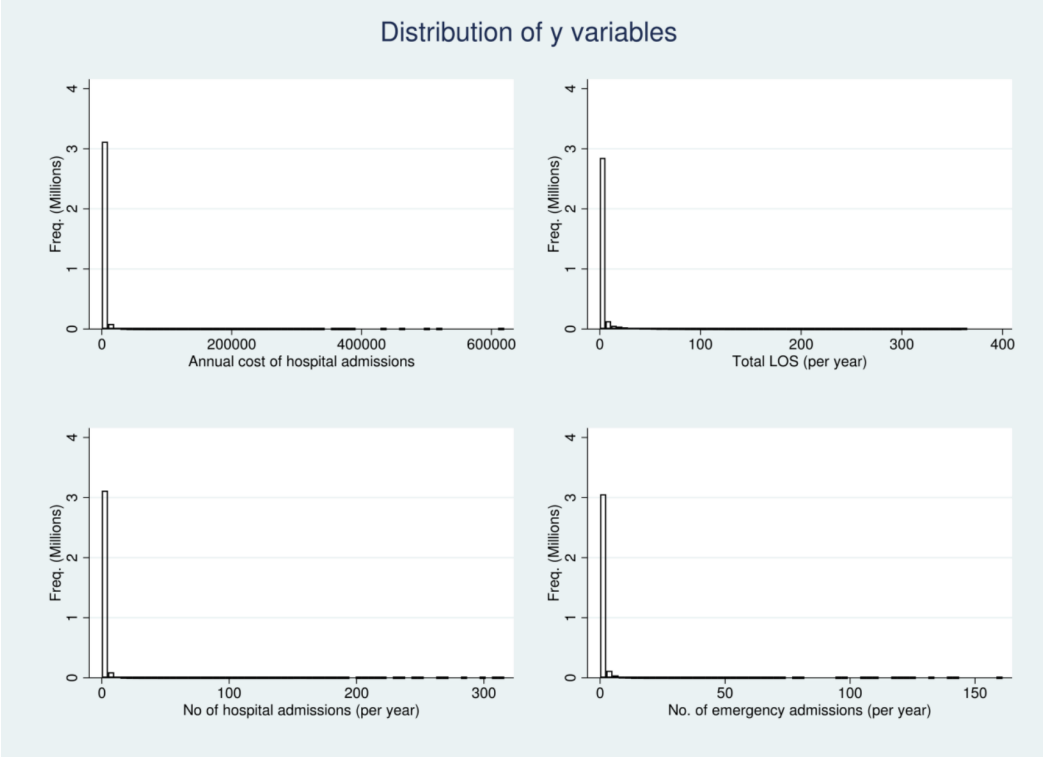


Figure 2i: Distribution of inverse hyperbolic sine transformation of the dependent variables

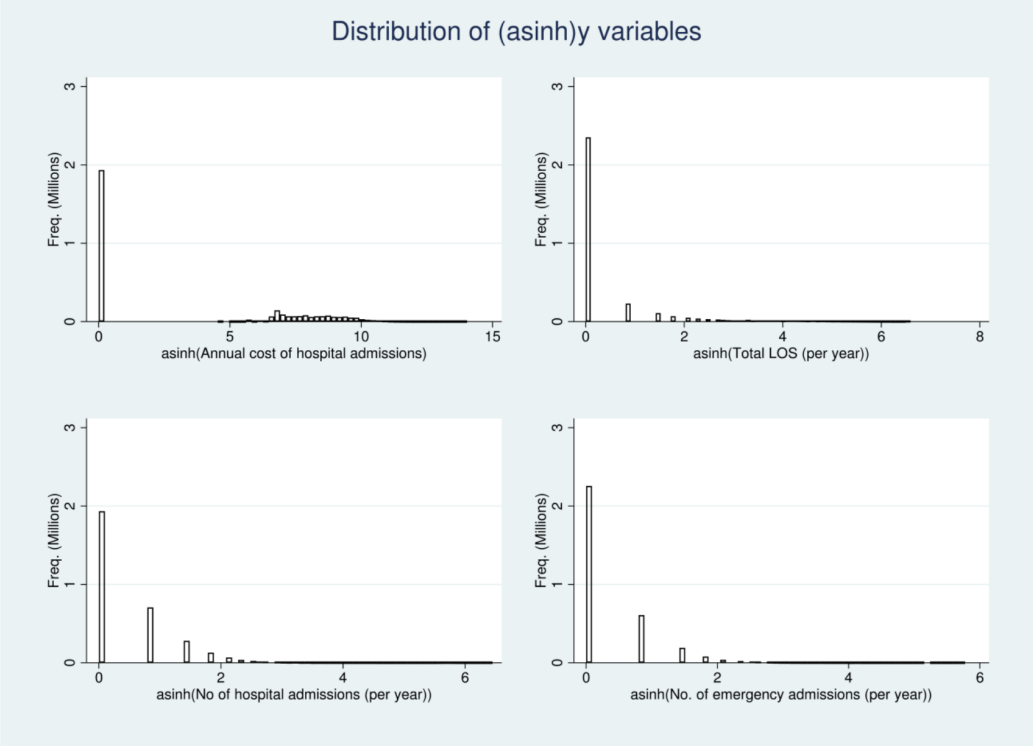
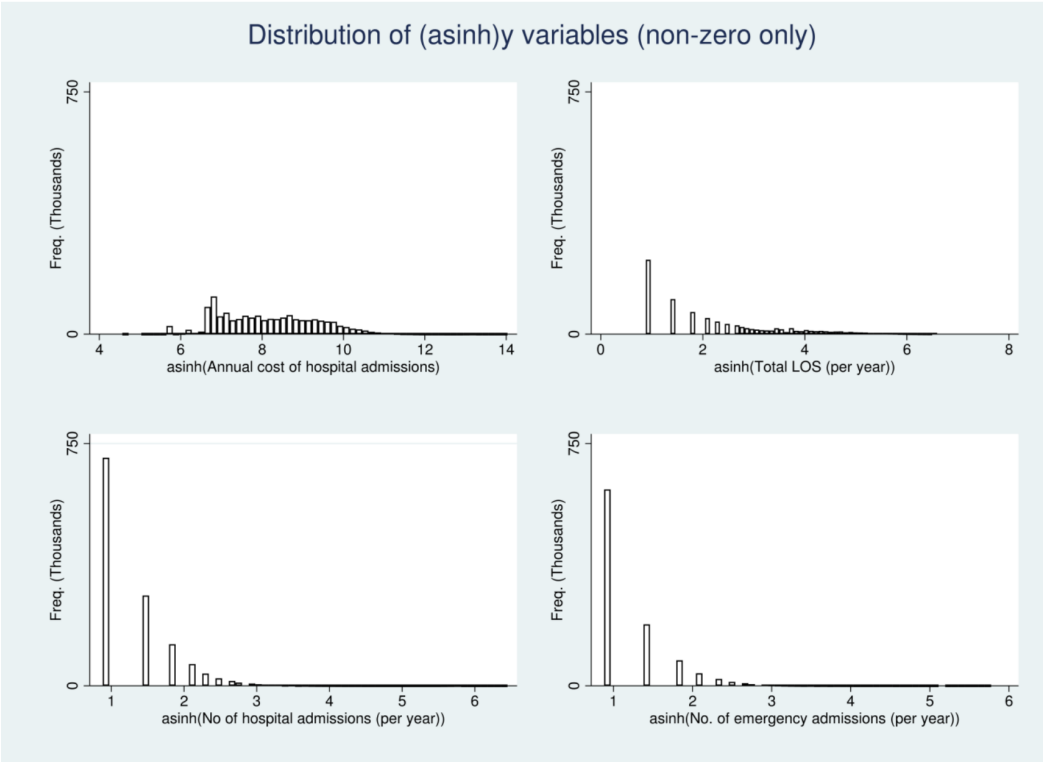


Figure 2j: Distribution of inverse hyperbolic since transformation of the dependent variables (non-zero values only)



2.2.2.2 Event Studies

Event studies are applied in economics as a framework in which treatment effects can be estimated for situations in which the timing of treatment varies across units. In this paper, I consider a panel of $i = 1, \ldots, N$ individuals in which the outcome Y_{it} is observed for $t = 1, \ldots, T$ periods (calendar time). Each individual has, at some point during the analysis period, a substance misuse related hospital admission. The year in which this occurs is defined as the event. This identification strategy allows for a comparison the hospital activity and costs of individuals before and after diagnosis of substance misuse in hospital (Dobkin et al. 2016).

I analyse the coefficients on indicator variables for time relative to the event (event time). The generalised specification takes the form:

$$Y_{it} = a + \gamma_t + X_{it}\beta + \sum_{r=j}^{J-1} \mu_r + \varepsilon_{it} \tag{2.20}$$

where Y_{it} is one of six outcome variables for individual i in year t; γ_t are coefficients on calendar time fixed effects, X_{it} represents a vector of control variables; μ_r are coefficients on binary indicators for event time $j = 1, \ldots, J$.

The key coefficients of interest in this study are the pattern on the μ_r dummies which estimate the outcome at a given r relative to the omitted category J.

The coefficients on μ_r for each r prior to the occurrence of the event represent "pre-trends". The consideration of pre-trends is important for studies in which the event represents a policy change which is assumed to be exogenous to the event (Freyaldenhoven et al. 2018).

In the econometric analysis of panel data, it cannot be assumed that the observations are independently distributed over time (Wooldridge 2009). Unobserved factors that affect outcomes in one time period are likely to affect the outcome in another time period. Failure to account for this can lead to correlation between the error term ε_{it} and the variables of interest. This violates the assumptions of techniques such as ordinary least squares (OLS) and prevents estimation of unbiased and consistent parameters. Inclusion of individual-specific effects (either random or fixed effects) allow for unbiased estimation in the presence of unit-specific heterogeneity (Wooldridge 2009).

A complicating factor in terms of controlling for individual heterogeneity relates to the collinear relationship that exists between the individual-specific event($\mu_{r=E}$), the year fixed effects (γ_t) and the individual fixed effect a_i. The relationship between the event and calendar time is fixed for each individual.

As stated by Borusyak & Jaravel (2017), this problem is effectively the same as the (well-known) age-year-cohort problem. Solutions to the problem of collinearity in such situations have been proposed (see for example Mason et al. 2015a; Deaton & Paxson 1994; Deaton 1997). Borusyak & Jaravel (2017) approve two approaches for dealing with the above under identification issue: restricting pre-trends and retaining the use of fixed effects; and replacing unit fixed effects with unit random effects. In this study, I primarily compare models with and without unit random effects and compare the effect of their exclusion on the parameter estimates.

2.2.2.3 Estimation of the generalised model using count data methods

The generalised event model (2.20) estimated using the various count approaches discussed previously. First, the generalised model (2.20) is estimated on the inverse hyperbolic sine transformation of each outcome:

$$asinh\, Y_{it} = a + \gamma_t + X_{it}\beta + \sum_{r=j}^{J-1} \mu_r + \varepsilon_{it} \tag{2.21}$$

(2.20) is also estimated using Poisson regression:

$$Pr(Y_{ijt} = s \mid \kappa) = \frac{\lambda_i^j e^{-\lambda_i}}{y!}; \quad s = 0, 1, 2.. \tag{2.22}$$

where $\kappa = a + \gamma_t + X_{it}\beta + \sum_{r=j}^{J-1} \mu_r + \varepsilon_{it}$; $\lambda_i = E[Y_i \mid \kappa] = Var[Y_i \mid \kappa] = e^{\kappa}$.

Negative binomial regression is also estimated on (2.20):

$$Pr(Y_i = s \mid \kappa) = \frac{\Gamma(\theta + y_i)}{\Gamma(y_i + 1)\Gamma(\theta)}\, r_i^{y_i}(1 - r_i)^{\theta}; \quad s = 0, 1, 2, 3 \ldots \tag{2.23}$$

where $\lambda_i = e^{(\kappa)}$; $r_i = \dfrac{\lambda_i}{\lambda_i + \theta}$.

In order to consider the impact of accounting for unobserved unit-specific heterogeneity, models are shown that are estimated with and without individual-specific random effects.

2.2.2.4 Sensitivity Analyses.

Two sensitivity analyses are carried out in this study. First, I test the sensitivity of the results to the assumption that the event is the individual's first hospital admission with a diagnosis of substance misuse. This assumption is violated for individuals who have an event in the few years leading up to the analysis period (for example in 2008). I therefore compare the sensitivity of the results to exclusion of all individuals whose event occurs in the first two financial years of the study period. Second, I compare the sensitivity of the results to the approach taken to costing activity that was not costed directly from the NTPS by excluding any such activity.

Regression analyses are estimated using Stata 15.1 SE. All regression models use cluster-robust standard errors (clustered on individuals) which control for both overdispersion and correlation over time for each individual (Greene 1997). Estimates are shown for different count data models, and the preferred model (interpreted in the narrative) is chosen based on whether the conditional variance exceeds the conditional mean - as indicated by the size and significance of the estimate of the overdispersion parameter from negative binomial regression. Akaike's and Bayes Information Criteria (AIC and BIC) are included - but are only directly comparable for models with identical likelihood functions and do not therefore inform selection of the preferred model between regression approaches.

2.3 Data and methods for Chapter 5: "The impact of paying care providers for outcomes: differences-in-differences analysis of the payment by results for drugs recovery programme"

This study evaluates the impact of the introduction of a P4P scheme for providers of addiction treatment services in eight geographical areas of England in April 2012. It uses individual-level difference-in-differences analysis comparing two treatment outcomes, successful (abstinent) completion and declining to continue with treatment,

for individuals located in intervention and control areas. The study population is individuals in treatment for problematic substance use in England for 2011/12 and 2012/13.

As previously noted in the declarations, this study was published as a research report in 'Addiction' in 2015 alongside multiple co-authors. Therefore, the following explanation of data and methods uses the personal pronoun 'we' instead of 'I'.

2.3.1 Data and Methods

We obtained anonymised individual-level data from the National Drug Treatment Monitoring System (NDTMS) from the first nine months of the financial years (1st April to 31st December) 2011-12 and 2012/13. NDTMS records details for all episodes of structured community-based and residential drug misuse treatment paid by the public sector in England. A total of 182,447 Service users received treatment for primary drug use during the financial years (1st April to 31st December) 2011-12, and 178,860 during 2012/13. We also exploit data for the three financial years prior to the introduction of the P4P scheme in order to assess historical trends in complexity between pilot and non-pilot areas.

2.3.1.1 Metrics

For payment purposes, the formal definition of service users who are free from drugs of dependence is those who are (HMG 2013):

"[d]ischarged from treatment successfully (free of drug(s) of dependence) and do not re-present in either the treatment system or in the criminal justice system in the following twelve months."

Full achievement of this outcome can only be measured twelve months after discharge. In this study, we consider achievement of successful completion of treatment, as this is the first stage of this measure and successful completion triggers the first payment to providers in the 'abstinence' domain. A 'successful completion' is defined by the National Treatment Agency (now Public Health England) as when an individual completes a planned treatment and is (HMG 2013):

"[j]udged by a clinician to be free of dependency from the drug for which the individual was being treated, and in addition not using either heroin or crack."

The client may be abstinent from or occasionally using other drugs within this definition. We measure the rate of successful completion by the number of treatment episodes that end with a discharge reason of either 'treatment completed – drug free' or 'treatment completed – occasional user (not heroin and crack)' divided by the total number of treatment episodes in a given year.

We also consider rates of treatment exit that resulted from service users declining to continue with their treatment ('declining'). A possible unintended consequence of paying providers based on completion rates might be a reduction in continued engagement with treatment services by service users, if the nature of treatment changes under the new payment system. Engagement with treatment services has been shown to reduce service users' likelihood of future hospitalizations and emergency department usage (Neighbors et al. 2005).

2.3.1.2 Comparator Groups

The eight intervention sites are commissioning localities known as Drug Action Team areas (DATs) that successfully applied to participate in the pilot programme. They were selected by the DH to be representative of all DATs on a range of characteristics such as the demographic and complexity of their local populations. The remaining 140 DATs could decide to adopt a similar payment mechanism for the providers in their areas but do not form part of the pilot programme. We obtained a database from the NTA (now PHE) of the actual and intended commissioning arrangements as at July 2012 for all 140 of these DATs and assigned them to three groups: (i) adopted P4P; (ii) plan to adopt P4P; and (iii) no plans to adopt P4P. We present our findings comparing the eight pilot DATs to all other DATs.

The eight pilot DATs are located across four of the eight Government Office Regions (GORs) of England (the North West; South East; East Midlands; and Yorkshire and the Humber). The market structure for provision of treatment services varies considerably across different localities. Generally, providers do not compete with each other for service users who require or seek treatment. Typically, in areas where multiple agencies exist and provide services, the services they provide are intended to complement each other and do not overlap (with, for example, one agency providing therapeutic interventions and another providing substitute prescribing). However, there are some limited instances of localities in which competition between providers exists, at least in principle. Across the eight pilot areas, the average number of providers in 2011-12 was 1.875; and (in theory) competition existed in only one of

these eight areas, and this area had two providers offering a distinct, stand-alone service.

2.3.2 Analysis

2.3.2.1 Difference-in-differences

This study seeks to exploit variation across geographical areas and over time periods to identify the impact of the P4P policy on hospital admissions for the population of interest. This setup is typically referred to as quasi-experimental, or as a natural experiment (Jones & Rice 2011), and the most common specific family of methods used is difference-in-differences (DiD).

DiD methods are a specific example of the statistical innovation that has sought to provide solutions to the problem identifying treatment effects using observational data. Ideally, it would be possible to identify the effects of policies or medical treatments by observing an individual both when they are treated and not treated, with identical conditions apart from the treatment. The solution typically used in the evaluation of medical treatments is to randomize individuals to treatment – but this is usually not possible in observational data, and for evaluation of policies for which randomization would be unethical. Observational data are characterized by the non-random assignment of treatment (Hess 2017). The statistical methods developed to adapt to this generally seek to include confounding variables so that the outcome variable and treatment can be considered to be independent, and a specific family of methods used widely in policy evaluation is the DiD.

DiD seeks to obtain causal effects by comparing the change in outcomes before and after the introduction of the intervention for the treatment and control groups (Ashenfelter 1978; Ashenfelter & Card 1985; Bertrand et al. 2004; O'Neill et al. 2016). DiD assumes that the mean outcomes for the intervention and control groups would have followed parallel trends over time in the absence of the intervention (Abadie 2005).

2.3.2.2 Model specification

The generalized difference-in-differences model estimated for each of the two outcomes (completion and declining) took the form:

$$y_{ijt} = \alpha + X_{it}\beta + \gamma_t + \rho_j + \theta D_{jt} + \delta \bar{y}_{j,t=0} + a_i + \varepsilon_{ijt} \quad 2.24$$

where y_{ijt} represents the value of the outcome variable for individual i, in year t, treated in DAT j during the study period; X_{it} represents a vector of control variables; γ_t are coefficients on calendar time fixed effects; ρ_j represents a binary indicator that equals one if DAT j was a pilot participants; D_{jt} represents an interaction between a binary indicator for the post-intervention period and ρ_j; $\bar{y}_{j,t=0}$ represents the baseline level of the outcome variable at DAT level; a_i represents a random effect; and ε_{ijt} is an idiosyncratic error. The key coefficient of interest is θ, which represents the treatment effect.

The analysis is conducted at the individual level. We adjusted for a range of client characteristics at the start of treatment that are known to influence both the probability of successful completion of treatment and declining to continue with treatment (PHE 2014): gender; age; opiate use; crack use; benzodiazepine use; injecting; accommodation status; whether service users were referred to treatment through the criminal justice system; and number of years since first use of the primary drug of presentation. These are contained in the vector of control variables $X_{it}\beta$. We also included the baseline level of the outcome variable at DAT level as an explanatory variable, in order to test whether the estimated effects might reflect 'regression to the mean' (i.e. whether there was differential capacity for improvement in outcomes between the treatment and control DATs). To allow for the nesting of service users within DATs, we used a random effects logistic regression using the xtlogit, re command in Stata v13.

Logistic regression is applied in order to reflect the binary nature of the two outcomes of interest: it allows for non-linearity in the relationship between the outcomes and the explanatory variables; and ensures that the predicted values of the outcome do not lie outside the range 0 to 1 (Wooldridge 2009).

The generalised regression (2.24) therefore takes the following form:

$$ln\left(\frac{p}{1-p}\right) = \alpha + X_{it}\beta + \gamma_t + \rho_j + \theta D_{jt} + \delta \bar{y}_{j,t=0} + a_i + \varepsilon_{ijt} \quad 2.25$$

where $p = E(y_{ijt} = 1)$.

2.3.2.3 Presentation of findings

Following estimation, we used the margins, pu0 command to calculate the magnitude of the estimated effect in terms of the percentage point change in the probability of the outcome. We show the estimated effect of participation in the pilot scheme using a difference-in-differences (DiD) estimator comparing baseline data from 2011-12 to pilot data from 2012-13. The associated odds-ratio and its confidence interval are presented. We also show the marginal effect estimated on the interaction between the indicator for the intervention group and the indicator for the 2012/13 financial year. We test for within-DAT correlation using a likelihood ratio test, testing the null hypothesis that the within-DAT correlation is equal to zero.

2.3.3 Sensitivity analyses

In the main analysis, we compared the eight DATs that participated in the pilot programme with all 140 other DATs. As secondary analyses, we examined the robustness of the results to:

i. comparison with a subset of DATs which were similar in terms of two characteristics of the local population. We identified these comparators using the ViewIt facility on the National Drug Evidence Centre (NDEC) website (PHE 2014a). We used (a), the percentage of the local population who are users of opiates or crack; and (b) the 2010 Index of Multiple Deprivation score. We only included 42 control DATs which were within ten places of at least one of the pilot DATs on both variables

ii. comparison to DATs located in GORs in which there was at least one pilot DAT. This led to 90 DATs in four regions (the East of England, the North East of England, the South West of England, and the West Midlands) being excluded from the comparators. These comparators were also identified using the ViewIt facility on the NDEC website.

We also repeated the main and secondary analyses for an alternative definition of the intervention and comparison groups, comparing all areas that adopted P4P with all areas that did not adopt P4P. This alternative definition was based on the NTA (now PHE) database compiled in July 2012.

2.4 Data and methods for Chapter 6: "Did paying addiction service providers for outcomes lead to unintended consequences? Difference-in-differences analysis of hospital admissions, costs and length of stay in England 2009-10 to 2015-16."

Chapter 6, the final empirical study, focuses on examining whether introduction of the P4P scheme in April 2012 led to unintended consequences in a wider substance misuse-related population using a quasi-experimental design applied to patient-level data, including individual characteristics, and an assessment of underlying trends.

2.4.1 Data

2.4.1.1 Creating the study population

In order to conduct the analyses in this study, I again require a panel dataset that captures changes in admission before and after the introduction of P4P for populations who are and are not exposed to the policy. I use hospital activity data from Hospital Episode Statistics (HES) for the financial years 2009-10 to 2015-16, allowing for a complete assessment of admissions in intervention and comparator populations. The hospital activity data are converted into panel form following the approaches set out previously, and the costing follows the same bottom-up methodology. Missing costs are addressed following the approach previously described.

The remaining individual-level activity and costs panel dataset is summarily the same as the one set out previously; but it contains more individuals as it does not need apply the same exclusions in terms of individuals admitted to hospital in 2009-10 for a drug-related diagnosis. These individuals can be included as the identification strategy in this study is based on geographic area of residence (rather than person-specific changes over time).

2.4.1.2 Fixing individuals to their initial residence

To conduct the analyses, individuals need to be identified as in the intervention or comparison group. This is defined based on initial residence, but it is possible that

individuals could change their residence over the study period and within financial years.

The identifying assumption is made due to concerns that individuals may have self-selected into either the control or intervention by changing their local authority of residence in response to or anticipation of the policy.

I therefore generate a binary variable for individuals which equals one if an individual is observed to change their area of residence during the study period. This is achieved in practice as follows: residence is coded as a numeric variable which takes one of 149 values representing local authorities (1, 2, .., 149). Variables are created for each individual which are equal to the maximum and minimum values of the residence identifier. In cases where these variables take different numeric values, this identifies individuals who migrate between areas during the study period. This migration identifier is then used to generate a subpopulation for testing the sensitivity of the results to the assumption made in fixing individuals to their initial residence.

Local authority of residence in HES maps directly onto DAT (the administrative geography in which the P4P policy was applied). Individuals' who were initially resident in the areas in which the scheme was introduced are allocated to the treatment group; and vice-versa for the comparison group.

2.4.2 Methods

This dependent variables in this study exhibit the same pronounced overdispersion that was discussed previously in this Chapter. These variables are:

i. the number of hospital admissions for each i in year t;

ii. the number of emergency hospital admissions for each i in year t;

iii. The number of drug related hospital admissions for each i in year t;

iv. The number of non-drug related hospital admissions for each i in year t;

v. the total length of stay for each i in year t; and

vi. the cost of the hospital admissions for each i in year t.

The approaches used for dealing with these counts is consistent with that outlined previously. Specifically, models estimated using negative binomial regression, Poisson regression and OLS estimated on the inverse hyperbolic sine are compared in terms of consistency. However, the preferred model is based on assessment of the conditional mean and variance as indicated by the overdispersion parameter included in negative binomial regression.

2.4.2.1 Difference-in-differences: assumptions and alternatives

This study adopts a difference-in-differences approach, following the same conceptualization outlined previously. However, it contains development of the approach insofar as it contains additional sensitivity tests and application of an alternative to the standard DiD model.

The assumption of parallel trends does not always hold (see for example Ryan et al. 2014) and in such cases the adoption of alternative approaches that do not require this assumption is warranted. An alternative family of methods assume that (O'Neill et al. 2016: 2):

> "in the absence of treatment, the expected outcomes for the treatment and control groups would have been the same, conditional on their past outcomes and covariates."

This can be called 'independence conditional on past outcomes' and does not require an assumption of parallel trends.

One such approach is multivariable matching, which can be used to construct a comparison group that is like the intervention group based on pre-treatment outcomes and covariates (Steventon et al. 2013; Kreif et al. 2015; Abadie 2005; Blundell and Costa-Dias 2009; Heckman et al. 1997). There are various specific alternatives that can be used to perform matching such as (Whittaker et al. 2016): nearest neighbour with replacement; radius matching; kernel weighting; and so on.

Matching techniques however require performance assessment in terms of the capacity of the matching variables available in the data for reducing or eliminating the bias that would existing in the absence of matching (Rubin 2001). Specifically, the reliability of regression adjustment must be based on the following conditions (Rubin 2001: 174):

i. "The difference in the means of the propensity scores in the two groups being compared must be small (e.g., the means must be less than half a standard deviation apart), unless the situation is benign in the sense that: (a) the distributions of the covariates in both groups are nearly symmetric, (b) the distributions of the covariates in both groups have nearly the same variances, and (c) the sample sizes are approximately the same."

ii. "The ratio of the variances of the propensity score in the two groups must be close to one (e.g., 1=2 or 2 are far too extreme)."

iii. "The ratio of the variances of the residuals of the covariates after adjusting for the propensity score must be close to one (e.g., 1=2 or 2 are far too extreme); "residuals" precisely defined shortly."

Regression adjustment performance can be assessed both before and after implementing matching to allow for discerning whether the data allow for sufficiently reliable matching of groups. This can be assessed by using Rubin's B and Rubin's R statistics, where the former gives the number of standard deviations between the means of the groups and the latter is the ratio of treatment variance to control variances (Rubin 2001).

An alternative to multivariable matching that can be used in the absence of parallel trends is the lagged dependent variable approach (LDV), which adjusts for pre-treatment outcomes and covariates explicitly in the regression model. The LDV approach has been demonstrated to produce the most efficient and least biased estimates of the average treatment effect when the assumption of parallel trends is violated (O'Neill et al. 2016). Unlike matching, the LDV approach does not rely on additional identifying tests – model reliability can be assessed using the standard regression output and assessment of significance tests.

In order to answer its research questions, this study: applies a standard DiD approach; formally tests the assumption of parallel trends; and then compares estimates from an LDV approach.

2.4.2.2 Model specification

The generalised difference-in-differences model estimated for each of the six outcomes takes the form:

$$y_{ijt} = \alpha + X_{it}\beta + \gamma_t + \rho_j + \theta D_{jt} + \varepsilon_{ijt} \tag{2.26}$$

where y_{ijt} represents the value of the outcome variable for individual i, in year t, resident in area j in their first hospital admission during the study period; X_{it} represents a vector of control variables; γ_t are coefficients on calendar time fixed effects; ρ_j represents a binary indicator that equals one if area j adopted P4P; D_{jt} represents an interaction between a binary indicator for the post-intervention period and ρ_j. The key coefficient of interest is θ, which represents the treatment effect.

The generalised DiD model is specified for the modelling approaches tailored to count data. First, (2.26) is estimated on the inverse hyperbolic sine transformation of each outcome:

$$asinh\, y_{ijt} = \alpha + X_{it}\beta + \gamma_t + \rho_j + \theta D_{jt} + \varepsilon_{ijt} \tag{2.27}$$

(2.26) is also estimated using Poisson regression:

$$Pr(Y_{ijt} = s \mid \kappa) = \frac{\lambda_i^j e^{-\lambda_i}}{y!}; \quad s = 0, 1, 2.. \tag{2.28}$$

where $\kappa = \alpha + X_{it}\beta + \gamma_t + \rho_j + \theta D_{jt}$; $\lambda_i = E[Y_i \mid \kappa] = Var[Y_i \mid \kappa] = e^{\kappa}$.

Negative binomial regression is also estimated on (2.26):

$$Pr(Y_i = s \mid \kappa) = \frac{\Gamma(\theta + y_i)}{\Gamma(y_i + 1)\Gamma(\theta)}\, r_i^{y_i}(1 - r_i)^{\theta}; \quad s = 0, 1, 2, 3 \ldots \tag{2.29}$$

where $\lambda_i = e^{(\kappa)}$; $r_i = \dfrac{\lambda_i}{\lambda_i + \theta}$.

2.4.2.3 Test of parallel trends

The assumption of parallel trends is tested for each of the six outcomes as follows:

$$y_{ijt} = \alpha + X_{it}\beta + t + \rho_j + \psi\upsilon\rho_j + \varepsilon_{ijt} \qquad \forall\, t < T_0 \tag{2.30}$$

where t represents a linear time trend; and T_0 is the first year in which the policy was introduced (the start of the post period). The size of the coefficient ψ on the interaction between υ and the indicator for treatment (pilot) areas ρ_j indicates the extent to which the trends are diverging. The following null hypothesis is tested for parallel trends against the alternative: $H_0 : \psi = 0$; $H_A : \psi \neq 0$.

2.4.2.4 Lagged dependent variable

The LDV models estimated in this study take the generalized form:

$$y_{ijt} = \alpha + X_{it}\beta + \sum_{k=1}^{T_0} \pi_k y_{ij,t=k} + \gamma_t + \rho_j + \varepsilon_{ijt} \qquad \forall\, t \geq T_0 \tag{2.31}$$

where $\sum_{k=1}^{T_0} \pi_k y_{ij,t=k}$ are the lagged values of the outcome variable (the values of the outcome in the pre-intervention period); and T_0 represents the first period in which the intervention was introduced (the post-intervention period).

This generalized LDV model is also estimated for the three alternative approaches to modelling count data, following the same structure as explained for the generalized DiD model. Linear and Poisson regression are estimated including random-effects for individuals in order to account for unobserved heterogeneity.

2.4.3 Sensitivity Analyses

2.4.3.1 Sensitivity to the assumption of fixed residence: intention to treat vs. treatment effect on treated

Individuals are assigned to the intervention according to their initial residence when they are first observed during the study period. This assumes that individuals do not subsequently change their local authority of residence. This assumption does not apply to the whole population of interest over the study period.

In order to test the sensitivity of the estimates to this assumption, the analyses above are repeated excluding individuals who are known to have changed their local authority of residence during the study period.

Estimated effects on the full study population represent an intention to treat analysis, whereas the estimated effects on the subpopulation that are not to have changed their local authority of residence during the study period represented the estimated treatment effect on the treated population.

2.4.3.2 Sensitivity: comparing intervention areas to matched comparison areas

In the main analysis, I compare individuals initially resident in the intervention areas with individuals initially resident in all comparison (i.e. other) areas in England. As a secondary analysis, I examine the robustness of the results to: comparison with a subset of areas which were similar in terms of two characteristics of the substance misuse treatment and resident population. These comparators were identified using the ViewIt facility on the National Drug Evidence Centre (NDEC) website (PHE 2014). I use (a), the percentage of the substance misuse treatment population who are users of opiates or crack; and (b) the 2010 Index of Multiple Deprivation score. I include only 42 comparison areas found to be within ten places of at least one of the intervention areas on both domains.

Matching techniques are explored using the psmatch2 command in Stata 15.1 SE, which allows for assessment of Rubin's R and Rubin's B statistics for assessing the reliability of the matching variables in the data. Use of matching techniques is conditional on obtaining appropriate values for:

i. Rubin's B: absolute standardised difference of the means of the linear index of the propensity score in the treated and matched non-treated group. Rubin (2001) endorses B<25;

ii. Rubin's R: ratio of treated to matched non-treated variances of the propensity score index. Rubin (2001) endorses R within [0.5; 2].

2.4.3.3 Estimation of treatment effects in separate portions of the post-intervention period

The P4P pilot officially ended in April 2014, at which point commissioners in all areas could take their own decision as to whether to retain/introduce P4P in their commissioning area and, if so, the form it would take. In order to examine whether

the estimated treatment effects are distributed differentially across different post-intervention years, I calculate estimates for DiD terms separating the post-period into two sub-periods (2012-13 and 2013-14 vs. 2014-15 and 2015-16):

$$y_{ijt} = \alpha + X_{it}\beta + \gamma_t + \rho_j + \eta_1\rho_j\varrho^1 + \eta_2\rho_j\varrho^2 + \varepsilon_{ijt} \tag{2.32}$$

where ϱ^1 and ϱ^2 represent binary indicators for 2012-13 and 2013-14; and 2014-15 and 2015-16 respectively. Models are estimated using Stata 15.1 SE, and cluster standard errors on individual. Average partial effects are calculated by predicting the outcome with DiD=1 and DiD=0, taking the difference and then calculating mean and standard deviation for the treated units. This allows for illustrating the absolute size of treatment effects, as calculation of marginal effects on interactions in non-linear models creates considerable computational and methodological difficulty.

3 The costs of substance misuse related hospital admissions in England 2009-10 to 2015-16: an exploration using Hospital Episode Statistics

3.1 Introduction

Drugs misuse is a global public health concern. The harms caused by drug use are significant: 0.6 percent of the global adult population were estimated to have a drug use disorder in 2015, with 28 million disability-adjusted life years (DALYs) lost as a result of drug-use and 17 million health years of life lost as a result of drug use disorders (UNODC 2017).

In England, there were: 7,545 hospital admissions with a primary diagnosis of drug-related mental health and behavioural disorders; and there were 82,135 hospital admissions with a primary or secondary diagnosis of drug-related mental health and behavioural disorders in 2015-16; over double the level in 2006-07 (NHS Digital 2018). This represents 0.5% of all hospital admissions.

Whilst this sort of public health information relating to substance misuse is updated annually in England, information on the economic costs of substance misuse is not systematically updated and the approaches used to estimate these costs vary from study to study.

3.1.1 Morbidity, mortality and costs related to substance misuse for England

Policymakers produce annual reports that summarise the numbers of substance misuse related hospital admissions for England (NHS Digital 2018). The most recent report has changed from previous reporting approaches and defines three categories of substance misuse related hospital admissions:

- hospital admissions with a primary diagnosis of drug-related mental and behavioural disorders;

- hospital admissions with a primary or secondary diagnosis of drug-related mental and behavioural disorders;
- hospital admissions with a primary diagnosis (only) of poisoning by illicit drugs.

All three measures count the number of "finished consultant episodes" for a given financial year for a set of substance misuse related diagnosis codes - the International Classification of Diseases, 10th revision (ICD-10). The same reports provide annual statistics on the number of registered deaths related to substance misuse. These statistics provide useful periodic updates on drug-related morbidity and mortality, but the total costs of drugs misuse are challenging to estimate due to the range of impacts. Drugs misuse affects the individual, family and friends, and wider society – including mental health, homelessness, unemployment, criminality and health care. Such information is typically found in cost of illness studies, but these studies represent an intensive undertaking and are not reported annually by government agencies.

The most recent estimates of the costs of hospital admissions related to substance misuse for England are contained in a Home Office (2013) report for the year 2011-12. They applied a version of top-down costing to three defined areas of substance misuse related hospital activity (defined as 'behavioural; poisonings; and neonatal' admissions); and then considered some additional costs in terms of related infectious diseases and lost economic output.

There have been significant changes to government policy since then and the epidemiological composition of the substance misusing population continues to evolve - with reducing numbers of younger and increasing numbers of older users (PHE 2016; Beynon et al. 2007; Gossop & Moos 2008; Gfroerer et al. 2003; Simoni-Wastila & Yang 2006).

The 2013 study divided the estimated costs to the NHS from substance misuse into five parts (Home Office 2013: 69):

i. mental and behavioural health due to illicit drug use;
ii. overdoses and poisoning due to illicit drug use;
iii. neonatal diagnoses due to illicit drug use;
iv. HIV/AIDS for injecting drug users;

v. deaths due to illicit drug misuse, costed in terms of lost productivity, the human cost and medical and ambulance costs.

The approaches taken for estimating these five areas of costs vary considerably in the study - and are explained in turn below.

Hospital Episode Statistics (HES) were used in order to identify inpatient episodes with specific primary ICD-10 diagnoses for mental and behavioural costs. The ICD-10 codes used to identify this activity were: F11.0 – F16.9; and F19.0 – F19.9. Average costs were then applied to the activity using estimated costs produced by the Personal Social Services Research Unit (PSSRU) for adult mental health inpatient bed days (PSSRU 2011).

Overdoses and poisonings were estimated by first identifying relevant episodes using ICD-10 codes (T40.1-T40.7; T40.9; and T43.6) and then applying NHS Reference Cost data for 2010-11 using estimates of the cost of treating 'poisoning, toxic, environmental and unspecified effects'. These estimated costs only cover activity in the year of the study (2010-11) and therefore provide no indication of how costs develop based on longer-term harm related to overdoses. The costs of drug-related neonatal diagnoses (such as babies born suffering from withdrawal of addictive drugs) were estimated by identifying episodes with ICD-10 codes P04.4 and P96.1 and applying NHS Reference Costs for treating major and minor 'neonatal diagnoses'. Costs relating to HIV and AIDs attributable to injecting drug use were estimated applying broad assumptions about the proportion of HIV cases attributable to drugs misuse (3%) to estimates of HIV related expenditure. Finally, the costs of deaths related to substance misuse were estimated by applying the Department for Transport (2011) estimates of the cost of a fatality resulting from road traffic accidents.

This Home Office study represents the most recent detailed research containing information on the costs of hospital care related to substance misuse. More recent reports have used the cost estimates contained in that report and have adjusted them for inflation to produce figures for more recent years (see for example: National Treatment Agency 2014; House of Commons 2017).

There are studies that incorporated measurement of the cost of use of health services related to substance misuse that were carried out prior to this. The Drug Treatment Outcomes Research Study (DTORS) (2009) was a study of outcomes and the cost-effectiveness of treatment for substance misuse. The study was national, multi-centre

prospective study that included an economic analysis of the costs and benefits associated with addiction treatment services (Davies et al. 2009). It included measurement of participants' use of health services through survey self-reported and costed activity using unit costs; and it was a wide ranging and systematic evaluation of treatment services but limited its scope to only a cohort of individuals in treatment for substance misuse. Both casual and problematic substance misuse is estimated to be much more prevalent in the general population than that captured by the addiction treatment cohort (Home Office 2018).

There have been significant changes to both government policy and in the epidemiology of substance misuse since these figures were last produced. UK Government policy towards illicit drugs has changed significantly in recent years (see Chapter 1), and the epidemiology of the substance misusing population has been changing for some time: globally, there are increasing numbers of people misusing substances at ages over 50 (Rao & Roche 2017). The management of this changing population represents a public health challenge for policymakers across the world (Degenhart et al. 2000; EMCDDA 2010).

There are therefore reasons to expect that the costs of substance misuse (and underlying patterns) may have changed in recent years both in terms of hospital services, and more widely.

3.1.2 Approaches for estimating the cost of illness

There are different approaches that exist for estimating disease or illness specific costs. Studies that typically seek to estimate the costs of a particular disease area or public health subject are known as 'cost of illness' studies (Rice 2000). They generally estimate disease specific costs by considering the costs that result directly from the illness, and the indirect costs that result from the illness.

Cost of illness studies are used by policymakers for a variety of purposes. They can be used to: define the magnitude of the disease or injury in monetary terms; justify interventions; assist in the allocation of research resources on specific disease areas; to provide a basis for policy and planning relative to public health schemes; and to provide an economic framework for policy evaluation (Rice 2000). Government agencies internationally have stated that cost of illness studies aid their decision making (Varmus 2000: 4):

> "COI [cost of illness] estimates do not provide a simple formula for the allocation of research resources. They cannot substitute for the well-informed judgement required to synthesize information about the broader dimensions of disease burden with knowledge of scientific opportunities in developing strategies and budgets for research and development programmes. However, COI estimates can provide order of magnitude indicators for particular diseases. While they should be interpreted with caution, COI estimates can help decision-makers in Congress and in the Administration anticipate and respond to public interests".

Cost of illness studies are therefore considered to be of use to policymakers given clear expression and understanding of their limitations.

There are nonetheless significant criticisms of the approaches used for conducting cost of illness studies. Kennelly (2017), for example, argues that they are not especially useful insofar as they simply attach a figure to what is already known in terms of the public health consequences in a given area such as the number of deaths resulting from road traffic accidents (Rice 2000). Controversy relating to COI studies has primarily focused on the inclusion and estimation of indirect costs such as 'productivity losses' – resulting in relatively limited scrutiny on estimates of direct costs (Tarricone 2006). There are a range of criticisms of cost of illness studies, that generally focus on COI studies that take a 'broad brush' approach (Clabaugh et al. 2008; Tarricone 2006).

It is sensible to consider the main characteristics of these alternative approaches to estimating COI to identify their strengths and limitations.

The first difference between studies relates to whether they take either an 'incidence based' or 'prevalence based' approach. The prevalence based approach comprises estimation of direct costs and production losses attributable to all cases occurring in a given year (Tarricone 2006). The incidence approach requires estimating lifetime costs of 'new cases' which have their onset in a given time period (Tarricone 2006). The incidence approach is much less common in the COI literature (Clabaugh et al. 2008).

COI studies also differ in whether they are carried out prospectively or retrospectively. Retrospective studies have practical strengths insofar as they are typically less expensive to conduct than prospective studies, and in that they allow for the investigation of disease over a long time period. This is especially useful for diseases that require long time periods in order to capture relevant outcomes and endpoints.

Prospective studies, however, allow for the database and questionnaires to be designed in order to capture all of the relevant data – and this can be useful in terms of acquiring information from participants in terms of diaries that record transportation costs, time taken off work, and so on. Unfortunately, in the cases of diseases with a long lifespan (Type 1 diabetes, for example), such studies are prohibitively expensive and retrospective studies are therefore the only practical option.

COI studies also vary in terms of the approach by which population costs are assigned - there are three general approaches that exist for assigning population costs (Tarricone 2006; Jo 2014). The first can be described as 'top-down costing', in which portions of known total expenditure are allocated to broad disease specific activity categories. The general form of top-down costing was developed by Rice (1966) and involved allocating national health expenditures for different types of care across disease categories of the International Classification of Diseases (ICD) according to primary diagnosis (Tarricone 2006). This approach to identifying relevant activity has limitations (Tarricone 2006: 54):

> "another problem with this method is that all costs are attributed to the primary diagnosis. This is a serious problem if we consider that a relevant part of all hospital discharges involve patients with multiple diagnoses."

In some cases, top–down approaches identify relevant disease specific activity by using what is referred to as the epidemiological or attributable risk approach – rather than use of primary ICD diagnoses. This method measures the proportion of a disease that is considered attributable to the exposure to the condition or particular risk factors (Bloom et al. 2001; Liu et al. 2002). It exploits aggregated data and applies it to a population-attributable fraction (PAF) to calculate the costs estimated to be attributable to the condition (Hogan et al. 2003; Hodgson & Cohen 1999; Reynaud et al. 2001; Morgenstern et al. 1980). The approach estimates the proportion of a given disease y (for example heart attacks) attributable to risk factor z (for example substance misuse) as follows:

$$PAF = \frac{p_y(rr_{zy} - 1)}{[p_y(rr_{zy} - 1) + 1]} \tag{3.1}$$

where p_y is the prevalence rate of disease y; and rr_{zy} is the relative risk for risk factor z for the population with disease y; compared with those without risk factor z (Hogan et al. 2003; Hodgson & Cohen 1999; Rockhill et al. 1998).

One limitation of this approach to identify disease specific activity is that it assumes that other factors must not affect the association between the two diseases. Confounding factors are likely to exist that affect both diseases such as age, gender, deprivation and so on. This is a classic case of omitted variables bias (Wooldridge 2012). For studies in which top-down estimates are based on attributable fractions taken from analyses lacking sufficient confounding, omitted variable(s) bias can result in an overestimate of the costs attributable to a particular condition (Jo 2014).

Another limitation of top-down costing studies is that they apply average/aggregated costs. However, the costs to hospitals for a given disease vary for each patient based on their use of health care resources whilst in hospital. Application of average costs to a particular population assumes that the average cost for the group of interest equals the average cost for the wider group. This assumption is unlikely to hold given that the use of health care resources varies considerably between patients based on their condition, severity, age, etc.

An alternative to the top-down approach for assigning population costs is known as 'bottom-up costing'. This approach considers the health inputs used (such as hospital admissions) and then applies the unit costs to these inputs (Jo 2014). This method requires detailed data on the use of services by the population of interest in addition to detailed data on the costs of the range of services. The strength of this approach is it allocates national health expenditures amongst major categories of diagnosis – and, unlike top-down studies, this ensures that the sum of the estimated costs for diseases cannot exceed the national health care budget (Drummond 1992). Bottom up studies have, however, frequently used primary diagnoses only to identify relevant activity for costing, which likely undercounts the activity attributable to the condition of interest (Jo 2014). In the case of substance misuse related hospital activity, primary diagnoses codes have been shown to relate to the presenting symptom (such as chest pain) rather than the underlying cause (such as cocaine use) in many instances. (Shah et al. 2011).

A common issue to both top-down and bottom-up studies is that approaches to the identification of relevant activity have particular limitations: use of ICD diagnoses can over or underestimate relevant activity; use of population attributable fractions can be limited by unobserved variables.

An alternative to the top-up and bottom-down approaches is sometimes described as the 'econometric approach'. The approach matches two populations – one with and one without the disease – based on a range of measured characteristics (Finkelstein

et al. 2003). It then produces an estimate of the costs that are attributable to the disease by including a binary variable for individuals who have the condition in a regression analysis in which costs are the outcome (Birnbaum et al. 2005; Swensen et al. 2003). This coefficient on the binary variable gives the average additional cost attributable to the condition. This method still requires some mixture of top-down and/or bottom-up approaches in order to generate these costs for inclusion in regressions and is also susceptible to omitted variables bias (failure to include controls correlated with the having the condition). It improves on population-level approaches by allowing for identification of confounding at the individual-level.

3.1.3 Research questions

Since the costs of substance misuse related hospital admissions were last estimated for England, there have been continued changes to the direction of policy and in the composition of the substance misusing population. There are a range of approaches that are typically used for estimating disease specific costs.

This study aims to estimate the substance misuse related hospital costs for England for 2009-10 to 2015-16 and to illustrate how different approaches to applying costs and defining related activity may impact on these estimates in practice. Specifically, I consider the following research questions:

i. what are the differences in the estimated costs of substance misuse related hospital admissions between bottom-up and top-down approaches?

ii. what is the difference in the estimated costs of substance misuse related hospital admissions depending on whether primary only or primary and secondary ICD diagnoses are used to identify relevant activity?

iii. what differences exist in the distribution of costs across population characteristics depending on the approach used for identifying relevant activity?

3.2 Results

The results of this chapter are presented as follows. First, I compare the differences in the estimated costs that result from using top-down and bottom-up methodology.

I then reflect on factors that partly cause differences between these approaches and impact on changes over the study period.

I then compare differences in estimated costs depending on whether relevant activity is based on primary diagnoses only, or primary and secondary diagnoses. These differences are compared first for bottom-up costing, and then the impact of these alternative approaches for identifying relevant activity is compared across the two costing approaches.

Finally, I compare how total costs, total admissions and costs per admission are distributed according to age, sex, ethnic origin and IMD score. These comparisons are shown visually depending on the costing approach, and the approach taken for identifying relevant activity.

Main result tables and figures are embedded in the text.

3.2.1 Comparison of top-down and bottom-up costing of admissions containing primary diagnoses of substance misuse

Table 3.1 shows differences in estimates of the costs of hospital admissions with a primary diagnosis related to substance misuse depending on the costing approach.

The top-down approach estimates that costs are £2,551,002 (36.74%) higher in 2009-10 than bottom-up costing of the same activity. This difference increases over the analysis period – to £9,068,170 in 2013-14 (100.08% higher); and £9,689,693 in 2015-16 (98.97% higher).

Table 3.1: Admissions, total costs and cost per admission by costing approach (primary diagnosis only)

Year	Admissions	Total Cost (£)		Cost per admission (£)	
		Top-down	Bottom-up	Top-down	Bottom-up
2009-10	11,093	9,494,489	6,943,487	855.90	625.93
2010-11	11,978	10,811,799	7,332,324	902.64	612.15
2011-12	11,973	11,646,474	7,166,555	972.73	598.56
2012-13	15,024	15,153,414	7,656,173	1,008.61	509.60
2013-14	16,957	18,128,701	9,060,531	1,069.10	534.32
2014-15	16,792	18,527,920	9,089,972	1,103.38	541.33
2015-16	17,135	19,479,951	9,790,258	1,136.85	571.36
% change	54.47%	105.17%	41.00%	32.83%	-8.72%

The mean cost per admission using bottom-up methodology starts from a lower baseline of £625.93 per admission compared with £855.90 per admission using top-down costing. This difference increases over the study period as top down mean costs increase by 32.83% over the study period, and bottom-up costs decrease by 8.72%.

3.2.1.1 Bottom-up costs

Bottom-up estimates of costs depend on the distribution of activity across HRGs, but also on whether activity is elective or non-elective and the LOS. Table 3.2 shows that the vast majority of the included activity is non-elective – and that this share of activity is stable over the study period.

The mean LOS for the activity included increases by 6.06% from 1.32 days in 2009-10 to 1.40 days in 2011-12; decreases to 1.35 in 2012-13; and subsequently increases to 1.44 in 2015-16 – an overall increase of 9.09% from 2009-10 (Table 3.3). The values of the median, upper and lower quartiles reflect the skewness of LOS – the standard deviation is around six times the value of the mean. The mean cost per admission calculated using the bottom-up approach also exhibits skewness, albeit less markedly (Table 3.3).

Estimates of bottom-up costs also depend on how activity is distributed across HRGs. Figure 3a gives the four HRGs for which there are the largest numbers of hospital admissions in each year for: the number of admissions; the average cost per admission; the total cost; and the percentage accounted for of the total annual cost for each of the four HRG codes.

Table 3.2: Admissions by elective vs non-elective activity (primary only)

Year	Elective		Non-elective		Total
	N	%	N	%	
2009-10	25	0.23%	11,068	99.77%	11,093
2010-11	36	0.30%	11,942	99.70%	11,978
2011-12	33	0.28%	11,940	99.72%	11,973
2012-13	31	0.21%	14,993	99.79%	15,024
2013-14	30	0.18%	16,927	99.82%	16,957
2014-15	38	0.23%	16,754	99.77%	16,792
2015-16	51	0.30%	17,084	99.70%	17,135
% change	104.00%		54.35%		54.47%

Table 3.3: Summary statistics by year - cost per admissions & LOS (primary only)

Year	Cost per admission (£) (bottom-up)					LOS				
	Mean	S.D.	Median	LQ	UQ	Mean	S.D.	Median	LQ	UQ
2009-10	625.93	671.19	509	382	574	1.32	3.69	1	0	1
2010-11	612.15	1198.21	416	366	508	1.41	5.11	1	0	1
2011-12	598.56	841.32	428	387	501	1.4	4.29	1	0	1
2012-13	509.6	711.28	362	260	362	1.35	3.59	1	0	1
2013-14	534.32	669.12	451	159	451	1.4	3.41	1	0	1
2014-15	541.33	728.08	445	282	473	1.4	3.62	1	0	1
2015-16	571.36	738.59	446	446	503	1.44	3.37	1	0	1

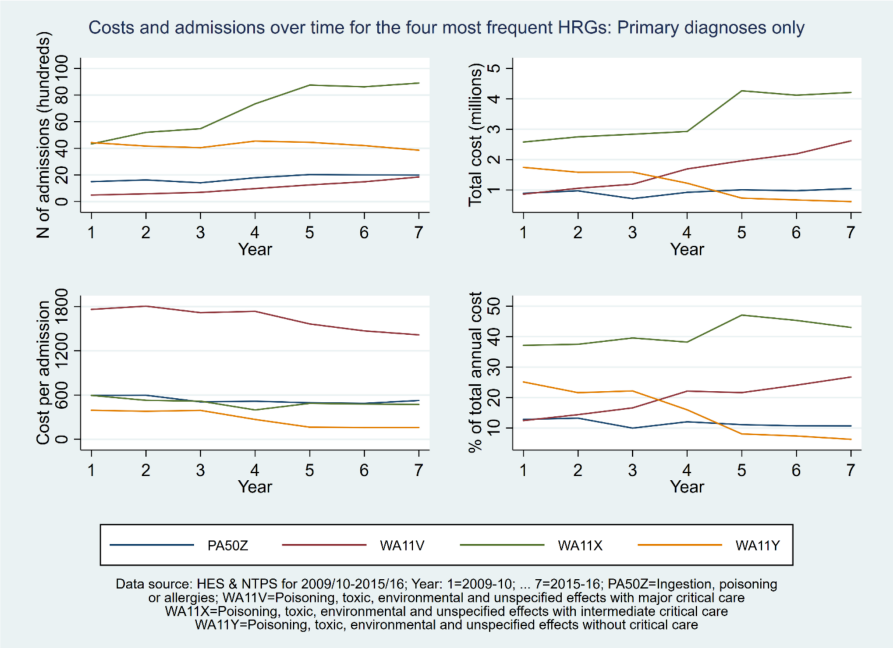
Admissions with HRG codes for 'poisoning, toxic and environmental and unspecified effects with intermediate critical care (code: WA11X) increase to 8,902 in 2015-16 – over double the initial value in 2009-10 (4,328) (Figure 3a). The average cost per admission decreases from £595.85 in 2009-10 to £517.75 in 2011-12; falls sharply to £398.45 in 2012-13; and subsequently increases to £473.02 in 2015-16. The share of total annual costs for activity with a primary diagnosis of substance misuse accounted for by HRG code WA11X increases over the period from 37.14% in 2009-10 to 43.01% in 2015-16.

Admissions with HRG codes for 'poisoning, toxic and environmental and unspecified effects without critical care' (code: WA11Y) decrease to 3,861 in 2015-16 – a 12% reduction since 2009-10. The cost per admission decreases slightly from £393.73 in 2009-10 to £392.47 in 2011-12; falls sharply to £269.34 in 2012-13; and subsequently decreases substantially to £160.23 in 2015-16. The share of total annual costs for activity with a primary diagnosis of substance misuse accounted for by HRG code WA11Y decreases substantially over the period from 25.13% in 2009-10 to 6.32% in 2015-16.

Admissions with HRG codes for 'ingestion poisoning or allergies' (code: PA50Z) increase over the study period from 1,499 in 2009-10 to 1,992 in 2015-16. The average cost per admission decreases slightly from £595.95 in 2009-10 to £526.64 in 2015-16. The share of total annual costs accounted for by this activity decreases slightly from 12.87% to 10.72%.

The fourth most frequent HRG for hospital admissions with a primary diagnoses of substance misuse is WA11V: 'poisoning, toxic and environmental and unspecified effects with major critical care'. The number of admissions assigned to this HRG increases dramatically (by 276.58%) from 491 in 2009-10 to 1,849 in 2015-16. The

Figure 3a: Patterns over time for HRGs assigned most frequently to activity with a primary diagnosis of substance misuse



average cost per admission decreases from £1,761.55 in 2009-10 to £1,416.55 – but the share of total annual costs accounted for by this HRG increase more than doubles nonetheless - from 12.46% in 2009-10 to 26.75% in 2015-16.

The combined share of annual total costs accounted for by activity assigned to these four HRGs is relatively stable over the study period – at 87.59% in 2009-10 to 86.8% in 2015-16.

3.2.1.2 Top-down costs

Top-down estimates of costs incorporate some information relating to length of stay – as non-elective admissions are stratified according to long stays, short stays, and day cases (Table 3.4).

The increase in non-elective long stays is differentially larger than for non-elective short stays, reflecting increases in the mean LOS over the study period (Table 3.4).

Estimated total costs using the top-down approach also reflect the changes in the average unit costs applied across the study period. Unit costs have increased by a similar percentage for elective inpatient stays (32.88%) and non-elective long stays

Table 3.4: Changes in activity stratified for top-down costing 2009-10 to 2015-16 (primary only)

Activity	09-10	10-11	11-12	12-13	13-14	14-15	15-16	% Ch.
Elective inpatient stays	14	36	29	31	29	36	50	257.1
Non-elective inpatient stays (long)	2,131	2,325	2,429	3,069	3,652	3,644	3,841	80.2
Non-elective inpatient stays (short)	8,120	9,617	9,511	11,924	13,275	13,110	13,243	63.1
Day cases	828	0	4	0	1	2	1	-99.9
Total	11,093	11,978	11,973	15,024	16,957	16,792	17,135	14.1

(32%), but less for non-elective short stays (17.78%) and day cases (11.93%) (these costs were tabled in Chapter 2; see Table 2.5).

3.2.2 Comparing costing of activity based on primary diagnoses only with primary and secondary (any) diagnoses

3.2.2.1 Bottom-up costs: primary only compared with any diagnoses

Primary and secondary diagnoses of substance misuse codes are at times referred to as 'any' diagnosis for ease of explanation.

Inclusion of relevant activity on the basis of primary and secondary diagnosis codes leads to around ten times as many hospital admissions being included across the study period compared to primary only. The relative increase in activity is less for admissions including any diagnosis related to substance misuse – increasing 22.5% from 128,647 in 2009-10 to 157,591 in 2015-16; whereas for primary diagnoses the increase is 54.47% over the same period (Table 3.5).

Differences in the total costs according to the decision rule for including activity are primarily a function of the difference in the number/volume of admissions – but also reflect higher average costs per admission. Costs per admission are 30% higher in 2009-10 - £625.93 per admission for primary diagnoses only, and £843.58 per admission for any diagnosis. Furthermore, the cost per admission increases for any

Table 3.5: Primary only vs. any diagnosis - admissions and bottom-up costs

Year	Admissions		Total Cost (£)		Cost per admission (£)	
	Primary	Any	Primary	Any	Primary	Any
2009-10	11,093	128,647	6,943,487	108,524,556	625.93	843.58
2010-11	11,978	137,034	7,332,324	115,347,410	612.15	841.74
2011-12	11,973	142,075	7,166,555	124,418,726	598.56	875.73
2012-13	15,024	139,259	7,656,173	119,692,963	509.60	859.50
2013-14	16,957	151,238	9,060,531	133,416,639	534.32	882.16
2014-15	16,792	149,505	9,089,972	138,794,822	541.33	928.36
2015-16	17,135	157,591	9,790,258	152,851,494	571.36	969.93
% change	54.47%	22.50%	41.00%	40.85%	-8.72%	14.98%

Table 3.6: Admissions by elective vs non-elective activity (any diagnosis)

Year	Elective		Non-elective		Total
	N	%	N	%	
2009-10	4,359	3.39%	124,288	96.61%	128,647
2010-11	5,546	4.05%	131,488	95.95%	137,034
2011-12	7,183	5.06%	134,892	94.94%	142,075
2012-13	7,482	5.37%	131,777	94.63%	139,259
2013-14	8,486	5.61%	142,752	94.39%	151,238
2014-15	9,664	6.46%	139,841	93.54%	149,505
2015-16	10,502	6.66%	147,089	93.34%	157,591
% change	140.93%		18.35%		22.50%

diagnosis of substance misuse – by 14.98% to £969.93 in 2015-16; whereas the cost per admission decreases by 8.72% for primary diagnoses only to £571.36 in 2015-16.

This reflects a combination of factors – partly differences in the proportion of admissions that are elective and non-elective (Table 3.6). A higher share of admissions with any diagnosis of substance misuse are elective admissions – 5.37% in 2012-13 compared with 0.21% for primary diagnoses only. This share increases differentially more over the study period for any diagnoses (104%) compared with primary diagnoses only (140.93%).

Differences in the mean cost per admission also reflects differences in the distributional properties of costs – which are on average more variable for all diagnoses compared with primary only: the average ratio of the standard deviation to the mean across years is 1.39 for admissions for primary diagnoses only, and 1.58 for any diagnoses. The median values are relatively similar comparing Tables 3.7 and 3.3, which reflects the influence of a small number of very high cost admissions (which is

Table 3.7: Summary statistics by year - cost per admissions & LOS (all)

Year	Cost per admission (£) (bottom-up)					LOS				
	Mean	S.D.	Median	LQ	UQ	Mean	S.D.	Median	LQ	UQ
2009-10	843.58	1,122.58	509	382	679	1.81	5.67	1	0	1
2010-11	841.74	1,325	416	416	594	1.88	6.09	1	0	1
2011-12	875.73	1,372.37	428	428	611	1.88	6.06	1	0	1
2012-13	859.5	1,404.95	362	362	580	2	5.92	1	0	2
2013-14	882.16	1,446.72	451	451	585	2.03	5.84	1	0	2
2014-15	928.36	1,586.53	445	445	663	2.15	6.6	1	0	2
2015-16	969.93	1,559.58	452	446	663	2.23	6.39	1	0	2

also illustrated comparing the mean and median values).

Admissions for any diagnoses of substance misuse have a longer length of stay compared with primary diagnoses only. Mean LOS for any diagnosis is 1.81 in 2009-10; compared with 1.32 for primary only. Mean LOS increases differentially for any diagnosis – by 23.22% to 2.23 in 2015-16 compared with an increase of 9.09% to 1.44 for primary diagnoses only. The distributional properties of LOS also vary depending on the activity included – as shown by differences in the standard deviation and values of the upper quartile (Table 3.7).

Estimates of bottom-up costs of activity including any diagnosis of substance misuse also depend on how the activity is distributed across HRGs – and Figure 3b presents the four most common HRGs for admissions with any diagnosis of substance misuse in terms of volume and costs. The four most frequent HRGs in terms of admissions are the same as for primary diagnoses only - but represent much smaller proportion of overall activity, and this proportion decreases sharply over the study period.

Admissions with HRG codes WA11X for 'poisoning, toxic and environmental and unspecified effects with intermediate critical care' are the most frequent. There were 42,036 such admissions in 2009-10; rising by 27.69% to 53,679 in 2015-16. The average cost per admission fell from £563.40 in 2009-10 to £482.29 in 2015-16; dropped sharply to £383.82 in 2012-13 and rose subsequently to £462.97 in 2015-16. The share of total annual costs accounted for by activity with this HRG code falls over the period from 21.82% in 2009-10 to 16.26% in 2015-16.

The second most frequent HRG in terms of admissions for any diagnoses is WA11Y 'poisoning, toxic and environmental and unspecified effects without critical care' in 2009-10 and this remains the case until 2014-15 – when PA50Z becomes the second most frequent HRG for the included activity (Figure 3b). The number of admissions

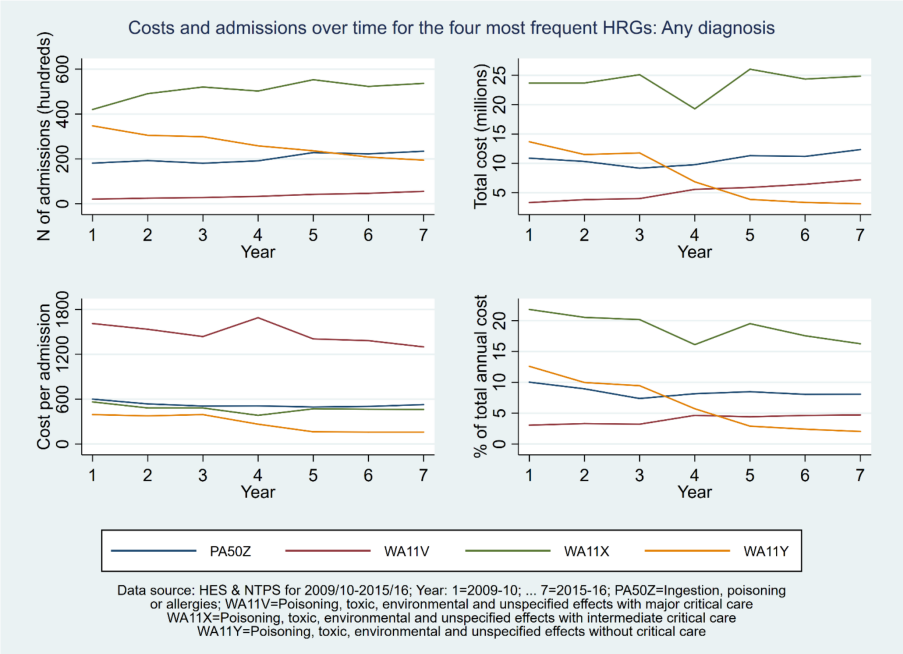
for HRG WA11Y falls 44.05% from 34,749 in 2009-10 to 19,441 in 2015-16. The average cost per admission is similar in 2011-12 (£393.75) as in 2009-10 (£393.73), but falls sharply in 2012-13 to £265.35. It then falls sharply again in 2013-14 to £163.66 – remaining relatively flat thereafter. These admissions comprise 12.61% of total annual costs in 2009-10 – falling sharply to 5.73% in 2012-13; and to 2.04% in 2015-16.

In 2009-10, the third most common HRG code in terms of total admissions is PA50Z – for 'ingestion, poisoning or allergies'. There are 18,124 such admissions in 2009-10, and this HRG becomes the second most common HRG code in 2014-15 when there are 22,240 admissions with any substance misuse related diagnosis assigned to this HRG. The average cost per admission decreases over the study period from £600.72 in 2009-10 to £495.60 in 2013-14; increasing to £527.19 in 2015-16. Admissions assigned to this HRG comprise a decreasing share of total annual costs over the study period – from 10.03% in 2009-10 to 8.08% in 2015-16.

Finally, the fourth most frequent activity sorted by HRG code is WA11V – for admissions with 'poisoning, toxic, environmental and unspecified effects with major critical care'. There are 2,058 such admissions in 2009-10, rising by 170% to 5,548 by 2015-16. These activities have the highest average cost within the four codes: the cost per admission decreases from £1611.65 in 2009-10 to £1299.13 in 2015-16. Nonetheless, these admissions comprise an increasing share of overall activity – rising from 3.06% in 2009-10 to 4.72% in 2015-16.

In terms of the total annual costs accounted for by these four HRG codes – the percentage decreases substantially from 47.52% in 2009-10 to 34.66% in 2012-13; and to 31.09% in 2015-16.

Figure 3b: Patterns over time for HRGs assigned most frequently to activity with a primary diagnosis of substance misuse



3.2.2.2 Top-down and bottom-up costs: primary only compared with any diagnoses

The scale of the estimated differences between the costing approaches changes over time according to the activity that is included – as outlined in Table 3.8.

Bottom-up total costs increase by a similar percentage comparing primary only (41%) with any diagnosis (40.85%) (Table 3.8). The reflects the fact that activity increases by a smaller percentage for admissions for any substance misuse related diagnosis (22.5%) than for primary diagnoses only (54.47%); which is offset by differences in the change in the mean cost per admission – an increase of 14.98% for all diagnoses compared with a decrease of 8.72% for primary only. In contrast, top-down total costs increase by a larger percentage (105.17%) for primary diagnoses only than for any diagnosis (86.06%).

Table 3.8: Comparison of activity included by costing approach

Year	Admissions	Total Cost (£)		Cost per admission (£)	
		Top-down	Bottom-up	Top-down	Bottom-up
Primary only					
2009-10	11,093	9,494,489	6,943,487	855.90	625.93
2010-11	11,978	10,811,799	7,332,324	902.64	612.15
2011-12	11,973	11,646,474	7,166,555	972.73	598.56
2012-13	15,024	15,153,414	7,656,173	1,008.61	509.60
2013-14	16,957	18,128,701	9,060,531	1,069.10	534.32
2014-15	16,792	18,527,920	9,089,972	1,103.38	541.33
2015-16	17,135	19,479,951	9,790,258	1,136.85	571.36
% change	54.47%	105.17%	41.00%	32.83%	-8.72%
Any diagnosis					
2009-10	128,647	120,562,177	108,524,556	937.15	843.58
2010-11	137,034	142,468,818	115,347,410	1039.66	841.74
2011-12	142,075	160,502,477	124,418,726	1129.70	875.73
2012-13	139,259	169,659,635	119,692,963	1,218.30	859.50
2013-14	151,238	196,288,072	133,416,639	1,297.88	882.16
2014-15	149,505	205,179,365	138,794,822	1,372.39	928.36
2015-16	157,591	224,322,734	152,851,494	1,423.45	969.93
% change	22.50%	86.06%	40.85%	51.89%	14.98%

3.2.3 Distribution of costs and admissions across population characteristics

3.2.3.1 Admissions and total costs

The median admission age is lower for admissions with a primary diagnosis of substance misuse than any diagnosis; and the median deprivation score is higher (Table 3.9). Admissions for primary diagnoses of substance misuse are disproportionately male (53.97%); whereas for any diagnoses females comprise a majority of admissions (50.93%).

Figure 3c shows the total number of costs and admissions by admission age according to whether primary or any diagnoses of substance misuse are included. Comparison between primary and any diagnosis reveals a considerable amount of similarity: there is a small but significant spike at very young ages (0-5 years old), followed by very small numbers of admissions between the ages of around 5 and 15 years old. There is a dramatic increase leading to a peak in admissions between the ages of 18 and 21 years old.

Table 3.9: Admissions by population characteristics

		Primary diagnoses			Any diagnosis		
		Mean	Median	S.D.	Mean	Median	S.D.
Admission Age		35.51	33	16.39	35.22	34	16.52
IMD Score		31.11	28.44	18.16	30.43	27.42	18.19
		N		%	N		%
Sex	Female	51,373		46.03	527,439		50.93
	Male	60,228		53.97	508,172		49.07
Ethnic Origin	Asian	2,187		1.96	26,812		2.59
	Black	1,518		1.36	19,657		1.9
	Mixed	1,131		1.01	12,188		1.18
	Other	2,085		1.87	23,383		2.26
	Unknown	7,808		7.00	60,090		5.8
	White	96,881		86.80	893,567		86.28

Figure 3c: Total admissions and costs according to age: primary vs any diagnosis

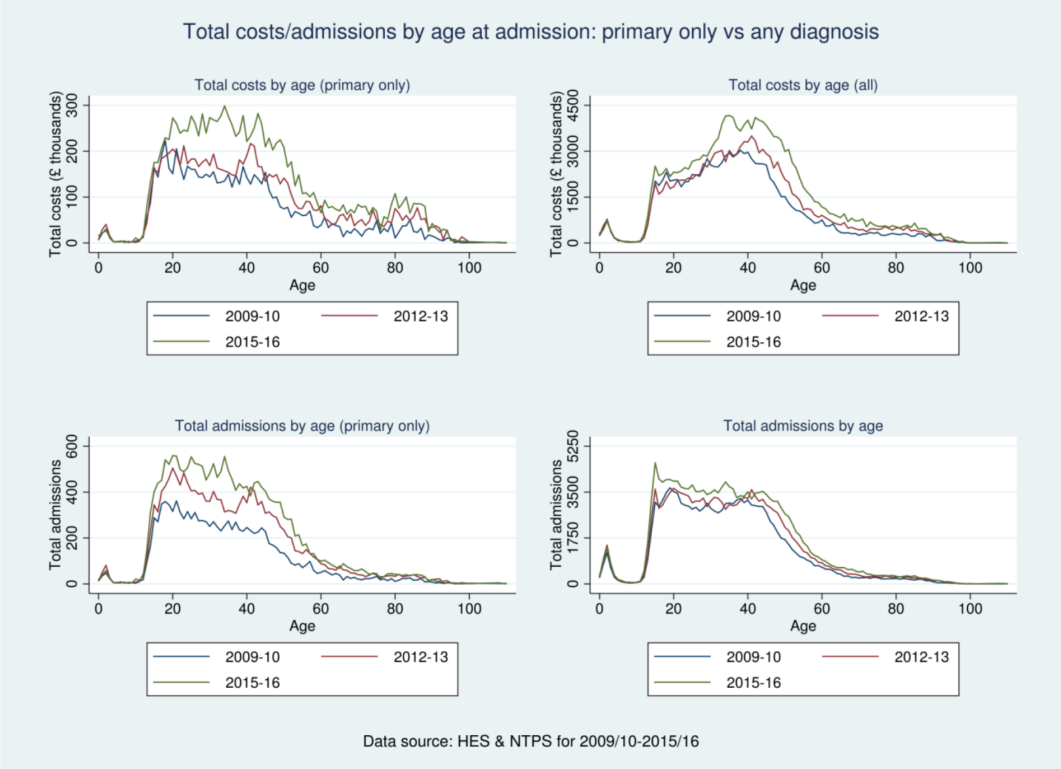
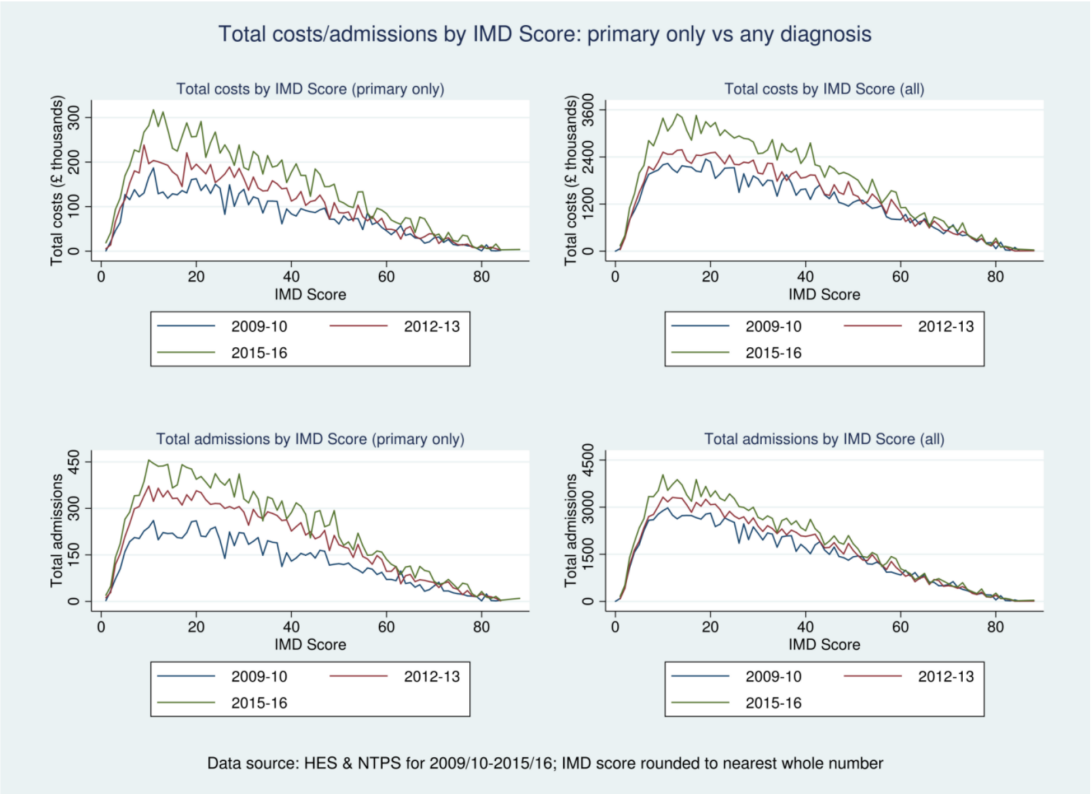


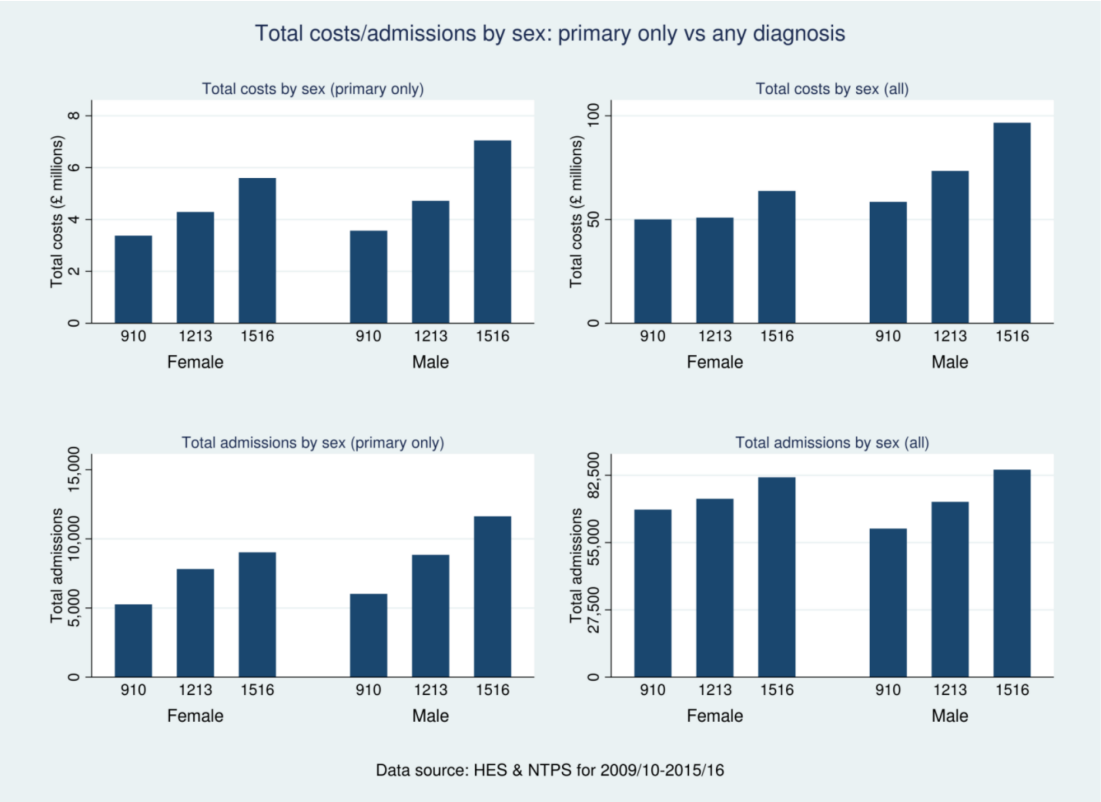
Figure 3d: Total admissions and costs according to IMD score: primary vs any diagnosis



However, there is a more pronounced decline between the ages of 20 and 40 for admissions for primary diagnoses only compared with a relatively flat gradient for all diagnoses. In fact, admissions for primary diagnoses of substance misuse decrease sharply between the ages of 35 and 60 – with the equivalent decrease compressed between the ages of 45 and 65 for any diagnosis of substance misuse. Both sets of activity have relatively few admissions at older ages. Profiles of total costs according to age at admission are relatively similar to the profile of admissions according to age for primary diagnoses – with the exception of older ages which show higher levels of total costs relative to the level of admissions. Patterns of total costs for any admission related to substance misuse are less similar than patterns for total numbers of admissions. Total costs increase between the ages of 18 and 45, and then decrease thereafter.

Total admissions according to IMD score (Figure 3d) simply reflect the distribution of admissions according the values of the IMD score (where higher values of the score represent higher levels of deprivation). Total costs are distributed similarly as total admissions according the IMD Score.

Figure 3e: Total admissions and costs by sex: primary vs any diagnosis



The disparity in total costs between males and females increases over the period, with males accounting for a differentially larger proportion of total costs by 2015-16 – regardless of the diagnoses used – although the disparity increases more for any diagnosis of substance misuse (Figure 3e). This is reflective of differential changes in activity – and these patterns are different depending on diagnoses used. There is a more pronounced sex difference for admissions containing primary diagnoses of substance misuse than there is for any diagnosis. Figures 3f and 3g show patterns of total costs and admissions by ethnic group. Regardless of the diagnoses used, these admissions comprise mainly white individuals – and there are many admissions with missing or null data on ethnicity.

Figure 3f: Total costs and admissions by ethnic group (non-white ethnic origin): primary vs any diagnosis

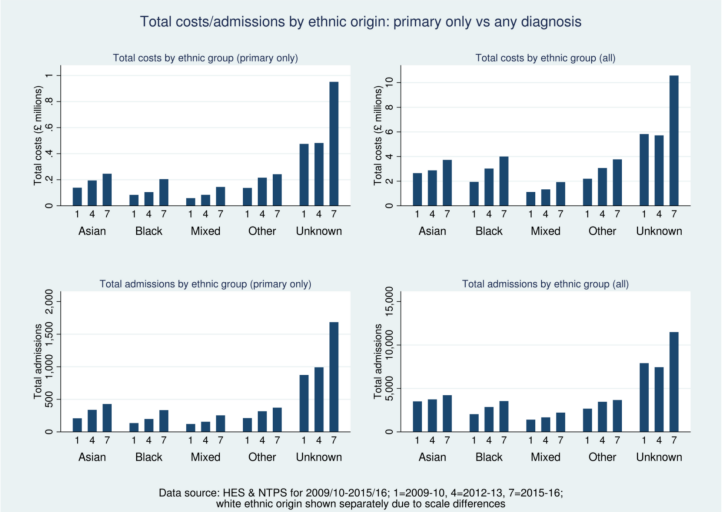


Figure 3g: Total costs and admissions by ethnic group (white ethnic origin): primary vs any diagnosis

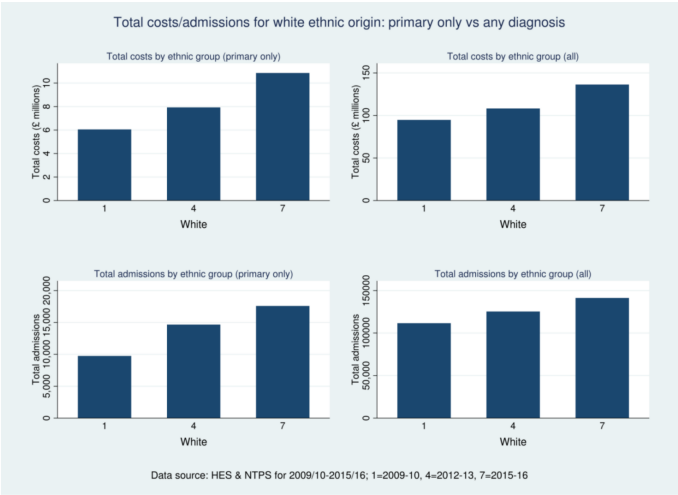
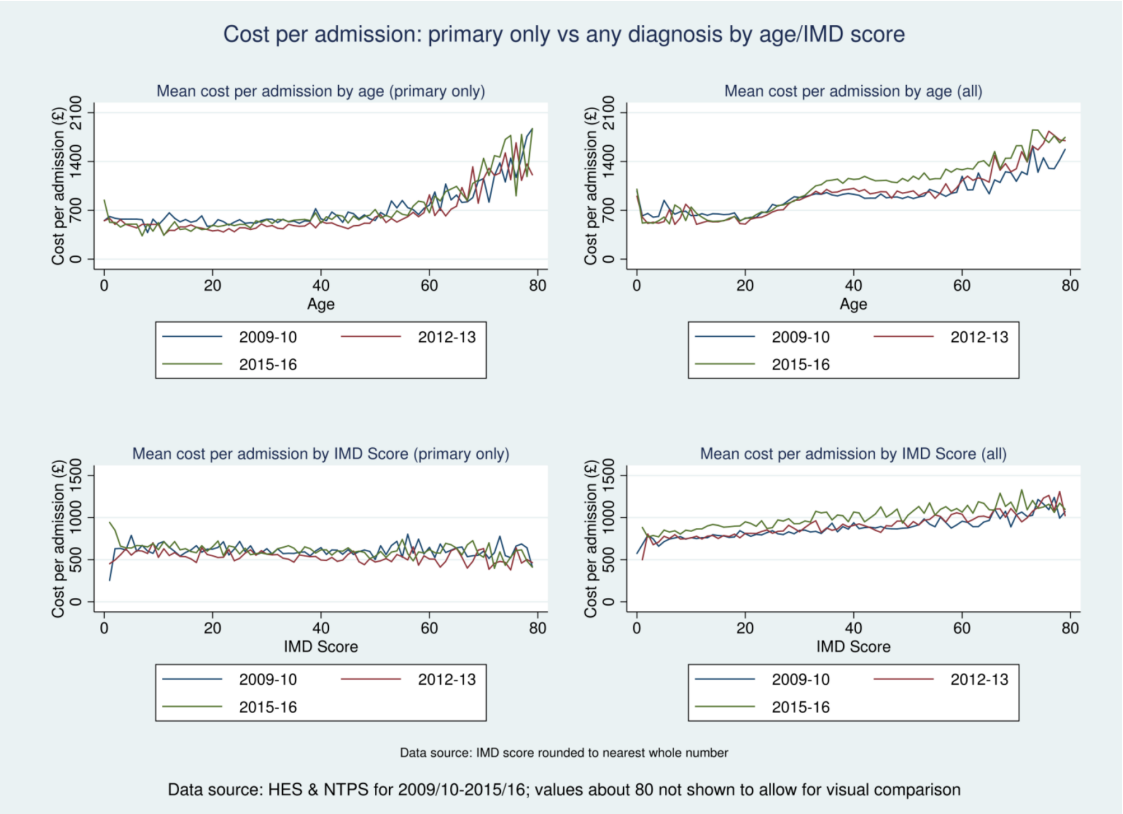


Figure 3h: Cost per admission according to age/IMD score: primary vs any diagnosis



3.2.3.2 Cost per admission

There is a differentially larger increase in the mean cost per admission for admissions containing any diagnosis of substance misuse between the ages of 20 and 40 than for primary diagnoses only - which are relatively flat between these ages (Figure 3h). The mean cost per admission is generally higher for any diagnosis of substance misuse over the age of 20 years old in general – primary diagnosis only exhibit very limited variation in average costs between the ages of 20 and 60. Admissions of both types are similarly expensive for ages greater than 80 years old; albeit the mean cost is noisier by age for primary diagnoses, reflecting the very small numbers of these admissions in older age. Mean costs by age are relatively similar in terms of changes over the study period.

Mean costs per admission are relatively similar by IMD score for activity with primary diagnoses of substance misuse, whereas the mean cost per admission increases slightly by the 2010 IMD score in an approximately linear manner for any diagnosis of substance misuse – that is, average costs are lower for admissions of residents of relatively more deprived areas.

Figures 3i and 3j show changes over the study period the average cost per admission by sex for both types of diagnoses. Admissions containing a primary diagnosis of substance misuse decrease between 2009-10 and 2012-13; and increase to a level slightly below their original level in females, and slightly above for males in 2015-16. However, the patterns over time differ by gender for admissions containing any diagnosis of substance misuse: for females, mean costs are lower in 2012-13 than in 2009-10 and then higher than at any other point in the final year 2015-16. In contrast, for men, mean costs per admission increase over the study period. Patterns by ethnic origin also show more uniform patterns in average costs (increases over the study period) for any diagnoses compared to primary diagnoses only (Figure 3k).

Figure 3i: Average costs per admission by sex (primary diagnoses)

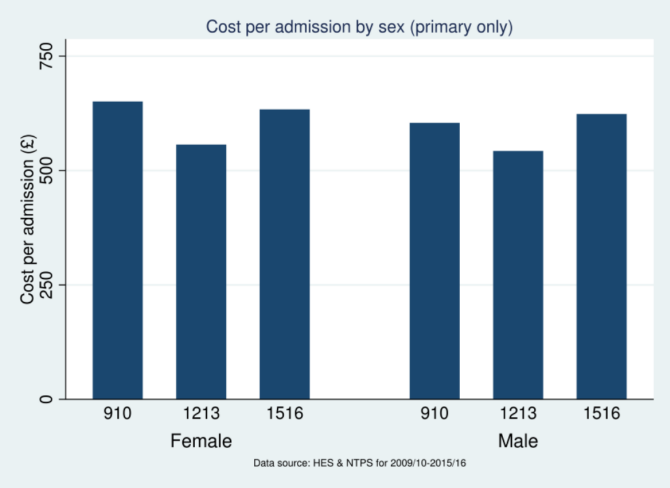


Figure 3j: Average costs per admission by sex (any diagnosis)

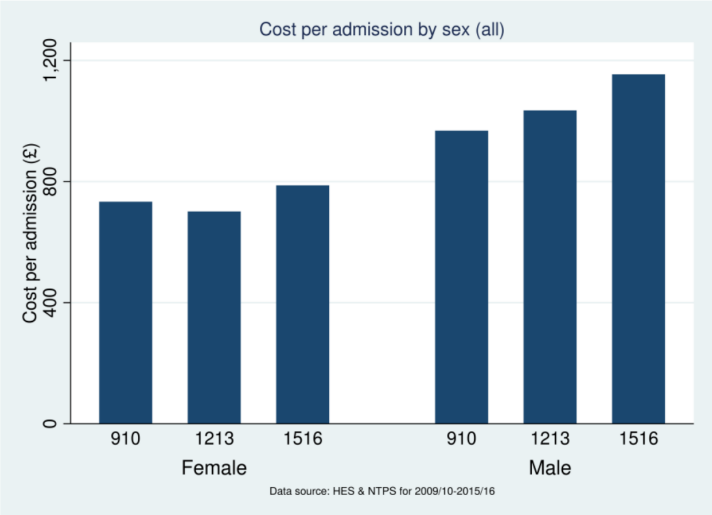
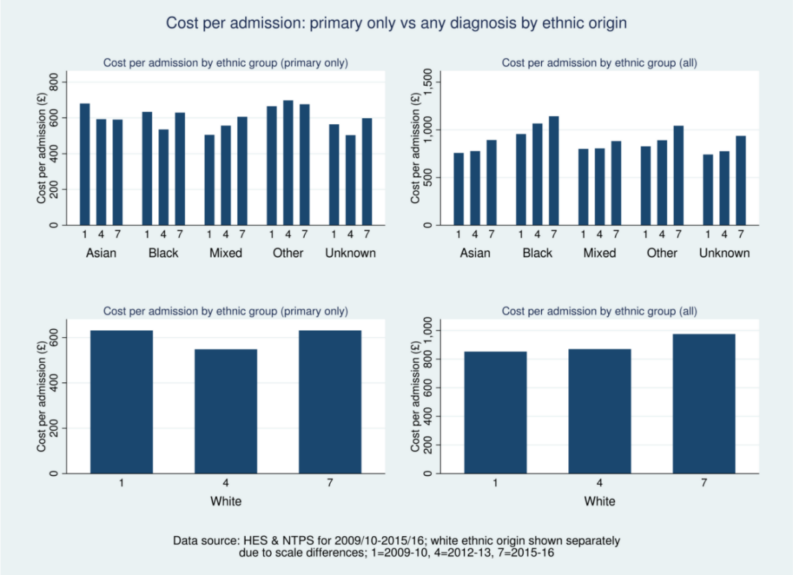


Figure 3k: Cost per admission according to ethnic origin: primary vs any diagnosis



3.3 Discussion

3.3.1 Summary of main findings

Hospital admissions with a primary diagnosis code indicating substance misuse have increased by 54.47% from 11,093 to 17,135 over the study period. Whilst there has been an 8.72% decrease in the cost per admission between 2009-10 and 2015-16, the rise in total admissions has meant that total costs have increased by 41% over the study period, from £6,943,487 to £9,790,258. Use of national average hospital costs for the same activity leads to estimated costs of £9,494,489 in 2009-10 and of £19,479,951 in 2015-16 – a 105.17% increase over the study period. The pronouncedly and differentially larger increase in top-down estimates of costs is attributable to a 32.83% increase in national average costs over the period.

Application of a version of top-down costing to this activity overestimated the costs of substance misuse related hospital admissions compared with bottom-up methodology – and the extent of the divergence increased over the study period. Over 99% of admissions including a primary diagnosis of substance misuse are non-elective, and the distribution of costs for these admissions exhibits marked skewness. Over 85% of all activity with a primary diagnosis of substance misuse is assigned to one of just four HRG codes used to determine hospital payment.

Extension of the definition of substance misuse related hospital activity to admissions including any diagnosis of substance misuse increases the estimated size of the related activity by a factor of around ten. Admissions with a primary diagnosis of substance misuse have increased by substantially more in relative terms over the study period than admissions where such diagnoses are in any position. However, the mean cost of admissions with a primary diagnosis of substance misuse has decreased over the study period – compared to increases for admissions with either primary or secondary diagnoses. Estimated total costs are higher by a factor of more than ten for any diagnosis of substance misuse. Activity including any diagnosis of substance misuse includes a higher proportion of elective admissions (around 5%), and longer mean LOS. Whilst the four most common HRG codes assigned to activity for any diagnoses are identical for primary diagnoses only, these HRGs represent a much smaller proportion of overall activity, and this proportion decreases sharply over the study period.

Admissions of primary diagnoses only are for a younger population overall than for

activity including any diagnosis of substance misuse, and there are more hospital admissions of individuals from more deprived areas regardless of diagnoses used. There are variations in the sex and ethnic profiles of the populations depending on diagnoses used.

3.3.2 Explanation of notable findings

Top-down costing approaches frequently apply average national costs – in some cases with some level of refinement – to disease specific activity. Critics of such approaches have noted that this often leads estimated costs to exceed budget allocations (Drummond 1992; Tarricone 2006). This illustration in this study was an application of national average costs to activity for substance misuse related hospital admissions. The estimated 'top-down' costs exceed the amounts that hospitals charge payers by a substantial margin. However, there are various potential explanations for this.

The first relates to differences in the demographic composition of the disease-specific population as compared to the overall population admitted to hospital. In the case of substance misuse, this population is relatively young (median age of 33 for primary diagnoses and 34 for any diagnosis) – compared with 53 years old in the total admitted population[3] (NHS Digital 2016b). There are also pronounced sex differences in the study population compared with all hospital admissions – the study populations contain a higher proportion of males compared with the overall admitted population which is 55.2% female[4].

Second, the average LOS of the study populations are substantially shorter than for all hospital admissions. Mean LOS in the overall admitted population was just under 5 days in 2015-16 (NHS Digital 2016b: 9); compared with between 1.4 and 2.3 days for the study populations in 2015-16. Estimates of top-down costs that do not incorporate precise measurement of the LOS and its distribution across activity will therefore produce misleading estimates – particularly if LOS is only incorporated using stratification of activity into elective, non-elective and day case admissions.

Differences in the underlying epidemiology of populations used to derive average costs drive differences in the distribution of activity across HRGs. This is especially important as there is considerable variation in the level of hospital reimbursement by HRG. These variations are not limited to the fixed component of the tariff (denoted

[3]Figures for 2015-16.

[4]This figure includes maternity care and with therefore overstate the difference

by π in equations 2.3 and 2.4), but also extend to an interaction with LOS in terms of the per diem rate for HRGs. This study demonstrated that, for the activity identified by the ICD codes used, over 85% was accounted for by 4 HRG codes for primary diagnoses only, and 47% at the study onset for any diagnosis – decreasing to 31% in 2015-16. The latter point also illustrates an explanation as to why the differences in estimated costs change over the study period when comparing the two approaches.

Restricting relevant activity to primary diagnoses for a particular condition only makes a material economic difference in terms of both the volume and average cost of the activity included. Inclusion of activity with any diagnosis of substance misuse includes a higher proportion of elective admissions, with longer LOS, and a much lower proportion of activity is assigned to specific HRG codes. The results on admission method, LOS and mean admission cost alongside demographic differences may be indicative of incorporation of activity from chronic effects of substance misuse in addition to acute activity in which substance misuse is the presenting condition.

3.3.3 Strengths and limitations

Two sources of costs were used in this study for the bottom-up and top-down costing applications. This represents an important limitation for this study and the findings should be interpreted in that context. The NTPS was used to estimate 'bottom-up' costs in order to accurately reflect the costs to commissioners of health care i.e. CCGs (previously PCTs). As such, this aligns with allocations from (and therefore the cost to) national government. The bottom-up approach matches the relevant activity with year and HRG-specific price algorithms that are used to calculate hospital reimbursement based on a several factors such as LOS and admission method. This costing approach was set out in detail and informs the costing approach taken for other studies in this thesis. Bottom-up costs were then compared for activity for primary diagnoses only substance misuse only and any diagnoses of substance misuse.

The principles of top-down costing were also illustrated in an application that used the NHS reference costs. The rationale for using the NHS reference costs in the illustration is as follows: top-down costing studies often apply disease-specific average costs to counts of activity. However, there are no comparable disease-specific costs for the general definition of substance misuse. The costs that were used were therefore just for average hospital costs (sourced from the NHS reference costs). The NTPS are HRG-specific prices, and as such do not have disease-specific average costs that

can be applied; and the combination of the above factors meant that the average cost figures from the NHS reference costs were used.

This difference in the cost sources used in the top-down and bottom-up approaches will be a partial driver of the differences in estimates of total costs and is an important limitation to consider when comparing the two sets of estimates. The NTPS are a set of HRG-specific prices, whereas the reference costs are based on self-reports of NHS providers from annual collections. The two sets of figures are related but distinct: adjustments are made the average of the reference costs in the construction of tariff prices. The reference costs used to produce the tariff are from three years previous, and as such an uplift is applied which reflects pay and other cost pressures in the NHS (DH 2012). There has also been an incorporation of an efficiency requirement for providers to reduce their costs in the prices that are set in the tariff.

Costs in this study are not deflated - they reflect inflation which partly explains trends over time. As such, the changes over time can be interpreted as reflecting nominal rather than real changes in costs.

This chapter is not intended to be a replication of a COI study. It is a demonstration of different approaches to costing applied to the disease area of interest in this thesis: substance misuse. The specific comparisons made are: top-down and bottom-up costing; and using primary/any diagnosis of the condition. As such, it is motivated by insights from the COI literature, and illustrates the potential materiality of the differences in these approaches. It also sets out in detail the costing methodology applied in the remainder of this thesis: bottom-up costing of any activity with any ICD-10 diagnosis. It is beyond the intended scope of this chapter to estimate the full cost of illness for substance misuse.

This study dropped data from the original sample of episodes identified as containing a diagnosis related to substance misuse for various reasons, outlined in detail. The removal of these data likely leads to an underestimate of the number of substance misuse related admissions and their costs. One such issue included the approach taken to spells that included unfinished episodes of care. This differentially impacted on the later years of the data, for which there is a much higher proportion of activity affected by unfinished care. This partly also reflect a limitation inherent in linkage of the data used for reimbursing providers, which in itself depends on the recording of providers in hospital. This study excluded maternity admissions and admissions assigned to mental health HRGs from the study sample. Differences in estimated costs are not attributable to these exclusions since the alternative costing approaches

were performed on identical sets of admissions. The omissions do however mean that the costing herein likely understates the true cost of admissions related to substance misuse, given that around 5% of all admissions containing any diagnosis of substance misuse is represented by such activity. In the case of admissions assigned to mental health HRGs, this activity represents a growing share of overall activity over the study period.

This study also only considers hospital activity in the 'admitted patient care' data in HES. It therefore does not include information on accident and emergency (A&E) attendances or outpatient visits. Whilst this does not affect comparisons of costing approaches, estimated costs will therefore understate the true hospital costs related to substance misuse.

Identification of relevant activity relies on use of ICD diagnosis codes. This may first affect the study findings insofar as activity included may simply reflect variations in coding practices – in some cases hospitals may only suspect drugs misuse and therefore not record it. Furthermore, underlying substance misuse may not be detected in hospital and therefore fail to be incorporated in diagnosis coding despite being an underlying factor causing admission to hospital. Second, there have been improvements in the recording of secondary diagnoses over since 2006-07 – and as such increase in the number of admissions for any diagnosis of substance misuse over time may reflect at least in part improvements in diagnoses.

Use of primary diagnosis codes only likely understates the activity attributable to substance misuse (Ward et al. 2005). Equally, use of secondary and primary codes likely overstates attributable activity - as the activity included contains other co-morbidities which cause the hospital admission and costs. This reflects a limitation inherent in approaches that rely on diagnoses codes alone.

Whilst there is no consensus on which ICD codes identify substance misuse related activity, inclusion of some X, Y and T codes may lead to inclusion of activity relating to adverse drug events and polypharmacy. Admissions at very young ages and older ages may in particular reflect this.

4 Exploring patterns of hospital activity and costs relative to admission with a diagnosis of substance misuse: an event study.

4.1 Introduction

The previous two chapters of this thesis have outlined that substance misuse represents a considerable public health challenge for policymakers internationally. Healthcare costs, along with crime costs and productivity losses, are consistently identified in international research as one of the principal contributors to the costs of substance use (ONDCP 2004; Harwood et al. 1999; Baumberg 2006; Wickizer 2013; McCollister et al. 2003; Olsson & Fridell 2015). Despite this, relatively few studies have investigated the role of substance dependence in healthcare utilization or its associated cost (Compton et al. 2007; Santora & Hutton 2008; Olsson & Fridell 2015). The evidence base suggests that drug users are admitted to hospital more frequently than the general population (Stein et al. 1993; Haapanen-Niemi et al 1999; French et al. 2011).

In the case of substance misuse, drug users are admitted to hospital for various reasons – typically the direct effects of substance use (for example, overdose and intoxication) or the medical complications (Haber et al. 2009). For example, injecting drug users are admitted for: soft-tissue infections, endocarditis, injuries sustained during intoxication, violence related to substance misuse, and psychiatric complications such as substance-induced psychosis or suicidal ideation (Haber et al. 2009). For elective admissions specifically, the most frequent causes are bacterial infections related to drug injection such as cellulitis and soft-tissue infections (Stein & Sobota 2001; Binswanger et al. 2000).

Previous studies which have considered the impact of substance misuse on measures of the use of hospital resources typically fall into two categories: prospective cohort studies and studies that observe longitudinal trends.

In the first of these (prospective cohort studies), consenting participants with substance misuse problems are followed as part of a dedicated study (Davies et al. 2009;

Jones et al 2009; Palepu et al. 2001; Solomon et al. 1991; Palepu et al. 1999; Stein et al. 1993; Baldwin et al. 1993; Parthasaray & Weisner 2005). These studies often use questionnaires relating to use of various services to quantify the use of services in the population of interest; following a sample of users for a period in which periodic information is gathered including aspects of health care use including primary and secondary care. In some cases, participants' health records are linked in order to eliminate issues around recall bias and improve accuracy in estimating the amount of use (Palepu et al. 2001; Olsson & Fidell 2015).

Davies et al. (2009) followed a cohort of individuals in treatment for substance misuse and used questionnaire information relating to the number of days (where appropriate) of use for particular health services at baseline and follow up to estimate the profile of service use and related costs over time in the treatment population. The study found that treatment was associated with reductions in drugs use, offending, harmful health behaviours, crime and improvements in social and mental functioning. Palepu et al. (2001) linked the hospital records of participants in a prospective cohort study of injecting drug users to estimate patterns of use over time and associate variations in use with particular characteristics (such as frequency of cocaine use). They found that, out of 598 injecting drug users, 265 presented to emergency departments three or more times, and 118 were admitted more than twice over 39 months. The most frequent sources of emergency care in this cohort are soft-tissue infections, overdose, intoxication and withdrawal (Haber et al. 2009). Olsson & Fridell (2015) followed 227 women remanded to compulsory care in Sweden and linked their hospital use from administrative records, and applied unit costs at the individual level to estimate costs of hospital care over a thirty year period. Findings indicated that the cohort had five to six times as many hospital admissions as the general population, and stays that were six to eight times that of the general population.

Prospective cohort studies which consider out-of-treatment drug users are unusual. Most studies in the literature only consider treatment populations and do not compare to the general population, which exhibit markedly different underlying characteristics (Hollander 1995; French et al. 2000). As an exception to this, French et al. (2000) recruited injecting drug users (IDUs), chronic drug users (CDUs) and non-drug users and estimated annual differences in the number of hospital admissions, the number of outpatient visits and the number of emergency room episodes. They found that IDUs and CDUs used more emergency services and had more hospital admissions overall than non-drug users, but used fewer outpatient services. The cost of services was over $1,000 more in drug users relative to non-drug users (French et

al. 2000: 1710).

The second type of study considers longitudinal trends in the use of particular services related to substance misuse using observational data (Santora & Hutton 2008; Schoenbaum et al. 1998; Stein 1994; John et al. 1999; Lange & Schacter 1989; Marik & Mohedin 1996; Orford et al. 1992; Taylor et al. 1986). These studies have shown that individuals with substance misuse problems have a high prevalence of medical and psychiatric illness (Buckley & Brown 2006; Hasin et al. 2007), are frequently sick and are admitted to hospital at a much higher rate compared to the general population. They also show how admissions and costs vary depending of specific characteristics such as human immunodeficiency virus (HIV) status and IDU status (Stein et al. 1993).

Santora & Hutton (2008) assessed longitudinal trends in hospital admissions related to addiction for the period 1994-2002 and found that admissions with co-occurring substance misuse problems increased by 50% over the analysis period and that the consequent costs amongst this group increased by 134%. Overall, longitudinal studies of hospital admissions related to substance misuse using large sample sizes are rarely conducted (Santora & Hutton 2008; Compton et al. 2007; Hoff & Rosenheck 1998; Piette et al. 1998; French & Drummond 2005). Furthermore, longitudinal studies often tend to seek to identify generalised trends over time at the population level (Giordano et al. 2014), rather than assessing patterns at the individual-level.

These longitudinal studies have commonly had short time frames (one year or less) and often rely on self-reported data on use (Olsson & Fridell 2015). They represent alternative approaches that can be used to identify the impact of a particular condition on costs.

The previous chapter of this thesis outlined traditional cost of illness approaches and explored their strengths and limitations. An alternative to identification of activity related to particular disease based on what is indicated by either primary or primary and secondary diagnosis codes is set out by Ward et al. (2000). It entails an incremental analysis of costs based on comparison of hospital activity of an age and gender matched cohort with and without the condition. The strength of such an approach is that it attempts to work out the portion of cost attributable to the specific condition in cases where these costs are the result of a number of comorbid conditions. However, the limitations of this approach, as set out by Jo (2014) is that it assumes that other factors must not affect the association between conditions. In other words, the matched cohort with and without the condition could

be unobservably different.

Despite the limitations of various approaches, epidemiological data that describe use of health services and costs over time in particular populations are relevant for policymakers for several reasons. First, understanding patterns in resource use can aid policymakers and decision makers in how to target, deliver and improve services (see for example Poikolainen (1982)). Second, this information can allow for enhanced understanding of the potential benefits of effective intervention for specific groups of substance misusers at specific time points. Third, information on use of services and subsequent costs are of interest to studies which only consider impacts on treatment outcomes (Olsson & Fridell 2015). Fourth, results from observational data aid the formulation of future studies (French & Drummond 2005). Finally, as Shah et al. (2011: 779) state, epidemiological evidence relating to substance misuse is required for guiding legislative decisions relating to the substances:

> "there is a public health need for epidemiological surveillance of morbidity due to acute recreational drug toxicity, in order to guide service provision and legislative decisions regarding the classification of both established and novel recreational drugs."

As set out, there are a range of approaches that can be used to generate this evidence on the use of health services in disease specific populations.

This chapter seeks to explore an alternative approach using an econometric method that uses the timing of admission related to a specific disease, in this thesis substance misuse, to understand how patterns and costs of service use change as a result of particular health conditions. The exploration of this method in this chapter will assess whether it has the potential to complement (or act as an alternative to) the existing approaches that have been more typically used to identify a diseases' impact on the use of resources.

Understanding whether there is a change in the use of hospital services once the use of substances has been diagnosed as being an underlying cause of admission will additionally provide evidence as to the potential benefits of intervention that is timely and effective enough to prevent admission to hospital where substance use is an underlying cause.

It will also show whether there is a predictable pattern of service use that is indicated by the occurrence of a hospital admission caused by substance misuse. An occurrence

could potentially be used as a flag for: first identifying substance misuse status for individuals; and then directing individuals towards treatment services and other public services (such as primary care, social services, and so on).

Comparably, Wiley et al. (2013) demonstrated the effectiveness of ICD-9 codes in identifying smoking status in a clinical population and concluded that ICD-9 code histories were a suitable substitute for ascertaining smoking status in the absence of full-text records.

The aim of this study is to characterise patterns of use of hospital inpatient services in England for individuals between 2009-10 and 2015-16 who are observed to have at least one substance misuse related hospital admission during the analysis period using an 'event study approach'. Specific questions are listed below:

1. Do individuals have relatively more hospital admissions after their first occurrence of being admitted to hospital for a diagnosis of substance misuse?
2. Do individuals have relatively higher costs and relatively more days in hospital after their first occurrence of being admitted to hospital for a diagnosis of substance misuse?
3. Do individuals have relatively more emergency admissions after their first occurrence of being admitted to hospital for a diagnosis of substance misuse?
4. What are the patterns of hospital resource use prior to an individual's first occurrence of being admitted to hospital for a diagnosis of substance misuse?
5. Are changes in hospital admissions after the first substance misuse related admission differentially higher for admissions that are substance misuse related compared with admissions that are not related to substance misuse?
6. Do the types of admissions that individuals have after their first occurrence of being admitted to hospital for a diagnosis of substance misuse tend to result in higher costs and increased length of stay?

4.2 Results

4.2.1 Summary statistics

Tables 4.1 to 4.3 summarise the dependent and independent variables used in the regression analyses herein. The study sample comprises N=3,233,056 observations on n=472,170 individuals for the period 2009-10 to 2015-16.

The mean annual cost of admissions is £1,504.73, with a standard deviation of £5,473.39 (Table 4.1). This reflects the skewness previously shown which is consistent with many health care and hospital related variables – a small minority of cases comprise the majority of costs. Each of the dependent variables exhibits marked skewness, although the degree varies slightly across the indicators.

The mean number of days spent in hospital per year for an individual is 4.62; the standard deviation is 21.46 with the longest stay equal to a whole year spent in hospital.

In terms of counts of hospital admissions, the mean is equal to 0.94 overall and 0.57 for emergency admissions only. For admissions without a drug-related diagnosis, the mean is 0.71 per person per year with a standard deviation of 3.38; and individuals have on average 0.23 drug-related admissions per year over the study period (with standard deviation of 0.68).

Table 4.2 shows the demographic breakdown of the sample: The average age is 36.88 and 52% of the sample are male. The sample is predominantly white (89.35% of individuals), with 3.36% of the sample comprised of individuals of Asian ethnic origin (Table 4.2).

I show the breakdown of observations by event time and year in Table 4.3. The observations are normally distributed around an event time mean of zero. Fewer observations at the low and high values of event time are the natural consequence of the design of the variable (which allows for 'treatment' to vary by individual and indexes time around the occurrence). Attrition due to death is observed over time as shown by decreasing numbers of observations as year increases.

Table 4.1: Summary statistics for dependent variables

Variable (per year)	Mean	Standard Deviation	Min	Max
Total cost of admissions	£1,504.73	£5,473.39	0	£619,790
Total LOS	4.62	21.46	0	365
Number of admissions	0.94	3.50	0	316
Number of emergency admissions	0.57	1.42	0	161
Number of non drug-related admissions	0.71	3.38	0	316
Number of drug-related admissions	0.23	0.68	0	159

N=3,233,056; n=472,170 (individuals)

Table 4.2: Summary statistics (population characteristics)

Categorical variables (ethnic origin)		
	N	%
Asian	108,764	3.36
Black	96,179	2.97
Mixed	41,888	1.3
Other	97,591	3.02
White	2,888,634	89.35

Continuous variable (age, proportion male)				
	Mean	SD	Min	Max
Age in 2009-10	36.88	16.58	16	95
(Age in 2009-10) squared	1,634.97	1,535.64	256	9,025
Male	0.52	0.49	0	1

N=3,233,056; n=472,170 (individuals)

Table 4.3: Observations by event time and year

Event Time	Frequency	%	Year	Frequency	%
-6	67,981	2.1			
-5	137,346	4.25			
-4	213,326	6.6	2009-10	472,170	14.6
-3	291,265	9.01	2010-11	472,167	14.6
-2	378,269	11.7	2011-12	469,813	14.53
-1	472,153	14.6	2012-13	465,433	14.4
0	472,153	14.6	2013-14	459,269	14.21
1	391,830	12.12	2014-15	451,805	13.97
2	317,005	9.81	2015-16	442,399	13.68
3	240,780	7.45			
4	165,997	5.13			
5	84,946	2.63			
Total	3,233,056	100	Total	3,233,056	100

4.2.2 Regression analyses

Table 4.4 outlines the regression output for the preferred count model used (negative binomial) for each of the six outcomes in this study. The preferred count data model was selected based on the significance of the dispersion parameter in negative binomial regression. The narrative of the results interprets the preferred count data models (Table 4.4).

Across all models, the reference categories are as follows: the final event year (event time=5); 2015-16 for the year dummies; females; the fifth IMD deprivation decile; and White ethnic origin.

Comparisons across count models for each outcome are set out in the Appendix Tables A4.1 to A4.5. In these Appendix Tables, the first two models show the coefficients for OLS estimated on the inverse hyperbolic sine transformation of the dependent variable (the second includes individual random effects). The third and fourth models use Poisson regression (and again the second includes individual random effects). The fifth model uses negative binomial regression. These count models are included to show differences in model performance for overdispersed dependent variables.

Coefficient plots (Appendix Figures A4a to A4e) provide a visual comparison of the path of dependent variables over event time and the similarity and/or differences between modelling approaches whilst controlling for the effect of the covariates.

Marginal effects are also estimated, giving the predicted value of each dependent variable by event time. These are computed at the following values of the control variables: year set equal to 2012-13 (the 'middle' year); IMD decile set equal to 5 (a 'middle' decile); age (in 2009-10) set at 35 years old (reflecting mean age); and for males (the majority sex).

4.2.2.1 Analyses of costs

Costs are shown to be increasing in age – with costs 1.027 times higher for each additional year of age ($p<0.001$) (Table 4.4). The squared age term is not significant, showing that there is no evidence of a quadratic trend in age. Costs for males are shown to be at 0.845 times the level for females ($p<0.001$). There is a clear gradient in deprivation – with the least deprived deciles having lower costs than the more deprived deciles compared with the 5th most deprived – except for the most

deprived decile. Costs are shown to be substantially higher for black (IRR=1.543; p<0.001), other (IRR=1.071; p<0.001) and mixed (IRR=1.167; p<0.001) ethnic groups compared with individuals of white ethnic origin. Estimates on the year effects show that costs are highest for the final year of the period (the reference year 2015-16).

Overall, the estimates on the event time dummies show a consistent trend and are estimated with precision. Compared with the base year (event time=5), costs are lower in the pre-event period – with costs at less than half the rate in the base year for 6 and 5 years before the occurrence, as demonstrated by incidence rates of 0.458 and 0.495 for these years respectively (p<0.001). These costs rise steadily in the pre-event period to 0.833 times the rate in the base in the year prior to the event (p<0.001). Costs then jump to 2.788 times the rate in the base year in the year of the event (p<0.001); before decreasing in the year after the event to a rate considerably higher in the pre-period but much lower than in the year of the occurrence (1.323; p<0.001). These costs remain higher in the post-period, declining steadily to the base level which equals one by definition in the base year.

Table 4.5 displays the predicted values of costs at the range of values of event time. Costs are shown to be rising from their previous level to £839.62 two years before the event [CI £811.71; £867.53]; and to £1,055.92 in the year before the event [CI £1021.98; £1089.86] – which is almost twice the level six years prior to the event (ME=£580.84 [CI £524.60; £619.09]). These costs fall considerably in the year after the event (ME=£1,677.69 [CI £1628.15; £1727.23]); and steadily decrease in the years after to £1267.86 [CI £1219.68; £1316.03]. There are relatively small differences between the predicted values of costs comparing count models.

Table 4.4: Preferred regression models (all outcomes)

	Costs	LOS	Admissions	Emergency Admissions	Non-drug admissions	Drug-related admissions
-6	.458***	.348***	.549***	.376***	.686***	
-5	.495***	.398***	.574***	.406***	.717***	
-4	.529***	.466***	.604***	.440***	.754***	
-3	.571***	.509***	.649***	.494***	.809***	
-2	.662***	.591***	.723***	.582***	.901***	
-1	.833***	.765***	.869***	.764***	1.079**	
Event=0	2.788***	3.101***	2.803***	3.596***	1.515***	8.451***
1	1.323***	1.494***	1.262***	1.329***	1.248***	1.391***
2	1.148***	1.222***	1.111***	1.137***	1.122***	1.118***
3	1.069***	1.126***	1.042**	1.053***	1.055**	1.032
4	1.027*	1.039	1.011	1.015	1.027	0.97
2009-10	.896***	1.02	.974*	1.039***	0.975	.387***
2010-11	.971**	1.056***	0.998	1.037***	0.996	1.010**
2011-12	.932***	1.037**	1.014**	1.028***	1.008	1.019***
2012-13	.724***	0.991	1.021***	1.015***	1.025***	1.002
2013-14	.821***	.977*	1.019**	0.999	1.033***	.985**
2014-15	.717***	.810***	1.012	.988*	1.044***	.959***
Age in 0910	1.027***	1.022***	1.023***	1.017***	1.026***	1.016***
(squared)	0.999	1.000***	.999**	0.999	0.999	.999***
Male	.854***	.980*	.798***	.935***	.736***	1.040***
IMD 1	.739***	.682***	.789***	.725***	.785***	.851***
IMD 2	.862***	.815***	.906***	.829***	.908***	.900***
IMD 3	.924***	.868***	.949**	.909***	.955*	.933***
IMD 4	.965*	0.982	0.979	.976*	0.979	.975*
IMD 6	1.039*	1.012	1.027	1.038**	1.028	1.018*
IMD 7	1.043**	1.004	1.036*	1.079***	1.029	1.045***
IMD 8	1.051**	1.028	1.050**	1.098***	1.043	1.063***
IMD 9	1.007	.959*	1.02	1.095***	0.997	1.084***
IMD 10	.897***	.794***	.947*	.965**	.913***	1.042***
Black	1.543***	2.448***	1.383***	1.071*	1.500***	.933***
Asian	1.031	1.281***	1.076*	.910***	1.141***	.833***
Other	1.071***	1.359***	1.007	1.008	1.029	.934***
Mixed	1.167***	1.799***	0.988	.910***	0.986	0.976
Constant	583.759***	1.543***	.404***	.259***	.287***	.109***
lnalpha	16.684***	11.695***	1.816***	1.452***	3.457***	.038***
N			3,233,056			
AIC	29,628,507	9,524,750	7,999,097	5,975,823	6,719,800	2,349,135
BIC	29,628,961	9,525,205	7,999,552	5,976,277	6,720,254	2,349,589

*p<0.05; **p<0.01; ***p<0.001; all models estimated using Negative Binomial regression; incident rate ratios displayed

Table 4.5: Marginal effects (annual cost of admissions)

Event Time	ME	95% CI	
-6	£580.84	£542.60	£619.09
-5	£627.77	£594.88	£660.67
-4	£671.24	£641.82	£700.65
-3	£724.14	£696.31	£751.97
-2	£839.62	£811.71	£867.53
-1	£1,055.92	£1,021.98	£1,089.86
0	£3,534.49	£3,435.78	£3,633.21
1	£1,677.69	£1,628.15	£1,727.23
2	£1,455.31	£1,411.05	£1,499.58
3	£1,355.44	£1,312.05	£1,398.82
4	£1,301.83	£1,257.15	£1,346.51
5	£1,267.86	£1,219.68	£1,316.03

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1.

4.2.2.2 Length of stay[5]

The next set of models were estimated for annual length of stay (Table 4.4). The broad patterns and scale of the results reflect those estimated for annual costs.

Length of stay is shown to be increasing in age, with 1.022 ($p<0.001$) times as many days for each year of age. Men have 0.98 as many bed days as women ($p<0.05$). Less deprived deciles have fewer bed days than the reference group, and individuals of black ethnic origin have 2.448 times as many bed days as white individuals ($p<0.001$). Individuals' bed days are shown to be increasing in the pre-event period. They are over double the level in the pre-event year (IRR=0.348; $p<0.001$) compared to six years prior (IRR=0.765; $p<0.001$). There are 3.101 times as many bed days in the event year compared with the reference year ($p<0.001$), falling to 1.494 times in the year after the event ($p<0.001$). Thereafter, the event time estimates steadily decline until the reference year (2015-16).

These differences are illustrated by the marginal effects shown in Table 4.6. Bed days rise from 1.198 six years prior to the event [CI 1.102; 1.295]; to 2.634 in the year before the occurrence [CI 2.533; 2.734]. Individuals spend 10.68 days in hospital in the event year [CI 10.315; 11.045]; 5.146 in the following year [CI 4.959; 5.333], before steadily declining to 3.443 days five years after the event [CI 3.253; 3.634]. These patterns are consistent across models although there are (relatively small) scale differences.

[5]Bed days and LOS used interchangeably.

Table 4.6: Marginal effects (total LOS per year)

Event Time	ME	95% CI	
-6	1.198	1.102	1.295
-5	1.369	1.283	1.456
-4	1.603	1.516	1.69
-3	1.753	1.671	1.834
-2	2.036	1.95	2.122
-1	2.634	2.533	2.734
0	10.68	10.315	11.045
1	5.146	4.959	5.333
2	4.207	4.046	4.368
3	3.876	3.71	4.042
4	3.579	3.41	3.749
5	3.443	3.253	3.634

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1.

4.2.2.3 All hospital admissions

The next set of results presented is for the annual number of hospital admissions per individual as the dependent variable (Table 4.4). The same broad patterns are observable in the control variables as for the previous two outcomes. There are 1.023 times as many admissions for each year of age ($p<0.001$); and there is a quadratic effect in age (IRR=0.999; $p<0.01$). Men have 0.798 times as many admissions as women ($p<0.001$); and the least deprived are less frequently admitted to hospital. Black individuals are admitted 1.383 times as often as individuals of white ethnic origin ($p<0.001$).

In terms of the estimates on event time, the broad pattern is consistent across modelling approaches and is consistent with the patterns seen for the two previous dependent variables: individuals have 0.549 times as many admissions six years prior to the event compared to the reference year ($p<0.001$). This increases to 0.869 times as many admissions one year before the event ($p<0.001$), followed by 2.803 times as many in the year of the occurrence ($p<0.001$). This decreases to 1.262 as many in the year following the event ($p<0.001$); and steadily decreasing thereafter.

The models translate the estimated effects into predicted values as shown in Table 4.7: 0.372 admissions per year six years before the event [CI 0.349; 0.394]; 0.587 one year before [CI 0.569; 0.606]; 1.895 in the year of the event [CI 1.839; 1.950] before steadily declining whilst remaining above the level prior to the event at 0.676 five

Table 4.7: Marginal effects (number of hospital admissions per year)

Event Time	ME	95% CI	
-6	0.372	0.349	0.394
-5	0.388	0.368	0.407
-4	0.408	0.39	0.426
-3	0.439	0.422	0.455
-2	0.489	0.472	0.506
-1	0.587	0.569	0.606
0	1.895	1.839	1.95
1	0.853	0.826	0.88
2	0.751	0.726	0.775
3	0.704	0.679	0.729
4	0.684	0.656	0.711
5	0.676	0.644	0.708

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1.

years later [CI 0.644; 0.708].

4.2.2.4 Emergency admissions

Models were then estimated on the count of emergency admissions (Table 4.4). Similar patterns are observed as for the other outcomes: the rate of emergency admissions increases in age (IRR=1.017; p<0.001); there is a clear deprivation gradient; men are less likely to be admitted than women (IRR=0.935; p<0.001). This is a smaller difference between genders than for the count of (any) admissions. There are also smaller differences between ethnic groups for emergency admissions only (for example no significant difference between black and white individuals). Asian individuals have fewer emergency admissions than whites (IRR=0.910; p<0.001).

In terms of event time, there is a slightly different trend when comparing the count of emergency admissions with the count of admissions of any type. There is a larger increase between six years (IRR=0.376; p<0.001) and one year before the event (IRR=0.764; p<0.001); and there is a more pronounced increase in the year of the event (IRR=3.596; p<0.001).

The predicted number of admissions per annum six years prior to the event is 0.162 [CI 0.156; 0.168]; increasing to 0.213 two years [CI 0.208; 0.218] and 0.251 one year before [CI 0.246; 0.256] (Table 4.8). This spikes to 1.553 in the year of the event [CI 1.526; 1.581]; decreasing to 0.574 [CI 0.563; 0.585] and 0.491 [CI 0.481; 0.501] one and two years after. The number of emergency admissions is still 0.432 five years after the event [CI 0.418; 0.445], over three times its level six years prior to the event (equivalent to forty-five emergency admissions per hundred of the sample).

Table 4.8: Marginal effects (number of emergency admissions per year)

Event Time	ME	95% CI	
-6	0.162	0.156	0.168
-5	0.175	0.17	0.181
-4	0.19	0.185	0.195
-3	0.213	0.208	0.218
-2	0.251	0.246	0.256
-1	0.33	0.323	0.336
0	1.553	1.526	1.581
1	0.574	0.563	0.585
2	0.491	0.481	0.501
3	0.455	0.445	0.466
4	0.438	0.427	0.45
5	0.432	0.418	0.445

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1.

4.2.2.5 Drug related and non-drug related hospital admissions

The final set of (main) analyses were performed for admissions with a drug-related diagnosis only; and admissions without any such diagnosis (Table 4.4). These were estimated using count data models only for enabling a simple visual comparison for admissions stratified according to whether they are defined as drug related.

The same patterns are observed for the covariates (age, sex, deprivation, ethnicity) comparing the two 'types' of admission. The estimates on the event dummies in the pre-period are essentially zero in the pre-period for the drug related admissions – which is by design, as the event year is defined by the occurrence of an admission with a drug-related diagnosis. The comparison of the post period is interesting - across types. The incidence rate remains comparatively higher for drug-related admissions in the year following the event: 1.391 ($p<0.001$); compared with 1.248 for non-drug admissions ($p<0.001$). The incidence rates converge in the years subsequent to this (Table 4.4).

Table 4.9 shows the predicted values calculated from the regression analyses. Non-drug admissions rise steadily from 0.348 per year six years prior to the event [CI 0.326; 0.371], to 0.548 in the year prior to the event [CI 0.528; 0.568]. They decline to 0.633 in the year following the event [CI 0.610; 0.657]; and decline to 0.508 five years after the event [CI 0.479; 0.537] – remaining well above the level in the period prior to the drug-related admission. In the year after the event, drug-related admissions are

Table 4.9: Marginal effects (Drug vs non-drug admissions)

Event Time	Non-drug admissions			Drug-related admissions		
	ME	95% CI		ME	95% CI	
-6	0.348	0.326	0.371			
-5	0.364	0.344	0.384			
-4	0.383	0.365	0.401			
-3	0.411	0.393	0.428			
-2	0.457	0.44	0.475			
-1	0.548	0.528	0.568			
0	0.769	0.742	0.796	1.296	1.28	1.312
1	0.633	0.61	0.657	0.213	0.209	0.217
2	0.569	0.547	0.592	0.171	0.168	0.175
3	0.535	0.513	0.558	0.158	0.154	0.162
4	0.521	0.497	0.546	0.149	0.144	0.153
5	0.508	0.479	0.537	0.153	0.146	0.16

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1. Note that there are no pre-event estimates for drug-related admissions as they are zero by definition: any drug-related admission defines the event time index (i.e. the first admission defines index (event) time=0 for each individual)

0.213 [CI 0.209; 0.217]. This decreases to 0.149 four years after the event [CI 0.144; 0.153], prior to increasing to 0.153 [CI 0.146; 0.160] in the fifth year following the event.

4.2.2.6 Sensitivity analyses

Analyses were conducted to test whether the results were sensitive to the removal of individuals whose event occurred in the first or second year of the analysis period (2009-10 and 2010-11). These analyses, estimated on total costs, showed the same pattern in the event time estimates as the full sample – and removal of these individuals increases the magnitude of the estimated effect in the event year and in the post period (Appendix Tables A4.6 and A4.7; Appendix Figure A4f).

In the case of activity directly costed from the tariff, estimated costs are generally lower (Appendix Table 4.7) than for the full model (Table 4.5) (for example, predicted costs are £607.88 lower in the event year [17.2%]). The proportional differences in the predicted costs comparing these models are generally larger at higher values of event time (Appendix Table 4.8). Excluding individuals whose is in either 2009-10 or 2010-11 very slightly reduces predicted costs for values of event time less than -1;

and slightly increases predicted costs at values of event time great than -2 (Appendix Table 4.7).

4.2.2.7 Comparing across outcomes

Table 4.10 shows that the effect of the event is more pronounced for emergency admissions and length of stay than for total admissions and total costs. There are several relevant observations to make when considering this. First, 66.53% of total admissions are emergency admissions. Second, 36.09% of total admissions are drug related. Third, 43.57% of emergency admissions are drug related. Finally, 89.54% of drug related admissions are emergency admissions.

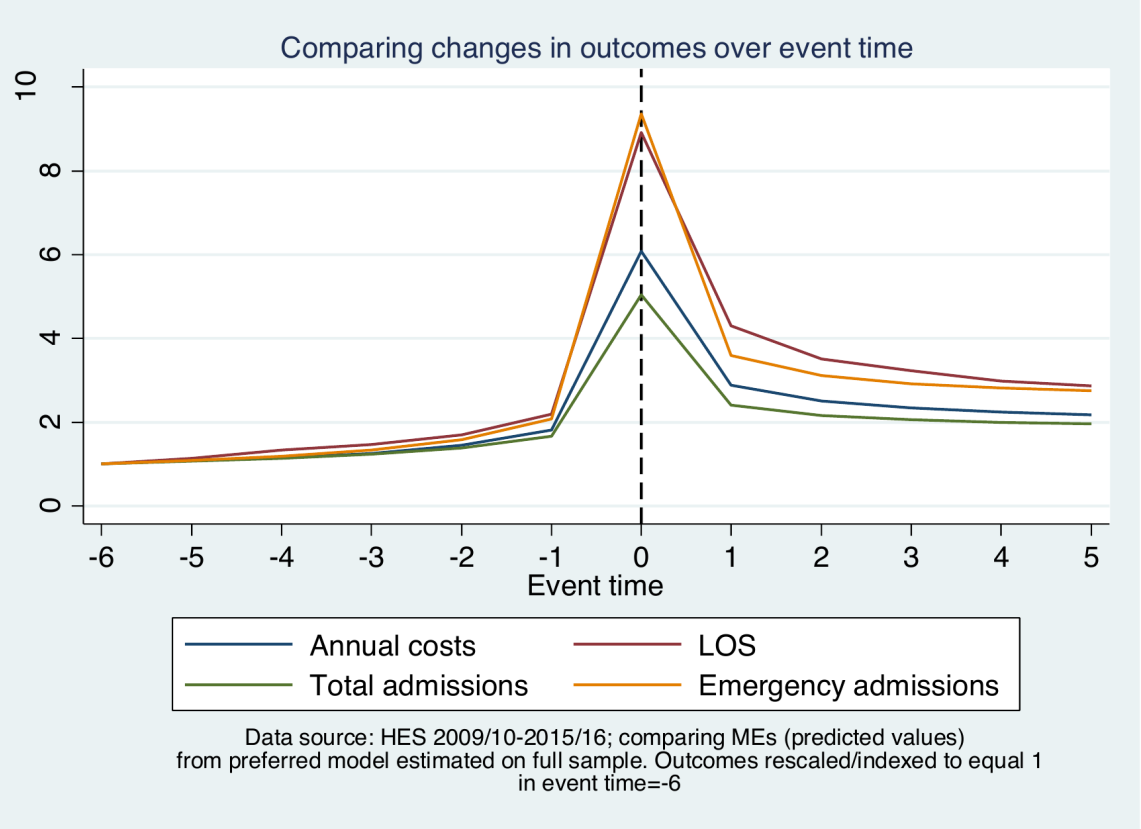
Figure 4a and Table 4.10 are constructed to compare the relative effects on the variables clearly – each is weighted to equal one in the first year of event time (i.e. when event time is equal to minus six). There is a differentially larger increase in the pre-period in emergency admissions and length of stay, than for total admissions and total costs. There is a more pronounced differential between the outcomes in the event year compared with six years prior (Table 4.10): emergency admissions are 836.2% higher; length of stay is 791.49% higher; whereas costs are 508.51% higher and total admissions are 404.07% higher. The comparison in the post-event period is remarkable: emergency admissions and the length of stay remain considerably higher than in the pre-event period – and this is a differentially higher level of persistence than for overall admissions.

Table 4.10: Comparing changes in predicted outcome by event time

Event Time	Cost		LOS		Admissions		Emergency admissions	
	ME	% change since -6	ME	% change since -6	ME	% change since -6	ME	% change since -6
-6	£580.84		1.198		0.369		0.163	
-5	£627.77	8.08%	1.369	14.27%	0.395	7.05%	0.178	9.20%
-4	£671.24	15.56%	1.603	33.81%	0.42	13.82%	0.194	19.02%
-3	£724.14	24.67%	1.753	46.33%	0.457	23.85%	0.219	34.36%
-2	£839.62	44.55%	2.036	69.95%	0.512	38.75%	0.258	58.28%
-1	£1,055.92	81.79%	2.634	119.87%	0.616	66.94%	0.339	107.98%
0	£3,534.49	508.51%	10.68	791.49%	1.86	404.07%	1.526	836.20%
1	£1,677.69	188.84%	5.146	329.55%	0.887	140.38%	0.586	259.51%
2	£1,455.31	150.55%	4.207	251.17%	0.798	116.26%	0.509	212.27%
3	£1,355.44	133.36%	3.876	223.54%	0.761	106.23%	0.476	192.02%
4	£1,301.83	124.13%	3.579	198.75%	0.739	100.27%	0.459	181.60%
5	£1,267.86	118.28%	3.443	187.40%	0.727	97.02%	0.45	176.07%

Compares full models (Tables 9, 11, 13 & 15) - total cost and LOS estimates from negative binomial; admissions and emergency admissions from Poisson regression with random effects.

Figure 4a: Visual comparison of changes in predicted outcomes over event time



4.3 Discussion

4.3.1 Summary of main findings

This study described annual patterns of hospital admissions and costs relative to a year in which an individual received a diagnosis related to substance misuse in hospital (the 'event'). It characterised the impact of this event on the following (annual) outcome variables per person: total costs, total LOS, and counts of hospital admissions classified as follows: all admissions; emergency admissions; drug related admissions and non-drug related admissions.

The results show a common pattern across the outcomes relative to event time: rising steadily in the pre-event period, then increasing dramatically in the year of the event, and decreasing after this - but to a level substantially higher than in the pre-event period.

The results display this same general pattern comparing across the study outcomes. These variables are closely related – a large proportion of the variation in the cost of

hospital admissions per year is explained by the number of admissions in the year, the admission method (elective vs. non-elective) for each admission, and the length of stay for these admissions. It is therefore logical that these variables exhibit similar relationships with characteristics such as age, sex; and with the event time indicators.

However, whilst the general patterns over event time are similar across outcomes, there are some differences of note. The effect of the event is more pronounced for emergency admissions and length of stay than for total admissions and total costs. There is also a differentially larger increase in the pre-period in emergency admission and length of stay, compared with total admissions and total costs.

It is also worth noting that there are differences in the impact of individual characteristics such as ethnicity and gender for admissions depending on whether this definition of admission was restricted to emergency only, or not. There is a qualitatively different relative rate of admissions per year comparing individuals of white and Asian ethnic origin: Asian individuals have 1.064 times as many admissions as whites but 0.915 times as many emergency admissions. The gender differential is much less pronounced for emergency admissions compared with any type of admission. There is also a more pronounced gradient in deprivation for emergency admissions only.

4.3.2 Explanation of notable findings

The size of the estimated effects in the event year are partially the consequence of the definition of the event. The event is defined as a year in which at least one drug related hospital admission occurred, and this is the primary cause of the size of the spike in the year of the event.

For example, in the case of admission counts, the increase in the event year over and above the definition of the event is given by subtracting one from the predicted value. Examination of non-drug related admissions allowed for examination of the size of impact of the event independent of how the event is defined.

The patterns on the event effects in the post-period show the persistence of differentially higher use of hospital services after the event. This is more pronounced for emergency admissions and length of stay in hospital. This overall increase likely reflects: the nature of substance misuse and addictive behaviours (which damage health in a persistent way over time); and the fact that drug related hospital admissions are much more likely to be emergency admissions. It also reflects the fact

that much of the activity classed as drug related by the definition used in this study does not relate to causes such as polypharmacy and drug interaction; the majority of these cases are for diagnoses that more likely indicate problematic addiction to substances. It is also likely that even 'one-off incidents' affect health in the event year and thereafter such that admissions are persistently higher; and differentially so in the case of emergency admissions.

The results show a pronounced effect on emergency admissions over event time. This is reflective of the very high probability that a drug-related admission is an emergency admission. This reflects the nature of admission to hospital for drugs-related causes. This study uses a definition for drug misuse related hospital admissions used by the ONS[6] for drug misuse death (ONS 2018). This therefore defines a drug misuse related hospital admissions as:

i. an admission where the underlying cause is drug abuse or drug dependence;

ii. an admission where the underlying cause is drug poisoning and where any of the substances under the Misuse of Drugs Act 1971 are involved.

Whilst a very high proportion of the drug related admissions are emergencies, this is not an indication that such admissions are overdoses. These admissions relate to the use of a range of substances, including legal and illegal drugs, prescription drugs (even where the prescription does not relate to the subject) and medicines available at retailers. The misuse of and dependence of medicines is a subject of increasing concern to policymakers (HMG 2017: 15). Use of medicines such as benzodiazepines and opioids can lead to dependence when used for extended periods of time. The admissions also include instances in which the activity is the result of complications of drugs misuse (such as heart disease relating to chronic cocaine use).

Finally, disproportionately higher use of emergency services in this group is consistent with other studies (see for example French et al. 2000). Substance misusers may exhibit an increased tendency to avoid preventative and ambulatory care for routine conditions. Neglected health care needs can then degenerate into more serious problems that subsequently require contact with emergency services and/or hospital admissions. Findings herein are consistent with this: showing increased hospital admissions once drugs misuse is recorded as a cause of admission – and differentially increased emergency admissions.

[6]this study also includes ICD-10 codes T40.1 - T40.9; and T43.6: Poisoning by, adverse effect of and underdosing of narcotic sand psychodysleptics [hallucinogens].

4.3.3 Comparison to existing research

This study is a comprehensive analysis of hospital admissions in a population with some diagnosis relating to substance misuse during the study period. The data are not subject to sampling variation as they are not drawn from a sample. Furthermore, random variation has very limited effect on the results as the sample size is in excess of three million observations; as demonstrated by the very small confidence intervals. This reflects the strength of using very large samples in statistical research (Li et al. 2012).

This study applies a relatively novel 'event study' approach to a cohort of drug misuse related individuals to reveal patterns of use of hospital services in this population, adjusting for other factors such as age, sex and calendar time.

It allows for reflection of more than simply annual changes in the number of admissions to hospital related to drugs misuse (ONS 2018; NHS Digital 2018). It considers changes in the number of days spent in hospital – showing for example whether admissions are for longer stays once a drug related admission occurs. It also considers costs, allowing for assessment of the severity and type of activity for each day spent in hospital, and how this affects use of economic resources. Costs are applied for each admission in the data, by using the payment data used for reimbursement of hospitals. Use of average unit costs is the approach typically used in the literature, which does not allow for refinement according to the factors which determine the cost of care such as admission method, length of stay, and health care resource group.

The application of this method also allows for a framework of comparison as to how a particular condition can have differential relative impacts in terms of the types of services affected. In this case, comparison was made between emergency admissions only; all admissions; and admissions with and without a diagnosis related to substance misuse. Other studies could consider an application which compared impacts across services such as primary and secondary care.

4.3.4 Strengths and Limitations

The general pattern of the pre-event dummies is notable – they show rising costs, (emergency) admissions and bed days. There are various reasons as to why this may be.

It may be indicative that an individual actually has drug use in this period which leads to an increased likelihood of being admitted to hospital, and increased likelihood that any admission would have greater severity or complexity than in the absence of drug(s) misuse – but that this drug use is not yet detected within hospital to the extent that it is entered as a cause of the admission in the diagnosis recording. It is also possible that such use is detected within hospital but is not considered as an underlying cause of the admission and as such is not recorded in the diagnosis codes. Another possibility is that the observed increase in the use of hospital services is reflecting deteriorating mental and/or physical health state(s) of individuals, which in turn increases their probability of drug(s) misuse.

Interpreting the coefficients on event time indicators after the hospital admission as the causal effect of the admission requires the identifying assumption that the timing of the admission is uncorrelated with the outcome (Dobkin et al. 2018). The pattern on the event dummies in the pre-period indicates that this assumption is unlikely to hold in this context. Using the timing of diagnosis in this study for identifying the attributable impact of a disease's (in this case substance misuse) impact on costs and activity is subject to limitations of endogeneity. First, it reflects a logical interpretation of the design insofar as the event is non-randomly allocated. Second, it reflects a similar limitation previous outlined for studies which determine attribution based on matching cohorts with and without a particular condition – specifically that the condition could be correlated with omitted variables.

Recording of substance misuse per se as an underlying medical cause in itself could not plausibly change future use of services. However, the pre-trends may indicate the weakness of assuming that a diagnosis of the condition in hospital in fact reflects the true health state of the individual – the individual could in this case have misused substances in the pre-event period but even in cases where they were admitted to hospital, it was not recorded in diagnosis codes. Despite this, the study nonetheless allows for the illustration of confounded patterns of use of hospital services in a specific population and to show how this pattern changes relative to a diagnosis being made.

There is no international consensus on what defines drug-misuse related hospital activity of death (ONS 2018). The definition used in this study allows for the inclusion of accidents, suicides and assaults relating to drug poisoning in addition to substance misuse and drug dependence. It does not, however, include particular consequences that may relate to drugs: for example transport accidents in which the driver was under the influence of drugs, or anaphylactic shock.

I am unable to provide breakdowns by specific substance in this study, due to the absence of specific codes for many common recreational drugs (Shah et al. 2011). This partly reflects the difficulty in representing the range of novel recreational drugs that exist and are introduced over time. Furthermore, many problem drug users take more than one substance with regularity (Jones at al. 2009; Davies et al. 2009; Jones et al. 2017).

ONS (2018) mortality records in some cases do not specify the year of death. This may cause the data to have observations that should be removed as a result of attrition. In this case, there are additional zero values in the dataset that are not valid which may cause the parameter estimates on the event year effects to be downwards biased.

These data depend on the information provided by doctors on diagnoses that are then recorded in hospital administrative records. There is non-trivial variation in the level of detail provided by doctors – some will systematically record underlying causes in greater detail than others (Shah et al. 2011). Some caution is necessary given that diagnosis recording practices have improved from 2006-7 onwards (NHS Digital 2018), and given that drug misuse may only be suspected but not recorded by the hospital. Diagnosis recording of substance misuse is generally under-recorded (Shah et al. 2011), leading to underestimation of the amount of activity that relates to substance misuse.

The data in this study considers inpatients only and therefore does not reflect all hospital activity. I provide no detail on use of outpatient services, although use of these services has been shown to be lower for drug users than for non-drug users (French et al. 2000). Furthermore, this study does not provide information on patterns in primary care and other sectors of health care provision. This was beyond the scope of this study, and the results are not considered generalizable across different sectors. A significant proportion of patients presenting with substance misuse related conditions are not admitted to hospital – and this study fails to capture the burden of provision relating to attendances to accident and emergency (A&E).

This study considers the period 2009-10 to 2015-16 and assumes that the event is the 'first' drug related admission. However, relapse is a known issue amongst individuals with addictive behaviours (see for example Shaffer et al. 2009; Jones et al. 2017) and refers to a recurrence in the addictive behaviour in an individual who was previously abstinent. Models were tested to consider the sensitivity of the results to exclusion of individuals whose event was in the first two years. This ensures that the results

do not change where it is known that the event did not occur in the two years before the sample period. These findings more generally are robust to various model specifications used to deal with the count properties of the outcome data, and they are robust to the inclusion of individual-specific effects. They are also found to be consistent with two sets of sensitivity analyses: the first excluded individuals whose event occurred at time $T < 3$; the second excluded observations that were not costed directly from the tariff.

The approach taken to costing data that did not have a matching value in the NTPS led to a higher estimate of the cost of admissions. This likely reflects the skewness of the cost data used to generate these values. Use of other measures of resource use (admissions, LOS) recognises this possibility, and exclusion of this activity was included to show the impact of this approach.

4.3.5 Implications for research and policy

Future studies may seek to consider further exploration of the methods applied in this chapter. Specifically, further exploration could combine an event study approach with an alternative in using matched cohorts with and without the condition. Other applications could consider health events that are more likely to be robust to an assumption of exogeneity (such as a road traffic accident).

This study shows the pattern of use of services around the year in which individuals are admitted to hospital where substance misuse is considered to be an underlying cause of the admissions. The use of hospital services is actually increasing prior to this event – and a differentially larger proportion of admissions are emergency admissions in this pre-period. Use of services remains persistently higher after the event, with a higher proportion of admissions being emergencies. In the case of individuals with substance misuse problems, the government is committed to (HMG 2017: 2):

> "offering people with a drug dependence problem the best chance of recovery support at every stage of their life."

Whilst increased use of elective hospital services may be the partial consequence of contact with addiction treatment services, the illustrated increased levels of emergency admissions are unlikely to reflect an intended increase in the use of hospital services. Emergency admissions are more likely to reflect either problematic use in

the population that are not in rehabilitation treatment, or relapse in the population that are in rehabilitation. As such, the patterns of use of emergency services after the event reflect either failures in recruitment and retention in rehabilitation.

This may reflect issues in access to and acceptability of rehabilitation services, which have partly shifted their strategic focus from harm reduction to abstinence (Mason et al. 2015b). It may also be indicative of failures in directing individuals towards the relevant services. As such, diagnosis coding of substance misuse may offer a useful indicator for policymakers. The potential for these codes as an indicator of addiction status has been previously highlighted by Wiley et al. (2013); and use of secondary as well as primary diagnosis codes recommended by Shah et al. (2011).

The patterns shown in this study are evidence that records of substance misuse in hospital diagnosis codes are potentially indicative of a persistent problem of substance misuse that increases use of hospital services for a population that are known to have higher use of health services than the general population (Stein et al. 1993; Haapanen-Niemi et al 1999; French et al. 2011).

At present, referral to addiction treatment services is more common through the criminal justice system and primary care than secondary care (Jones et al. 2017). There is potential for exploration of how these diagnosis codes can be better used to direct individuals towards the treatment services they require and prevent the social, health and economic costs that are caused by substance misuse. Such possibilities might mirror approaches taken for other health behaviours – such as provision of smoking cessation advice

5 The impact of paying care providers for outcomes: differences-in-differences analysis of the payment by results for drugs recovery programme

5.1 Preface

This chapter was published as a peer-reviewed research report in the journal 'Addiction' on 7th June 2015, first-authored by the author of this PhD thesis and co-authored by several collaborators. The analysis and interpretation of the data and the design of the study in this research report was undertaken by Thomas Mason, Andrew Jones, Matthew Sutton and William Whittaker. The paper was drafted in full by Thomas Mason. All other authors contributed to and approved the final version of the paper. The paper is available online at: https://onlinelibrary.wiley.com/doi/pdf/10.1111/add.12920

5.2 Introduction

The adoption of pay-for-performance (P4P) schemes is becoming more widespread internationally in health care (Van Herck et al. 2010) and across the public sector in the UK (HMG 2010). Most of these schemes reward providers exclusively, or in large part, on the basis of their performance on 'best practice' process measures rather than outcomes (Ryan & Doran 2010). There is a growing evidence base that seeks to identify the effects of P4P schemes in health care (Van Herck et al. 2010), but there is limited evidence that the introduction of these schemes is associated with better outcomes (Flodgren et al. 2011) and a substantial gap remains in the evidence base regarding the effects of these programs (Mehrotra et al. 2009).

In many previous P4P schemes, payments have typically been linked to process measures of clinical quality rather than outcomes. This has, in some instances such as the National Health Service (NHS) Quality and Outcomes Framework (QOF),

been because it is difficult to link remuneration to 'hard' outcomes such as mortality (which is rare) (Kontopantelis et al. 2014). Studies have therefore typically focused on schemes' effects on clinical processes rather than outcomes (Van Herck et al. 2010), albeit with some notable exceptions (Sutton et al. 2012). Reviews of the P4P literature have also noted that many studies have either been unable to adopt a quasi-experimental design or failed to do so even where this was possible (Flodgren et al. 2011 & Scott et al. 2011).

The P4P schemes introduced by the UK Government from around 2010 are focused more clearly on paying for 'outcomes', for example finding employment for people claiming unemployment benefit and preventing re-offending by released prisoners (HMG 2010). This focus on outcomes is more consistent with simultaneous developments in English health policy (HMG 2012).

Commissioning of drug treatment services has focused on retention in treatment as the principal measure of effectiveness for many years and service providers have been paid using 'block' or 'activity' contracts. In April 2012 the Department of Health in England introduced a pilot programme under which providers of drug and alcohol misuse services received payments based on "recovery-focused" outcomes (HMG 2010a). This pilot in eight geographical areas was introduced in relation to a central pillar of a new Drug Strategy (HMG 2010a), which emphasised the desirability of service users recovering from their dependency on drugs or alcohol and successfully completing treatment free from drugs of dependence, without continued prescribing of substitute drugs.

Outcome-based payments are promoted as paying providers for 'the value they produce' and are intended to incentivise innovation rather than increased adoption of particular processes by providers. However, P4P schemes can also generate unintended effects (Newhouse 1984; Newhouse 1994; Smith 1995; Ellis 1998; Doran et al. 2011). In particular, outcome-based payments transfer more risk from payers to providers.

In this study, we compare successful completion of treatment by service users between pilot and non-pilot areas over time using a national administrative dataset.

5.3 Intervention

Prior to the introduction of P4P in pilot areas, drug treatment was partly funded through allocations from the public health budget to Primary Care Trusts (PCTs) from central government. The funds were allocated by a formula which recognised key deprivation factors and the size of areas' treatment populations. The remainder of funding for drug treatment was distributed from the Department of Health (DH) to DATs who then awarded funding to local providers. Typically, prior to 2012-13 providers' budgets were calculated simply as the product of the predicted size of the treatment population and the average price per service user (Maynard et al. 2011), although local flexibility for funding existed.

Under the new P4P scheme, DATs link providers' income to performance indicators. Commissioners were given limited scope for local design, provided their payment scheme adhered to a 'national outcomes framework'. Local commissioners have some limited flexibility regarding the payment weights attached to indicators in the national framework (for example, successful completion of treatment and non-representation outcome always comprise a large share of P4P income). Providers in the pilot areas are paid for outcomes in three nationally set domains: (i) progression towards abstinence from presenting drug(s) of dependence; (ii) reduction in offending; and (iii) improved health and wellbeing. We focus on the first of these, as it is the primary objective of the Government's Drug Strategy (HMG 2010a); and within this domain, we focus specifically on successful completion as the key proxy measure of recovery used by Public Health England (PHE 2014) (Appendix Tables A5.1 and A5.2).

5.4 Results

Small, statistically significant differences were identified in client complexity between the intervention and comparator areas prior the introduction of P4P ($p < 0.001$) (Table 5.1).

Successful completions fell by 0.29 percentage points comparing 2011-12 and 2012-13 in non-pilot areas, compared with 2.46 percentage points in pilot areas for the same period (Table 5.2). Declining to continue with treatment decreased by 0.12 percentage points comparing 2011-12 and 2012-13 in non-pilot areas, compared with an increase of 0.56 percentage points for pilot areas over the same period (Table 5.2).

Table 5.1: Historical DAT Characteristics pre-P4P by pilot status

	Non-pilot				Pilot			
	2010/11		2011/12		2010/11		2011/12	
	Total	%	Total	%	Total	%	Total	%
Females	46,098	27.00%	44,974	26.90%	3,168	26.40%	3,013	26.20%
Age 18-29	60,150	35.30%	56,165	33.60%	4,781	39.80%	4,324	37.60%
Age 30-49	103,122	60.50%	103,590	61.90%	6,822	56.80%	6,742	58.70%
Age 50-69	7,106	4.20%	7,555	4.50%	415	3.50%	420	3.70%
Age 70+	50	0.00%	62	0.00%	1	0.00%	2	0.00%
Opiate Users	140,953	82.70%	136,814	81.70%	9,496	79.00%	9,273	80.70%
Crack Users	51,785	30.40%	50,117	29.90%	2,890	24.00%	2,763	24.10%
Benzodiazepine Use	123	0.10%	100	0.10%	13	0.10%	9	0.10%
Injected Recently	46,153	27.60%	43,645	26.60%	3,501	30.50%	3,322	30.00%
No Fixed Abode	41,206	25.30%	39,743	24.70%	2,721	23.50%	2,618	23.50%
Referred by CJS	18,333	10.80%	18,420	11.00%	1,047	8.70%	1,039	9.00%
Drug using career								
Up to 5 yrs	33,785	21.70%	32,591	21.10%	2,646	24.00%	2,427	23.00%
6 to 10 yrs	36,582	23.40%	34,661	22.50%	2,681	24.40%	2,438	23.10%
11 to 20 yrs	61,262	39.30%	61,598	39.90%	4,346	39.50%	4,285	40.70%
21 to 30 yrs	20,134	12.90%	20,928	13.60%	1,129	10.30%	1,154	10.90%
Over 30 yrs	4,263	2.70%	4,606	3.00%	205	1.90%	235	2.20%

Table 5.2: Treatment Completion and Declining for pilot and non-pilot areas pre and post-P4P

	Non-pilot areas		Pilot areas	
	2011-12	2012-13	2011-12	2012-13
Successful completions	12.50%	12.21%	13.73%	11.27%
Interquartile Range	9.7%; 14.1%	9.4%; 14.1%	11.4%; 20.9%	6.7%; 16.1%
Treatment declinations	0.55%	0.43%	0.47%	1.03%
Interquartile Range	0.21%; 0.68%	0.15%; 0.63%	0.11%; 0.43%	0.15%; 0.97%
Remained in treatment	72.93%	73.66%	72.13%	75.18%
N	170,428	167,372	12,019	11,488

Table 5.3: Logistic regression results for completion

Completion	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
Effect of Pilot Scheme	0.859	0.038	-3.42	0.001	[0.788	0.937]
Financial Year 2012-13	0.947	0.011	-4.75	<0.001	[0.927	0.969]
Pilot Area	0.939	0.071	-0.84	0.4	[0.810	1.088]
Baseline Level of Completions	46.039	17.384	10.14	<0.001	[21.964	96.501]
Female	1.056	0.013	4.28	<0.001	[1.030	1.082]
Age	1.003	0.001	3.89	<0.001	[1.002	1.005]
Opiate User	0.167	0.002	-143.74	<0.001	[0.163	0.171]
Crack User	0.857	0.012	-11.02	<0.001	[0.834	0.881]
Injected Recently	0.621	0.01	-28.72	<0.001	[0.601	0.642]
No Fixed Abode	0.912	0.012	-6.73	<0.001	[0.888	0.937]
Years of Drug Use	0.985	0.001	-17.62	<0.001	[0.983	0.986]
Benzodiazepine User	0.574	0.135	-2.37	0.018	[0.362	0.909]
Referred by CJS	0.926	0.017	-4.33	<0.001	[0.894	0.959]
Constant	0.402	0.021	-17.11	<0.001	[0.363	0.447]
σ_u	0.184	0.013			[0.161	0.211]
ρ	0.01	0.001			[0.008	0.013]
Number of individuals	346,300					
Number of DATs	149					

5.4.1 Completions

We found that service users located in pilot areas were significantly less likely to complete treatment. Service users treated in pilot DATs were 1.3 percentage points (OR: 0.859; 95% CI [0.788; 0.937]) less likely to complete treatment compared to those treated in comparison DATs (Table 5.3 and Appendix Table A5.3). We also found significant within-DAT correlation (ρ=0.010; $\bar{\chi}^2$ = 713.73; p < 0.001) (Table 5.3).

This result was robust across both secondary analyses. For secondary analysis I, we found that service users located in pilot DATs were 1.0 percentage point less likely to complete treatment as those located in comparison DATs (OR: 0.895; 95% CI [0.812; 0.988]) (Appendix Table A5.3). For secondary analysis II, we found that service users located in pilot DATs were 1.6 percentage points less likely to complete treatment as those located in comparison DATs (OR: 0.834; 95% CI [0.764; 0.912]) (Appendix Table 5.3).

Table 5.4: Logistic regression results for declined to remain in treatment

Declined to Remain in Treatment	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
Effect of Pilot Scheme	2.934	0.505	6.25	<0.001	[2.094	4.113]
Financial Year 2012-13	0.781	0.04	-4.85	<0.001	[0.707	0.863]
Pilot Area	0.786	0.247	-0.77	0.444	[0.425	1.455]
Baseline Level of Declining	$2.36\text{x}10^{31}$	$2.64\text{x}10^{32}$	6.47	<0.001	[$7.46\text{x}10^{21}$	$7.50\text{x}10^{40}$]
Female	1.074	0.06	1.29	0.197	[0.964	1.197]
Age	0.997	0.003	-0.9	0.369	[0.990	1.004]
Opiate User	0.227	0.012	-27.14	<0.001	[0.204	0.253]
Crack User	1.005	0.062	0.07	0.941	[0.890	1.133]
Injected Recently	0.801	0.057	-3.15	0.002	[0.697	0.919]
No Fixed Abode	1.203	0.068	3.26	0.001	[1.077	1.344]
Years of Drug Use	0.986	0.004	-3.59	<0.001	[0.978	0.994]
Benzodiazepine User	1.389	0.996	0.46	0.646	[0.341	5.664]
Referred by CJS	0.754	0.063	-3.35	0.001	[0.640	0.890]
Constant	0.008	0.001	-30.81	<0.001	[0.006	0.011]
σ_u	0.732	0.06			[0.623	0.860]
ρ	0.014	0.02			[0.106	0.183]
Number of individuals	346,300					
Number of DATs	149					

5.4.2 Declining to continue with treatment

We found that service users located in pilot areas were significantly more likely to decline to continue with treatment. Service users treated in pilot DATs were 0.6 percentage points (OR: 2.934; 95% CI [2.094; 4.113]) more likely to decline to continue with treatment compared to those treated in comparison DATs (Table 5.4 and Appendix Table A5.4). We also found significant within-DAT correlation (ρ=0.014; $\bar{\chi}^2 = 578.36$; $p < 0.001$) (Table 5.4).

These results were robust across both secondary analyses. For secondary analysis I, we found that service users located in pilot DATs were 0.6 percentage points more likely to decline to continue with treatment as those located in comparison DATs (OR: 3.014; 95% CI [2.059; 4.411]) (Appendix Table 5.4). For secondary analysis II, we found that service users located in pilot DATs were 0.5 percentage points more likely decline to continue with treatment as those located in comparison DATs (OR: 2.827; 95% CI [1.999; 3.998]) (Appendix Table A5.4).

For our alternative definition of the intervention and comparison groups, the results for both outcome measures were consistent. Service users treated in P4P DATs were 1.0 percentage points (OR: 0.896; 95% CI [0.854; 0.939]) less likely to complete treatment compared to those treated in non-P4P DATs. Service users treated in P4P DATs were 0.14 percentage points (OR: 0.896; 95% CI [0.854; 0.939]) more likely to decline to continue with treatment compared to those treated in non-P4P DATs.

5.5 Discussion

5.5.1 Summary of Main Findings

We found that, after the introduction of P4P, service users located in pilot areas were significantly less likely to complete treatment and were significantly more likely to decline to continue with treatment than those located in non-pilot areas. The main results were robust to a range of secondary analyses. In the case of completion rates, we performed analyses which restricted the sample of comparison DATs to those that were similar to the pilot DATs. In the first case in which the comparison areas were similar both in terms of the percentage of the treatment population using opiates or crack and their 2010 IMD index of multiple deprivation score, we found that the negative treatment effect was slightly diluted but still strongly significant. When we

then included only comparison DATs which were similar based on geographic location, the negative treatment effect was larger in magnitude and strongly significant. For declining to continue with treatment, we found very similar effect sizes for both sets of secondary comparisons, and these effects remained statistically significant. These results were also robust to alternative definitions of the intervention and comparison groups.

5.5.2 Explanation of Notable Findings

P4P changed the way providers of drug recovery treatment were paid, putting a proportion of providers' incomes at risk, compared to activity-based financing which pays providers on the basis of activity indiscriminately of quality or outcomes. Whilst the P4P scheme allows for an element of local variation, non-representation for treatment (Tables A5.1 and A5.2) comprises a substantial share of P4P income in the eight pilot areas. Successful completion of treatment also comprises a substantial share of P4P income. However, to trigger payment for a successful completion, a service user must have completed a planned treatment and be considered as free of dependency from their presenting substance in addition to crack and opiates. Introduction of the P4P scheme and the structure of the payments may have led providers to be more risk-averse in discharging service users from treatment and thus recording them as successfully completed, compared with previous financing arrangements. The non-representation payment comprises the largest share of providers' P4P income, and this may have led to providers retaining more service users in treatment as a result of becoming less willing to risk service users re-presenting for treatment.

An increased likelihood of service users declining to continue with treatment in pilot areas is a probable unintended consequence of the P4P scheme. The introduction of P4P, linking payment to recovery-based outcomes, is likely to have changed the nature of treatment – resulting in some service users declining to continue with treatment. We believe this is an unintended consequence as studies have shown that engagement with treatment services increases the probability of recovering from drug addiction and reduces the probability of service users' likelihood of future hospitalizations and emergency department usage (Neighbors et al 2005).

5.5.3 Comparison with Existing Research

This study is the first statistical evaluation of this P4P scheme. The scheme is unique insofar as payments were linked to treatment outcomes and not process measures of care quality. The majority of previous research has evaluated the impact of P4P schemes on quality measures, typically because the schemes have linked payments to these measures. There have been limited previous attempts to introduce PbR funding for providers of drug misuse treatments (Commons et al. 1997; Shen & Ellis 2002; McLellan et al 2008) and the evidence base is equivocal as to its effects in this context.

The wider literature on the use of P4P in health care is also ambiguous as regards its impacts. Several previous studies have found improvements in selected quality measures and have suggested that P4P can potentially be effective, but convincing evidence is still lacking (Van Herck et al. 2010; Flodgren et al. 2011; Eijkenaar et al. 2013). Furthermore, many previous studies have been unable to adopt a tight, controlled, quasi-experimental design (Van Herck et al. 2010; Ryan & Doran 2010; Scott et al. 2011; Maynard et al. 2011; Eijkenaar et al. 2013). In some cases (such as the QOF), this has been because the policy implementation has not aided itself to such an evaluation. We use a controlled, quasi-experimental design and control for unmeasured confounding. We also check whether our findings are robust for a set of secondary comparisons based on areas similar to the intervention areas on three different bases (treatment population complexity, area level deprivation and location).

5.5.4 Study Limitations

Non-pilot areas were subsequently permitted to implement their own, fully locally-designed P4P scheme after the introduction of P4P in April 2012. We obtained information regarding adopters and non-adopters, but there is no available information regarding the extent to which P4P schemes in these areas correspond to the pilot scheme.

At the introduction of P4P, not every client located in providers in pilot areas was technically a 'P4P client'. This remained the case for a short period whilst complexity reassessment was undertaken for the purposes of re-tariffing.

Our results are for the first year following the launch of the PbR pilot sites and

rates of treatment completion in these areas may improve in the longer-term. A considerable amount of system change was required to implement PbR. In pilot areas, recommissioning took place to coincide with the introduction of P4P and this led to new providers taking over from old ones in many cases. In some cases, this equated to a considerable reconfiguration of services (in one area, a single new provider replaced five providers). Introduction of the P4P pilot also required considerable changes to areas' existing performance monitoring and measurement systems and processes.

This study is not a cluster-randomised trial, although we have sought to control for differences in treatment populations and unmeasured confounding.

5.5.5 Implications for Policy and Future Research

An important lesson from this study, consistent with findings from the wider P4P evidence-base, is that improving performance via P4P is not straightforward. Whilst many previous studies have shown some limited improvements in quality under P4P, we find that the introduction of payment for outcomes had a significant, negative impact on successful treatment completion.

We focus on the payment domain which comprises the highest proportion of provider income and that is the primary objective of government policy. Future studies should consider the effects of the P4P scheme on other payment domains (offending, accommodation, self-reported physical health and wellbeing). In addition to this, future studies should consider the impact on measures of quality to which payments have not been explicitly linked aside from declining to continue with treatment (such as waiting times and additional measures treatment engagement).

6 Did paying addiction service providers for outcomes lead to unintended consequences? Difference-in-differences analysis of hospital admissions, costs and length of stay in England 2009-10 to 2015-16.

6.1 Introduction

Previous discussions in this thesis have described how pay-for-performance schemes have become more widespread in health care and across public services more generally as governments seek to switch towards more active forms of reimbursement for providers of public services. Whilst the evidence base evaluating the effects of such schemes has grown considerably over the past 15 years, there are no simple conclusions as to their impacts (Scott et al. 2011; Van Herck et al. 2011; Alhamsan et al. 2010; Eijkenaar et al. 2013; Mendelson et al. 2017; Scott et al. 2016). Overall, studies have suggested that their effects depend on the scheme's design, the context and the implementation. In health care, studies evaluating these schemes generally consider their impact on indicators of providers' activities: typically, whether a provider performed a particular task which is considered to be clearly beneficial in achieving the objective for the service in question. Schemes have tended to link payment to these processes, because linking to outcomes can be more problematic (particularly for health outcomes which can be rare).

The policy change introduced by the UK government in 2012 for the providers of substance misuse service has been set out in detail in this thesis. It linked payment to treatment outcomes for providers in eight geographical areas in England. Studies, including the previous chapter, have shown that it led to reduced treatment completion and increased rates of disengagement with treatment services in the eight pilot areas. One possible explanation for this finding is that providers' behavior changed in response to an increase in the proportion of their income being uncertain – possibly changing the nature of treatment and potentially changing the inclination of providers to discharge individuals on the basis that re-entry in substance misuse services would lead to a loss of income for providers. Jones et al. (2017) showed that the

policy led to a range of consequences which had not been set out as objectives of its introduction by policymakers, including differentially reduced treatment initiation; treatment completion; treatment completion without subsequent re-presentation; and higher waiting times.

Whilst other studies and the previous chapter considered impacts of the policy for a population in structured treatment for their substance misuse problem, it did not consider impacts across the wider drug misusing population. Restricting evaluation of the policy's impacts to those in-treatment only potentially overlooks assessment of impacts on access to, engagement with and retention in substance misuse treatment services. The following chapter of this thesis therefore focuses on considering whether introduction of the P4P scheme in question in April 2012 led to unintended consequences in a wider substance misuse-related population using a quasi-experimental design applied to patient-level data, including individual characteristics, and an assessment of underlying trends. Specifically, it addresses the following research questions. For individuals admitted to hospital with a drug-related diagnosis at least once between 2009-10 and 2015-16 (the study population), did the introduction of P4P for providers of substance misuse services in the area in which they were initially resident:

i. lead to a differential change in the probability of hospital admission?

ii. lead to a differential change in the probability of emergency hospital admission?

iii. lead to a differential change in the number of days spent in hospital?

iv. lead to a differential change in the cost of their hospital admissions?

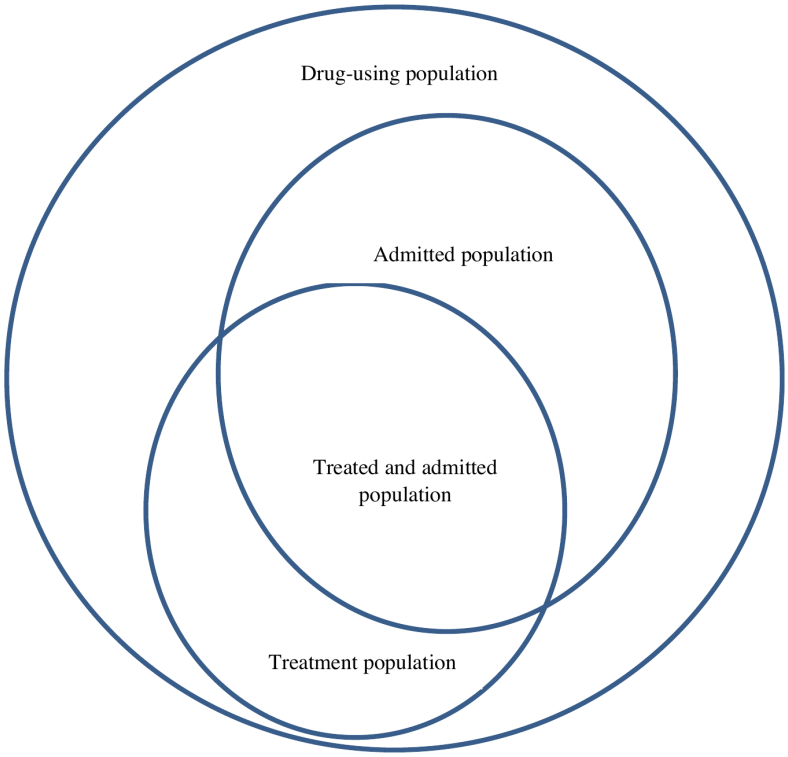
v. lead to a differential change in the probability of drug-related hospital admissions?

vi. lead to a differential change in the probability of hospital admissions not diagnosed as being drug-related?

vii. do any differential changes in the outcome above vary across the four years after the introduction of P4P in April 2012?

The consideration of all persons with a drug-related hospital admission allows for examination of the health care activity of both those in substance misuse treatment

Figure 6a: Study populations



(considered in Chapter 4), and those not in treatment for substance misuse (see Figure 6a).

Addressing the above research questions will provide evidence on whether the P4P scheme in question led to unintended consequences in terms of hospital admissions in a substance misuse related population. It also addresses the question of unintended consequences in a setting that is relatively unique in the literature:

i. in the substance misuse setting (rather than primary or secondary care), in which outcomes are jointly determined by providers and individuals, and treatment is a complex and lengthy process that frequently uses a range of approaches (Day 2018);

ii. for a policy that linked payment to outcomes and not processes;

iii. in a setting in which provision of the service is characterized by a diverse market of providers from the public, private and voluntary sectors;

iv. in a setting in a P4P scheme has been introduced for a condition (substance misuse) that is unlike the context for a majority of previous schemes (such as

stroke and hip replacement); in which providers observe individuals over a more prolonged period for a condition that is unobservable, and diagnosis requires substantive personal interaction between subjects and practitioners.

The remainder of this chapter is set out as follows: I set out the results of the analyses; and then discuss the results and their implications in detail.

6.2 Results

6.2.1 Summary statistics and trends

Summary statistics for the six outcome variables are shown in Table 6.1. Each of the outcomes exhibit marked skewness, with the standard deviation at least three times larger than the mean for each of the outcomes.

Tables 6.2 to 6.4 illustrate the changes over time in the characteristics of the study population in the intervention and comparison areas in the pre-intervention period (2009-10 to 2011-12). The intervention areas have less ethnic diversity than the comparison areas, and generally have a higher percentage of individuals resident in areas of lower deprivation (Table 6.2). The populations are generally similar in terms of age and gender in both the intervention and comparison areas (Table 6.3). Table 6.4 compares the number of individuals observed in the intervention and comparison areas in each year. A similar proportion of individuals exit the sample as a result of death in both the intervention and comparison areas.

Table 6.1: Summary statistics (dependent variables)

Variable (per annum)	Mean	Standard Deviation	Min	Max
Number of admissions	0.980	3.497	0	316
Number of emergency admissions	0.608	1.514	0	161
Number of non drug-related admissions	0.722	3.353	0	316
Number of drug-related admissions	0.258	0.757	0	159
Cost of admissions (£)	1,573.648	5,619.960	0	798,425
Total LOS	4.793	21.926	0	365

N=3,893,503 observations on 572,585 individuals

Table 6.2: Individuals' characteristics depending on initial residence prior to the introduction of P4P (Ethnicity & IMD)

	Intervention areas						Comparison areas					
	2009-10		2010-11		2011-12		2009-10		2010-11		2011-12	
	n	%	n	%	n	%	n	%	n	%	n	%
Ethnicity												
Asian	75	1.28%	95	1.66%	90	1.65%	2,730	3.31%	2,654	3.31%	2,635	3.38%
Black	92	1.58%	75	1.31%	72	1.32%	2,329	2.83%	2,253	2.81%	2,296	2.95%
Mixed	62	1.06%	67	1.17%	64	1.17%	989	1.20%	1,049	1.31%	1,024	1.31%
Other	128	2.19%	117	2.05%	117	2.15%	2,295	2.79%	2,263	2.83%	2,255	2.90%
White	5,482	93.89%	5,360	93.80%	5,104	93.70%	74,016	89.87%	71,859	89.74%	69,665	89.46%
IMD Decile												
1	719	12.31%	739	12.93%	742	13.62%	8,179	9.93%	7,954	9.93%	7,464	9.58%
2	748	12.81%	736	12.88%	691	12.69%	8,295	10.07%	7,971	9.95%	7,526	9.66%
3	706	12.09%	690	12.08%	720	13.22%	7,996	9.71%	7,861	9.82%	7,663	9.84%
4	663	11.35%	606	10.61%	594	10.91%	8,114	9.85%	7,930	9.90%	7,670	9.85%
5	565	9.68%	571	9.99%	484	8.89%	8,133	9.88%	8,105	10.12%	7,837	10.06%
6	558	9.56%	589	10.31%	541	9.93%	8,264	10.03%	8,047	10.05%	7,806	10.02%
7	614	10.52%	564	9.87%	553	10.15%	8,157	9.90%	7,799	9.74%	8,061	10.35%
8	563	9.64%	510	8.93%	458	8.41%	8,203	9.96%	7,995	9.98%	7,836	10.06%
9	372	6.37%	347	6.07%	350	6.43%	8,417	10.22%	8,168	10.20%	8,069	10.36%
10	331	5.67%	362	6.34%	314	5.76%	8,601	10.44%	8,248	10.30%	7,943	10.20%

Table 6.3: Individuals' characteristics depending on initial residence prior to the introduction of P4P (Age & Gender)

	Intervention areas						Comparison areas					
	2009-10		2010-11		2011-12		2009-10		2010-11		2011-12	
	N	%	N	%	N	%	N	%	N	%	N	%
Age (years)												
16-25	1,679	28.75%	1,477	25.85%	1,355	24.88%	22,530	27.36%	20,370	25.44%	18,510	23.77%
26-35	1,195	20.47%	1,209	21.16%	1,171	21.50%	18,304	22.22%	18,280	22.83%	18,328	23.54%
36-45	1,225	20.98%	1,301	22.77%	1,189	21.83%	17,884	21.71%	17,679	22.08%	17,438	22.40%
46-55	710	12.16%	779	13.63%	747	13.71%	10,159	12.34%	10,678	13.33%	11,268	14.47%
56-65	385	6.59%	351	6.14%	386	7.09%	5,077	6.16%	5,092	6.36%	5,088	6.53%
66-75	261	4.47%	252	4.41%	255	4.68%	3,496	4.24%	3,365	4.20%	3,000	3.85%
76 and over	384	6.58%	345	6.04%	344	6.32%	4,909	5.96%	4,614	5.76%	4,227	5.43%
Male	2,981	51.05%	2,883	50.46%	2,766	50.78%	42,805	51.97%	41,537	51.87%	40,341	51.81%

Table 6.4: Number (and %) of individuals by initial residence, by year

	Intervention areas		Comparison areas		Total	
	N	%	N	%	N	%
Year						
2009-10	37,784	14.71%	534,801	14.71%	572,585	14.71%
2010-11	37,625	14.65%	532,453	14.64%	570,078	14.64%
2011-12	37,312	14.53%	528,210	14.52%	565,522	14.52%
2012-13	36,898	14.37%	522,481	14.37%	559,379	14.37%
2013-14	36,359	14.16%	515,208	14.17%	551,567	14.17%
2014-15	35,767	13.93%	506,881	13.94%	542,648	13.94%
2015-16	35,047	13.65%	496,677	13.66%	531,724	13.66%
Total	256,792	100%	3,636,711	100%	3,893,503	100%

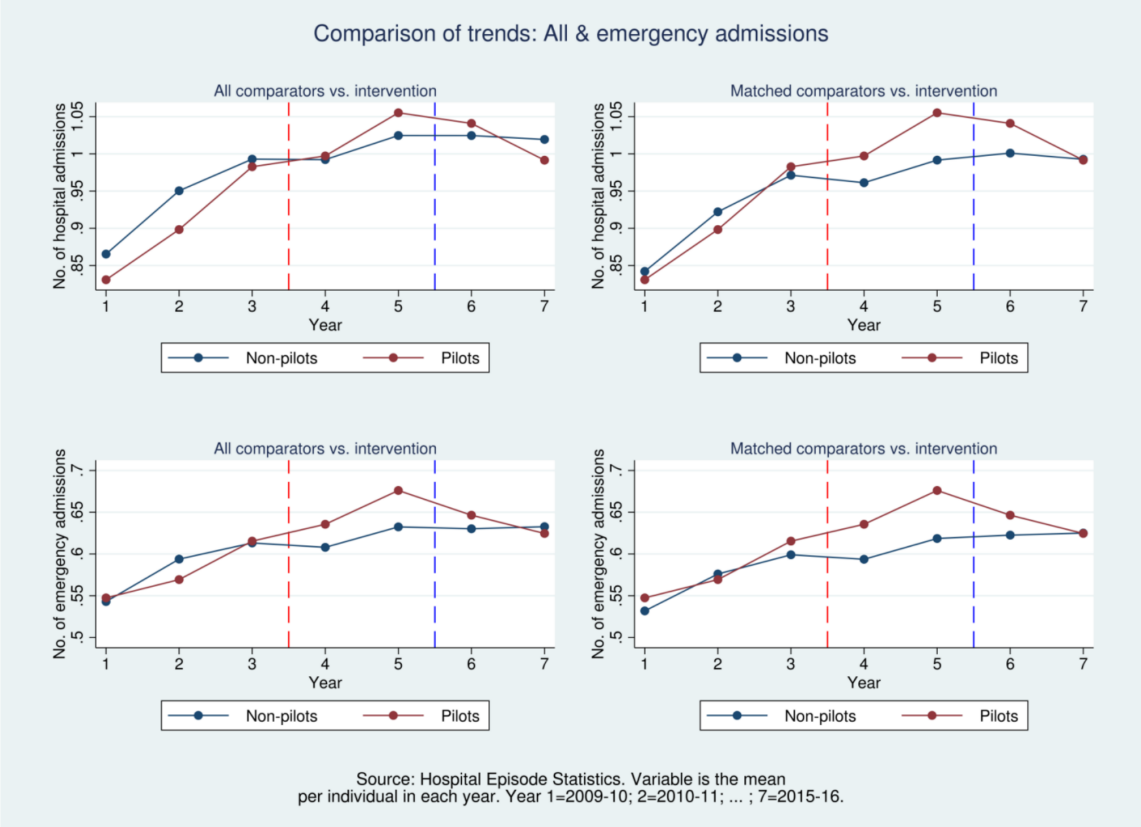
Note: individuals exit the sample as a result of death; total represents person-years

The trends in the outcomes are illustrated in Figures 6b to 6d for: 8 intervention areas versus all 141 comparators; and 8 intervention versus 42 matched comparators. For all hospital admissions (shown in the north-west and north-east quadrants), there is a more pronounced increase in the mean number of hospital admissions per person per year for the intervention areas relative to matched comparators only (Figure 6b). This is similarly the case for emergency admissions (shown in the south-west and south east quadrants), albeit less markedly. The pre-intervention trends are relatively similar in each quadrant of Figure 6b, and it is noticeable that there appears to be a differential increase in the mean of the outcomes in the intervention areas in 2011-12 and 2012-13; and then a return to a lower level in 2014-15 and 2015-16. Relative to this, the comparison areas exhibit more similarity in terms of the level of the outcome in the period between 2011-12 and 2015-16.

For admissions that are classified into drug related and non-drug related groups, the patterns over time are more similar for all comparators versus intervention areas compared with the matched comparisons which show a more obvious divergence in trends over the study period (Figure 6c). It is again clear that divergence is greatest in 2011-12 and 2012-13; after which the levels of the outcome in the intervention areas return to a lower level.

For costs and LOS, an examination of the trends over time reveals a somewhat different pattern (Figure 6d). The comparison areas have higher baseline mean LOS – and the mean in the intervention areas is 'catching up' for the period 2009-10 to 2011-12. The means are then higher for intervention areas in the first year during P4P (2012-13), followed by a reduction to lower levels until 2015-16. Costs show patterns

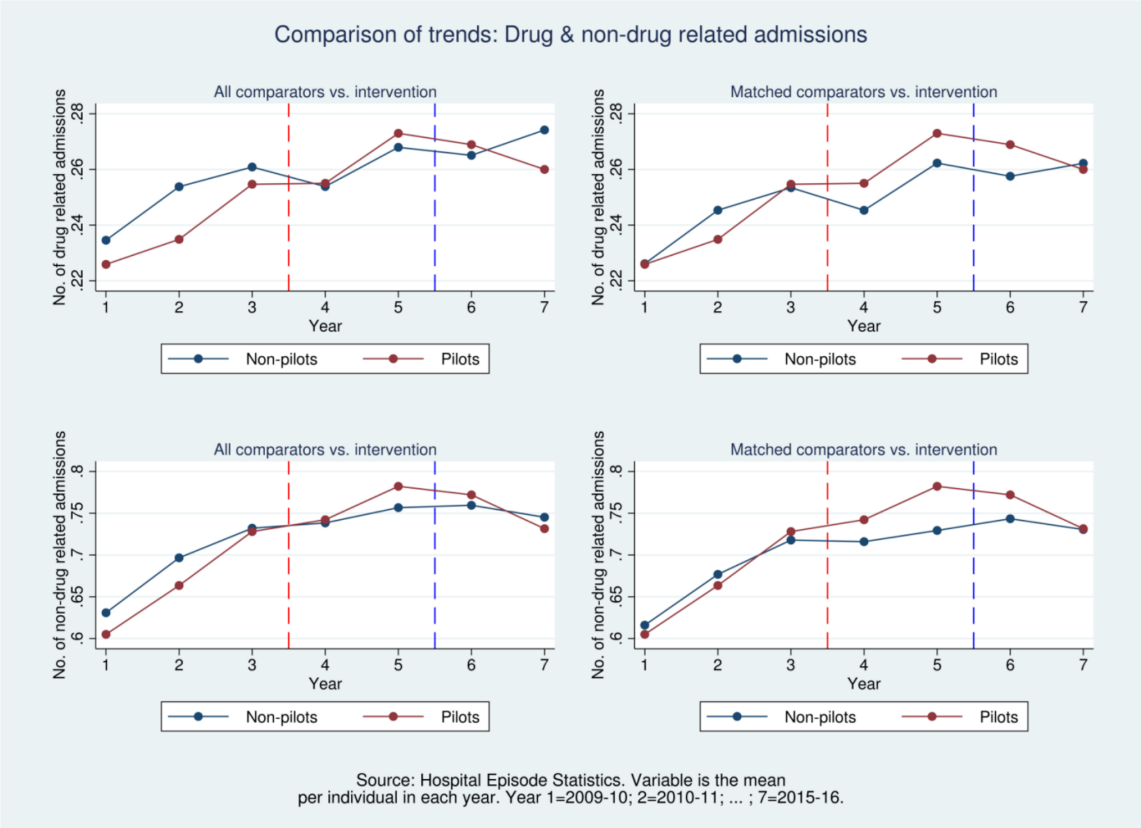
Figure 6b: Comparison of trends in admissions and emergency admissions



Note: dashed lines represent pilot introduction and discontinuation dates.

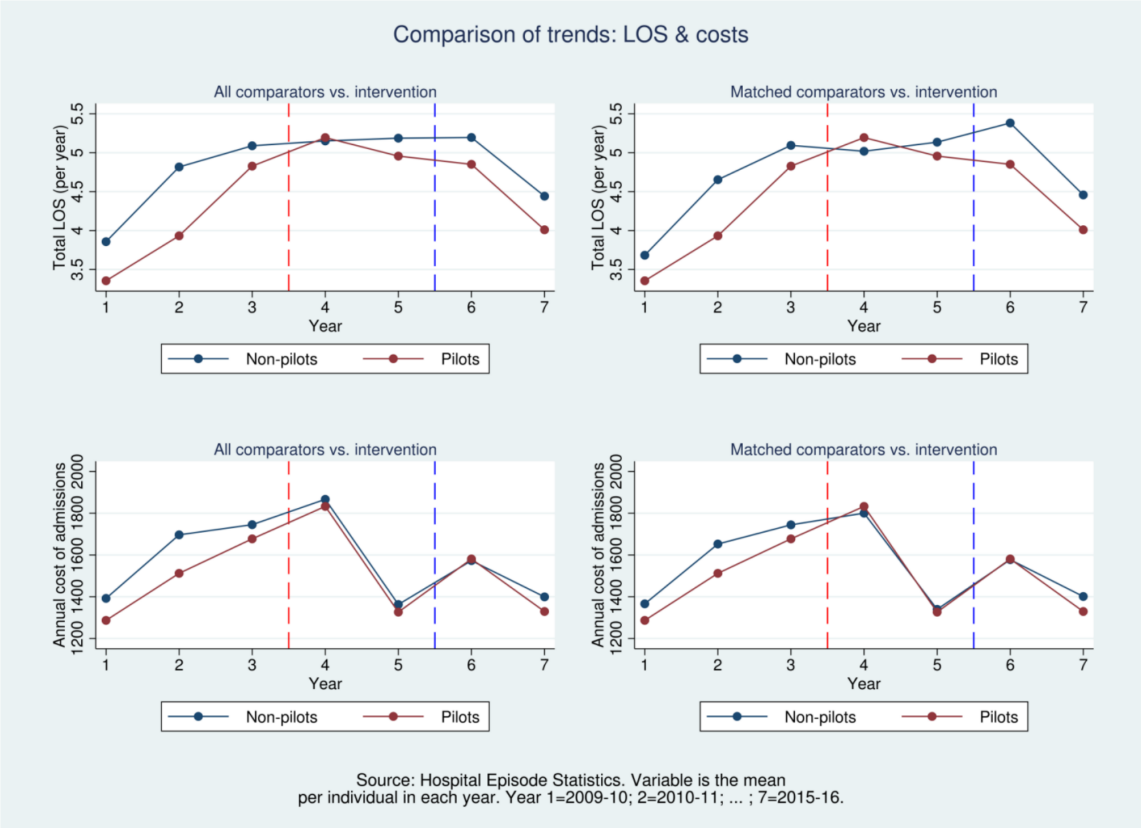
that essentially reflect the combined effects of LOS and numbers of admissions. The component of variation that can change in costs between intervention and comparison areas other than LOS and the number and type (i.e. admission method) of the admissions; is the distribution of admissions across HRGs (i.e. the prices categories that the admissions fall into).

Figure 6c: Comparison of trends in drug related and non-drug related admissions



Note: dashed lines represent pilot introduction and discontinuation dates.

Figure 6d: Comparison of trends in costs and total LOS



Note: dashed lines represent pilot introduction and discontinuation dates.

6.2.2 Regression analyses

The preferred model specification for accounting for the count data properties of the outcomes was negative binomial regression. This choice is supported by comparing the variance and mean for each variable, and the size and significance of the overdispersion parameter in negative binomial regression (see for example Table 6.5). The sensitivity of the estimated treatment effects to this choice can be compared in Appendix Tables A6.1 to A6.8.

Appendix Tables A6.1 to A6.8 also compare estimates from DiD and LDV estimation; and set out the results of testing the assumption of parallel trends. The alternative approaches both in terms of DiD and LDV; and in terms of count data estimation are demonstrated for combinations of the sensitivity checks (matched comparison areas; treatment effect on treated). LDV methods were preferred to matching approaches based on failing to obtain reliable matching variables according to Rubin's R and Rubin's B statistics (Rubin's B $>$ 50 for all models estimated) (Rubin 2001).

6.2.2.1 All hospital admissions

Males had 0.822 times as many hospital admissions per years as females (95% CI [0.811; 0.832]) (Table 6.5). Admissions are higher across age groups and deprivation and were generally higher in later years. Black individuals had a higher rate of hospital admissions compared with White individuals (IRR=1.256; 95% CI [1.174; 1.344]).

For individuals initially resident in intervention areas, hospital admissions were 1.044 times higher after the introduction of P4P compared with non-intervention areas (95% CI [1.016; 1.073]). This equates to an average partial effect of 0.043 hospital admissions per person per year in the study population (95% CI [0.028; 0.058]); or 430 additional admissions per year per 10,000 population (Table 6.6). This finding is robust to both sample restrictions (Appendix Tables A6.1 to A6.4).

The test of parallel trends shows that the null hypothesis of parallel trends could not be rejected (P>Chi-sq=0.849). LDV estimation, however, reduced both the estimated size and significance of the treatment effect – the extent to which it does so depends on the sensitivity restrictions applied – the most restricted LDV model, estimated for non-movers and matched comparison areas only, individuals residing in

Table 6.5: Full regression results - number of admissions per year

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	0.977	0.015	-1.520	0.129	0.949	1.007
DiD	1.044	0.015	3.070	0.002	1.016	1.073
Year						
2009-10	0.906	0.005	-19.360	0.000	0.897	0.915
2010-11	0.982	0.005	-4.020	0.000	0.973	0.990
2011-12	1.016	0.004	4.300	0.000	1.009	1.023
2013-14	1.026	0.004	7.530	0.000	1.019	1.033
2014-15	1.013	0.004	2.930	0.003	1.004	1.022
2015-16	0.999	0.005	-0.200	0.839	0.989	1.009
Ethnicity						
Asian	0.991	0.025	-0.350	0.730	0.944	1.041
Black	1.256	0.043	6.610	0.000	1.174	1.344
Mixed	0.985	0.035	-0.420	0.676	0.918	1.057
Other	0.969	0.020	-1.580	0.114	0.931	1.008
IMD Dec.						
1	0.774	0.011	-18.020	0.000	0.753	0.796
2	0.878	0.014	-8.460	0.000	0.851	0.905
3	0.927	0.013	-5.340	0.000	0.902	0.953
4	0.969	0.014	-2.210	0.027	0.942	0.996
6	1.038	0.016	2.480	0.013	1.008	1.069
7	1.041	0.014	2.940	0.003	1.014	1.070
8	1.069	0.016	4.590	0.000	1.039	1.100
9	1.033	0.014	2.350	0.019	1.005	1.062
10	0.976	0.016	-1.490	0.135	0.946	1.008
Age						
16-25	0.627	0.006	-46.700	0.000	0.615	0.640
26-35	0.702	0.007	-34.190	0.000	0.688	0.717
36-45	0.814	0.008	-20.440	0.000	0.798	0.830
56-65	1.325	0.020	18.620	0.000	1.286	1.365
66-75	1.667	0.033	25.750	0.000	1.603	1.733
76+	1.569	0.029	24.200	0.000	1.513	1.627
Male	0.822	0.005	-30.250	0.000	0.811	0.832
Constant	1.279	0.016	19.170	0.000	1.247	1.311
ln α	0.811	0.006			0.798	0.823
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

Table 6.6: Summary of DiD estimates for preferred specification

	Compared to all comparison areas (N=3,893,165)				Compared to matched areas (N=1,452,303)			
Variable	IRR	95% CI	APE	95% CI	IRR	95% CI	APE	95% CI
Number of hospital admissions	1.044**	[1.016; 1.073]	0.043	[0.028; 0.058]	1.044**	[1.014; 1.075]	0.043	[0.029; 0.057]
Number of emergency admissions	1.043***	[1.021; 1.066]	0.027	[0.019; 0.034]	1.036***	[1.012; 1.059]	0.022	[0.016; 0.028]
Number of drug related admissions	1.041**	[1.015; 1.067]	0.01	[0.009; 0.012]	1.039**	[1.012; 1.067]	0.01	[0.009; 0.012]
Number of non-drug related admissions	1.044*	[1.007; 1.082]	0.032	[0.016; 0.048]	1.043*	[1.004; 1.084]	0.031	[0.017; 0.046]
Total cost of admissions	1.056**	[1.024; 1.089]	80.659	[39.01; 122.314]	1.054**	[1.021; 1.089]	78.626	[38.273; 118.979]
Total LOS	1.085***	[1.039; 1.132]	0.369	[0.165; 0.573]	1.058*	[1.011; 1.107]	0.258	[0.113; 0.404]

Notes: *p<0.05; **p<0.01; ***p<0.001;; models estimated using negative binomial regression (preferred specification).

intervention areas had 1.018 (95% CI [1.006; 1.031]) times as many hospital admissions after the introduction of the policy compared with matched comparison areas only (Appendix Table A6.4).

6.2.2.2 Emergency admissions

Males had 0.925 times as many hospital admissions per years as females (95% CI [0.917; 0.933]) (Table 6.7). White individuals had the highest rates of emergency admissions comparing across ethnic groups.

For individuals initially resident in intervention areas, emergency admissions were 1.043 times higher after the introduction of P4P compared with non-intervention areas (95% CI [1.021; 1.066]). This equates to an average partial effect of 0.027 hospital admissions per person per year in the study population (95% CI [0.019; 0.034]); or 270 additional emergency admissions per year per 10,000 of the treated population (Table 6.6). The estimated treatment holds for various sensitivity restrictions – and is estimated to be larger for a comparison between the intervention areas and the matched comparison areas for non-movers (IRR=1.039; 95% CI [1.015; 1.065] (Appendix Table A6.4). The null hypothesis of parallel trends could not be rejected (P>Chi-sq=0.317) (Appendix Table A6.4). LDV estimation produces a similar estimate of the treatment effect as DiD (IRR=1.034; 95% CI [1.02; 1.047]) (Appendix Table A6.4).

Table 6.7: Full regression results – emergency admissions per year

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	1.004	0.010	0.420	0.676	0.984	1.025
DiD	1.043	0.012	3.840	0.000	1.021	1.066
Year						
2009-10	0.926	0.004	-18.170	0.000	0.919	0.934
2010-11	0.998	0.004	-0.510	0.607	0.990	1.006
2011-12	1.021	0.004	5.850	0.000	1.014	1.028
2013-14	1.034	0.003	9.990	0.000	1.027	1.041
2014-15	1.015	0.004	4.040	0.000	1.008	1.023
2015-16	1.010	0.004	2.390	0.017	1.002	1.018
Ethnicity						
Asian	0.856	0.010	-13.810	0.000	0.837	0.875
Black	0.982	0.026	-0.680	0.494	0.933	1.034
Mixed	0.892	0.016	-6.480	0.000	0.862	0.923
Other	0.951	0.013	-3.820	0.000	0.927	0.976
IMD Dec.						
1	0.725	0.007	-34.650	0.000	0.712	0.739
2	0.822	0.008	-21.020	0.000	0.807	0.837
3	0.900	0.009	-11.130	0.000	0.883	0.917
4	0.959	0.010	-4.200	0.000	0.940	0.978
6	1.038	0.010	3.840	0.000	1.018	1.058
7	1.068	0.011	6.700	0.000	1.048	1.089
8	1.094	0.011	9.100	0.000	1.073	1.116
9	1.081	0.011	7.620	0.000	1.060	1.103
10	0.990	0.010	-1.020	0.305	0.971	1.009
Age						
16-25	0.661	0.005	-58.180	0.000	0.652	0.670
26-35	0.740	0.005	-44.420	0.000	0.730	0.750
36-45	0.876	0.006	-21.040	0.000	0.865	0.887
56-65	1.171	0.010	19.160	0.000	1.152	1.190
66-75	1.376	0.013	33.510	0.000	1.351	1.402
76+	1.538	0.011	58.640	0.000	1.516	1.561
Male	0.925	0.004	-17.860	0.000	0.917	0.933
Constant	0.746	0.007	-33.180	0.000	0.733	0.759
ln α	0.906	0.004			0.899	0.914
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

6.2.2.3 Drug related admissions

There were no significant differences across gender in rates of admissions with a diagnosis of drugs misuse (IRR=0.997; 95% CI [0.990; 1.004]) (Table 6.8).

Drug-related admissions exhibit a different pattern across age to all types of hospital admissions. As such, the estimated IRRs for age groups follow a U-shape (Table 6.8). The IRRs indicate the multiplicative difference between age groups in terms of rates of drug related admissions – and the rates are relatively highest for 36-45 year olds (IRR=1.07; 95% CI [1.056; 1.084]. The next highest are for the reference group 46-55 year olds (IRR=1); and the rates are relatively lowest for the oldest age groups (66-75 and 75+). White individuals had the highest rates of drug related admissions comparing across ethnic groups.

For individuals initially resident in intervention areas, hospital admissions were 1.041 times higher after the introduction of P4P compared with non-intervention areas (95% CI [1.015; 1.067]). This equates to an average partial effect of 0.01 hospital admissions per person per year in the study population (95% CI [0.009; 0.012]); or 100 additional drug related admissions per year per 10,000 of the treated population (Table 6.6).

The estimated treatment effect is similar both in terms of significance and scale for a range of sensitivity restrictions and the test of parallel trends found that the null hypothesis of parallel trends could not be rejected (P>Chi-sq=0.916). LDV estimation reduces the size and precision of the treatment effect estimated.

Table 6.8: Full regression results – drug related admissions per year

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	0.974	0.010	-2.660	0.008	0.956	0.993
DiD	1.041	0.013	3.130	0.002	1.015	1.067
Year						
2009-10	0.931	0.005	-13.030	0.000	0.922	0.941
2010-11	1.003	0.005	0.500	0.619	0.993	1.013
2011-12	1.031	0.005	6.320	0.000	1.021	1.040
2013-14	1.053	0.005	11.090	0.000	1.044	1.063
2014-15	1.038	0.006	6.930	0.000	1.027	1.050
2015-16	1.065	0.006	11.550	0.000	1.054	1.076
Ethnicity						
Asian	0.773	0.006	-31.420	0.000	0.761	0.786
Black	0.867	0.011	-11.590	0.000	0.846	0.888
Mixed	0.949	0.015	-3.420	0.001	0.920	0.978
Other	0.890	0.009	-10.950	0.000	0.872	0.909
IMD Dec.						
1	0.808	0.006	-28.300	0.000	0.796	0.820
2	0.865	0.007	-18.230	0.000	0.852	0.879
3	0.905	0.007	-12.570	0.000	0.891	0.919
4	0.952	0.009	-5.140	0.000	0.934	0.970
6	1.023	0.008	2.740	0.006	1.006	1.039
7	1.055	0.008	6.650	0.000	1.038	1.072
8	1.082	0.009	9.500	0.000	1.065	1.100
9	1.100	0.010	10.960	0.000	1.081	1.119
10	1.077	0.009	8.690	0.000	1.059	1.095
Age						
16-25	0.850	0.006	-23.970	0.000	0.839	0.862
26-35	0.997	0.007	-0.440	0.660	0.984	1.010
36-45	1.070	0.007	10.150	0.000	1.056	1.084
56-65	0.875	0.007	-15.750	0.000	0.860	0.889
66-75	0.750	0.006	-34.160	0.000	0.738	0.762
76+	0.752	0.007	-32.070	0.000	0.739	0.765
Male	0.997	0.004	-0.810	0.420	0.990	1.004
Constant	0.274	0.002	-155.170	0.000	0.270	0.279
ln α	0.584	0.008			0.567	0.601
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

Table 6.9: Full regression results – non-drug related admissions per year

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	0.980	0.020	-1.000	0.316	0.943	1.019
DiD	1.044	0.019	2.330	0.020	1.007	1.082
Wave						
2009-10	0.899	0.006	-16.110	0.000	0.887	0.911
2010-11	0.974	0.006	-4.360	0.000	0.963	0.986
2011-12	1.011	0.005	2.310	0.021	1.002	1.020
2013-14	1.015	0.004	3.480	0.001	1.007	1.024
2014-15	1.002	0.006	0.380	0.704	0.991	1.013
2015-16	0.974	0.006	-4.040	0.000	0.962	0.987
Ethnicity						
Asian	1.075	0.034	2.320	0.020	1.011	1.143
Black	1.430	0.063	8.180	0.000	1.313	1.558
Mixed	0.999	0.050	-0.030	0.980	0.905	1.102
Other	0.999	0.026	-0.050	0.960	0.950	1.050
IMD Decile						
1	0.760	0.014	-14.790	0.000	0.733	0.788
2	0.879	0.018	-6.390	0.000	0.845	0.915
3	0.935	0.017	-3.610	0.000	0.902	0.970
4	0.975	0.018	-1.370	0.172	0.939	1.011
6	1.044	0.021	2.160	0.031	1.004	1.086
7	1.034	0.019	1.860	0.063	0.998	1.072
8	1.063	0.021	3.130	0.002	1.023	1.105
9	1.005	0.019	0.270	0.790	0.969	1.043
10	0.929	0.021	-3.330	0.001	0.890	0.970
Age						
16-25	0.550	0.007	-45.420	0.000	0.536	0.564
26-35	0.602	0.008	-36.670	0.000	0.586	0.618
36-45	0.728	0.010	-23.530	0.000	0.709	0.748
56-65	1.473	0.027	21.180	0.000	1.421	1.526
66-75	1.958	0.045	29.390	0.000	1.872	2.048
76+	1.824	0.040	27.380	0.000	1.748	1.905
Male	0.744	0.006	-34.290	0.000	0.732	0.757
Constant	1.033	0.017	1.930	0.053	1.000	1.067
ln α	1.264	0.007			1.251	1.278
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

6.2.2.4 Non-drug related admissions

Males have considerably fewer hospital admission without a diagnosis relating to drugs misuse (IRR=0.744; 95% CI 0.732; 0.757]) (Table 6.9).

Non-drug related admissions exhibit a more typical age profile with respect to admission to hospital, with older individuals admitted more often. Black (IRR=1.43; 95% CI [1.313; 1.558]) and Asian (IRR=1.075; 95% CI [1.011; 1.143]) individuals have more admission compared to white individuals (reference group).

For individuals initially resident in intervention areas, hospital admissions were 1.044 times higher after the introduction of P4P compared with non-intervention areas (95% CI [1.007; 1.085]). This equates to an average partial effect of 0.032 hospital admissions per person per year in the study population (95% CI [0.016; 0.048]); or 320 additional drug related admissions per year per 10,000 of the treated population (Table 6.6).

The estimated treatment effect is similar both in terms of significance and scale for a range of sensitivity restrictions and the test of parallel trends found that the null hypothesis of parallel trends could not be rejected (P>Chi-sq=0.406) (Appendix Table A6.1). LDV estimation reduces the size and precision of the treatment effect estimated when estimated for all individual (the ITT effect); but the effect is estimated more precisely by LDV for the restricted sample of individuals who were not observed to have changed their area of residence during the study period (IRR=1.029; 95% CI [1.013; 1.046]) (Appendix Table A6.2).

6.2.2.5 Length of Stay

The relative effects of older age on days in hospital per person per year (LOS) are more pronounced than for the previous outcomes discussed (Table 6.10). For example, individuals aged 66-75 have 2.245 times as many bed-days as the reference group aged 46-55 (95% CI [2.193; 2.299]).

Across all non-white ethnic groups, total LOS is relatively higher. For example, Black individuals have 2.18 times as many days in hospital as White individuals (95% CI [2.1; 2.263]). Estimated IRRs for year effects indicate that total LOS is highest in 2011-12 and 2012-13.

Results from the preferred model (Table 6.10) indicate that individuals initially resident in intervention areas had 1.085 more days in hospital after the introduction of P4P compared with individuals located in non-intervention areas (95% CI [1.04; 1.132]). This equates to an average partial effect of 0.369 days in hospital per person per year in the study population (95% CI [0.165; 0.573]); or 3,690 days in hospital per year per 10,000 of the treated population (Table 6.6). This estimated effect is smaller when restricted to matched comparison areas only (APE=0.258; 95% CI [0.113; 0.404]).

However, LDV regression does not support the direction or precision of the estimated treatment effect (IRR=0.997; 95% CI [0.964; 1.03]) for this outcome (Appendix Table A6.5).

Table 6.10: Full regression results - LOS

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	0.894	0.018	-5.490	0.000	0.859	0.930
DiD	1.085	0.024	3.760	0.000	1.040	1.132
Wave						
2009-10	0.792	0.007	-25.170	0.000	0.777	0.806
2010-11	0.969	0.009	-3.550	0.000	0.952	0.986
2011-12	1.016	0.009	1.780	0.076	0.998	1.034
2013-14	0.994	0.009	-0.670	0.503	0.978	1.011
2014-15	0.971	0.009	-3.360	0.001	0.954	0.988
2015-16	0.807	0.007	-24.810	0.000	0.793	0.820
Ethnicity						
Asian	1.152	0.025	6.470	0.000	1.104	1.202
Black	2.180	0.042	40.700	0.000	2.100	2.263
Mixed	1.641	0.053	15.450	0.000	1.541	1.747
Other	1.193	0.026	8.130	0.000	1.144	1.245
IMD Decile						
1	0.676	0.012	-22.870	0.000	0.654	0.699
2	0.799	0.014	-13.100	0.000	0.773	0.827
3	0.878	0.015	-7.840	0.000	0.850	0.907
4	0.961	0.016	-2.360	0.018	0.931	0.993
6	1.023	0.017	1.370	0.170	0.990	1.056
7	1.014	0.016	0.890	0.372	0.983	1.047
8	1.022	0.016	1.390	0.166	0.991	1.055
9	0.967	0.015	-2.070	0.039	0.938	0.998
10	0.833	0.013	-11.580	0.000	0.807	0.859
Age						
16-25	0.522	0.006	-52.670	0.000	0.510	0.535
26-35	0.722	0.008	-31.170	0.000	0.708	0.737
36-45	0.835	0.008	-18.730	0.000	0.819	0.851
56-65	1.462	0.017	33.220	0.000	1.430	1.496
66-75	2.245	0.027	67.180	0.000	2.193	2.299
76+	2.622	0.027	94.950	0.000	2.570	2.674
Male	1.019	0.008	2.510	0.012	1.004	1.033
Constant	5.691	0.086	115.390	0.000	5.526	5.862
ln α	2.508	0.002			2.505	2.512
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

6.2.2.6 Total Costs

The relative effects of older age on total hospital costs year follow a similar pattern (Table 6.11) to the relative effects for length of stay. For example, individuals aged 66-75 have 2.117 times the amount of costs as the reference group aged 46-55 (95% CI [2.125; 2.222]).

Black individuals have 1.393 times the amount of annual hospital costs as white individual (95% CI [2.1; 2.263]). Estimated IRRs for year effects indicate that total costs are highest in 2012-13.

Results from the preferred model (Table 6.11) indicate that individuals initially resident in intervention areas costed 1.056 times more after the introduction of P4P compared with individuals located in non-intervention areas (95% CI [1.024; 1.089]). This equates to an average partial effect of £80.56 per person per year in the study population (95% CI [£39.01; £122.314]); or £805,659 per 10,000 of the treated population (Table 6.6). This estimated effect is smaller when restricted to matched comparison areas only (APE=£78.626; 95% CI [£38.273; £118.979]).

Whilst a null hypothesis of parallel trends could not be rejected (P>chi-sq=0.425), LDV regression does not support the direction or precision of the estimated treatment effect (IRR=1.002; 95% CI [0.987; 1.019]) for this total costs (Appendix Table A6.7).

Table 6.11: Full regression results -Total costs

	IRR	Robust se	z	P>z	95% CI	
Pilot=1	0.939	0.016	-3.730	0.000	0.909	0.971
DiD	1.056	0.017	3.430	0.001	1.024	1.089
Wave						
2009-10	0.776	0.005	-43.250	0.000	0.767	0.785
2010-11	0.943	0.006	-9.770	0.000	0.932	0.954
2011-12	0.914	0.004	-18.770	0.000	0.906	0.923
2013-14	0.715	0.004	-63.600	0.000	0.708	0.723
2014-15	0.815	0.004	-38.390	0.000	0.807	0.824
2015-16	0.704	0.004	-63.820	0.000	0.696	0.711
Ethnicity						
Asian	0.952	0.018	-2.560	0.011	0.917	0.989
Black	1.393	0.032	14.480	0.000	1.332	1.457
Mixed	1.159	0.030	5.630	0.000	1.101	1.220
Other	1.018	0.016	1.150	0.249	0.987	1.051
IMD Dec						
1	0.723	0.009	-25.520	0.000	0.705	0.741
2	0.832	0.011	-14.340	0.000	0.812	0.854
3	0.900	0.011	-8.600	0.000	0.879	0.922
4	0.955	0.012	-3.750	0.000	0.932	0.978
6	1.042	0.013	3.300	0.001	1.017	1.068
7	1.054	0.013	4.410	0.000	1.030	1.079
8	1.058	0.013	4.650	0.000	1.033	1.084
9	1.014	0.012	1.170	0.241	0.991	1.038
10	0.927	0.012	-5.730	0.000	0.904	0.952
Age						
16-25	0.523	0.005	-71.960	0.000	0.513	0.532
26-35	0.665	0.006	-47.470	0.000	0.654	0.677
36-45	0.816	0.007	-24.680	0.000	0.802	0.829
56-65	1.398	0.016	29.630	0.000	1.367	1.429
66-75	2.117	0.028	57.080	0.000	2.063	2.172
76+	2.173	0.025	68.040	0.000	2.125	2.222
Male	0.909	0.005	-17.690	0.000	0.899	0.919
Constant	2317.064	26.779	670.410	0.000	2265.168	2370.148
ln α	2.808	0.001			2.806	2.810
N			3,893,165			

Models estimated using negative binomial regression; IRR=incidence rate ratio; all models clustered by individual; reference groups: 2011-12; White; 5th IMD decile; Aged 46-55; Female.

6.2.2.7 Comparison of treatment effects across all outcomes in 2012-13 and 2013-14 with 2014-15 and 2015-16

Table 6.12 shows the treatment effects for the preferred model in which distinct DiD effects are estimated separately for the two year periods after the introduction of P4P. A pattern of results emerges that is consistent with a visual comparison of the trends illustrated in Figures 6b to 6d.

That is, there is a relatively more pronounced effect for all of the outcomes for the years 2012-13 and 2013-14; whereas the two subsequent years are generally not found to exhibit a differential change between intervention and comparison areas relative to the pre-intervention period. For example, the estimated treatment effect for emergency admissions for 2012-13 and 2013-14 is 1.068 (95% CI [1.044; 1.093]); compared with 1.043 (95% CI [1.021; 1.066]) for the average estimated effect post-intervention.

Table 6.12: Treatment effects on separate DiD terms for during P4P and after P4P

	DiD (2012-13 & 2013-14)		DiD (2014-15 & 2015-16)	
Variable	IRR	95% CI	IRR	95% CI
Number of hospital admissions	1.057***	[1.028; 1.086]	1.031	[0.998; 1.066]
Number of emergency admissions	1.068***	[1.044; 1.093]	1.018	[0.992; 1.044]
Number of drug related admissions	1.057***	[1.029; 1.086]	1.024	[0.994; 1.056]
Number of non-drug related admissions	1.055**	[1.017; 1.093]	1.033	[0.989; 1.078]
Total cost of admissions	1.059***	[1.026; 1.093]	1.053**	[1.015; 1.092]
Total LOS	1.135***	[1.081; 1.191]	1.034	[0.983; 1.087]

Notes: *p<0.05; **p<0.01; ***p<0.001; all models clustered standard errors on individual; models estimated using negative binomial regression (preferred specification) - see equation 5.6.

6.3 Discussion

6.3.1 Summary of main findings

Overall, hospital admissions were differentially higher than expected for individuals initially resident in the intervention areas when the policy was in place. Applied to the study population located in intervention areas, there were an estimated 1,587 additional hospital admissions in 2012-13[7].

This finding is robust: to estimation on only those who didn't change their area of residence; to estimation on only matched comparison areas; and both. The findings are robust to formal testing of the parallel trends assumption of DiD, and results from LDV estimation are generally consistent in terms of the direction of results.

Comparison of the average partial effects with the means of the outcome variables shows that the relative increase in all admissions for individuals in intervention areas (4.39%) is roughly equal to the relative increase when restricting the analysis to emergency admissions only (4.44%). This is indicative that the increase in admissions was roughly equally distributed between elective and emergency admissions. Comparison of admissions that were and were not related to drugs shows that the differential increase in hospital admissions was similar for both classifications.

The direction of DiD estimates of the relative differential change in LOS and costs were not robust to LDV estimation, which generally estimated that there was no statistically significant effect.

The treatment effects are concentrated in 2012-13 and 2013-14 – to the extent that: treatment effects for these two years are considerably larger than the average DiD effects; and there is generally no statistically precise effect estimated for 2014-15 and 2015-16. This is consistent with the unconfounded trends illustrated.

6.3.2 Explanation of notable findings

This study defined a population with a diagnosis of substance misuse according to their hospital records and analysed their activity over a seven-year period. During this period, a P4P scheme was introduced in eight areas that policymakers intended

[7] APE(0.0430)x36,898=1,586.61

to act as a hard mechanism in refocusing substance misuse treatment towards 'recovery'. Specifically, this policy explicitly linked providers' payments to several substance misuse related outcomes including: cessation of injecting; reliable reduction in substance misuse; within-treatment abstinence; discharge abstinent from drugs (successful completion); non-representation for treatment following discharge.

Previous studies have considered impacts on the substance misuse treatment population (Mason et al. 2015b; Jones et al. 2017; Disley et al. 2017). These studies found that, for the population in treatment for substance misuse, P4P areas had differentially: lower rates of treatment initiation; lower rates of treatment completion; lower rates of treatment completion without subsequent representation; higher waiting times; higher rates of declining to continue with treatment; higher rates of within treatment abstinence; and higher rates of within treatment injecting cessation.

The differentially higher numbers of hospital admissions in this study population is consistent with all of the above findings. As such, the combined evidence from this chapter and the studies mentioned can be potentially explained in terms of changes in access to treatment.

Access can be defined as (Cleary et al. 2012: 1):

> "the degree of fit between population needs and health system responses in terms of affordability [..], availability (fit between patient needs and the type, place and time of services provided) and acceptability (fit between the provider and patient attitudes towards and expectations of each other)."

Two of the three dimensions of this conceptualisation of access are applicable to the setting for this chapter (availability and acceptability) and provide a useful framework for setting out potential reasons for the findings of this and other studies.

First, differentially lower rates of treatment initiation and higher waiting times are consistent with reduced availability of treatment in P4P areas. Providers and commissioners in these areas reported difficulty in implementing the scheme and problems in terms of the time (and monetary) cost involved in the bureaucracy and administration generated by the P4P scheme (Disley et al. 2017: 109).

Second, higher rates of declining to continue with treatment are consistent with issues around the acceptability of treatment. Specifically, the policy was intended to be a

'hard' mechanism for diverting providers' priorities towards abstinence relative to harm reduction. Harm reduction has represented a key pillar of substance misuse treatment over recent decades and its efficacy has been widely demonstrated. For example, retention of individuals in treatment and substitute prescribing of opioid agonist pharmacotherapy (OAP) has been demonstrated to be effective at retaining patients in treatment and reducing heroin use (Mattick et al. 2009; Mattick et al. 2014; Pierce et al. 2015). However, the scheme was intended to reduce incentives for providers to focus their activities on the retention of individuals in treatment, and qualitative (thematic) analysis of interviews with treatment populations in P4P areas after its introduction identified changes to substitute prescribing in these areas as a problematic issue (Disley et al. 2017: 110):

> "I've been stable and just doing my prescribed medication for over ten years now. After a couple of false starts of actually getting clean; I've devised a different strategy. I had their word that they weren't going to touch my prescription until certain things had been done. Basically, I went in a fortnight later and they just cut me down and didn't even offer me anything to replace it with. They've totally turned my life upside at a time when I was actually seeing light at the end of the tunnel. My worker has been straight with me and said, 'It's down to money' ".

This is actually also compatible with observed increases in within treatment abstinence and higher rates of injecting cessation (Jones et al.: 2016). As such, the policy intended to reduce substitute prescribing, and move policy away from retention in treatment and towards discharging individuals abstinent. However, it was not stated that an intentional consequence of the policy would be to alter the acceptability of treatment such that there would be reduced engagement with treatment services.

Reduced access leads to higher levels of unmet need, where need is defined as the capacity to benefit from treatment (Culyer & Wagstaff 1993). The overall evidence from this chapter and previous studies is indicative of a scheme that: reduced access to substance misuse treatment in the prevalent population - leading to higher levels of unmet need - which in turn led to increased use of hospital services – more pronouncedly for emergency hospital admissions. This picture is consistent with other studies which have demonstrated that engagement with treatment services for substance misusers reduced hospitalisations (Neighbors et al. 2005).

Table 6.12 shows that the differential increase in hospital admissions was concentrated in 2012-13 and 2013-14; and that differences in 2014-15 and 2015-16 were

not statistically significant. This is consistent with how policy developed after the two official 'pilot' years. After the pilot ended, commissioners in all areas were given autonomy regarding whether they retained or adopted P4P into their local reimbursement model (and all specifics – to what extent, what was incentivised, and so on). Generally, commissioners in P4P areas indicated in interviews that they intended to incorporate at least an element of P4P into their reimbursement model – but only one area retained a scheme that was fully based on P4P, and other areas substantially revised its scale and scope, with greater emphasis placed on process indicators rather than outcomes (Disley et al. 2017: 112-114). It is unclear as to why there were not statistically significant differences in 2014-15 and 2015-16 in hospital admissions. It could possibly indicate: that P4P generally ceased in practice in the eight intervention areas; that P4P was improved based on lessons learned during the pilot period; or that the intervention period was characterised by implementation issues that were subsequently addressed. It is nonetheless notable that the unconfounded trends in 2014-15 and 2015-16 are characterised by pronounced decreases in hospital admissions in the intervention areas, and non-pronounced increases in hospital admissions in the comparison areas.

6.3.3 Comparison with existing research

Previous studies of this scheme have evaluated its effects on several treatment outcomes and other measures (Disley et al. 2017; Mason et al. 2015b; Jones et al. 2017). The focus of previous studies was limited to the population in treatment for substance misuse. However, analysis of hospital administrative data has allowed for analysis of effects beyond the treatment population to a wider population with a history of substance misuse. Estimated contact with illicit drugs in the general population is many times higher than the size of the treatment population (NHS Digital 2018).

This study provides evidence relating to unintended consequences for a scheme that was relatively unique insofar as it linked remuneration of providers to treatment outcomes. It adopts a quasi-experimental design for a large sample using patient-level data for England. It formally tests the assumptions of DiD and compares DiD estimates to an alternative, LDV estimation. It costs the relevant hospital activity using prices from the national tariff, which is used to reimburse hospitals. As such, the study is relatively novel both in the approach taken to evaluating the P4P policy, but also the specific policy itself and the setting in which it was introduced. It

complements other evaluations of the scheme and estimates results that fit into a consistent overall story.

6.3.4 Strengths and limitations

The novelty of the policy and the setting in which it was introduced may impact on the generalisability of the findings to other settings. The study was unable to identify individuals who dropped out of the study population as a consequence of moving out of England during the study period.

This study defines a study population based on patient history of a drug-related hospital admission. This definition includes a broad range of ICD-10 codes that have been used in other studies for examination of drug-related mortality (Pierce et al. 2015). Alternative definitions of the study population could include fewer ICD-10 codes related to substance misuse; or could restrict examination to only instances in which a substance misuse related ICD-10 code populates the primary diagnosis field in HES. This study includes primary and secondary codes based on the findings of other studies which suggest that primary codes are typically related to a presenting symptom rather than the drug use which is the underlying cause (Shah et al. 2011). However, use of secondary codes allows for inclusion of activity not only attributable to substance misuse, but also other co-morbidities and interactions between different conditions and health states.

The data analysed in this chapter are for inpatients only and therefore do not reflect all hospital activity. It is also possible that changes over time are attributable to change in the recording practices in hospitals – for example improvements in the recording of secondary diagnoses. This particular concern is material only for situations in which recording practices changed differentially in intervention and comparison areas. Some caution is required in considering HES data – drug use may only be suspected and therefore not recorded by hospitals (NHS Digital 2018a).

Individuals are fixed according to their initial residence, and then allocated to intervention or control areas. This assumption was made due to concerns about endogenous selection of area of residence by individuals. The results were not found to be sensitive to this when restricting the analysis to individuals who were not recorded as changing their area of residence within the study period.

6.3.5 Implications for policy and research

The findings of this study imply that the introduction of P4P schemes for providers of public services can lead to unintended consequences which can have economically meaningful impacts on other public services. It is therefore generally advisable that assessment of policy impacts should consider the wider implications across other settings, rather than a partial assessment of the sector delivering the service.

This study specifically focuses on a service that relates to treatment for a heterogeneous population with complex needs across a range of areas of public policy (including health care, psychosocial support, housing support, social services, etc). The introduction of a P4P scheme which was primarily focused on paying treatment providers for the successful achievement of particular treatment endpoints such as successful completion of treatment and abstinence from drugs, had a range of consequences. Overall, the evidence from this and other studies would suggest that the policy had negative consequences both in terms of treatment outcomes and in terms of the economic impact on the need for hospital care. It is debateable to what extent these negative consequences were unintended. The policy intended to shift the focus of treatment away from retention in treatment towards abstinence or 'recovery'. Whether it was foreseeable that this shift would create reduced access to treatment is not a worthwhile speculation – but the overall evidence clearly points in this direction. Given the context, the evidence from this chapter leads to an inference that is consistent with the wider P4P literature: P4P is most effective when targeted towards conditions that have clear 'best practice' pathways which are applicable to a clear majority of the target population (such as emergency care for Stroke) (Mason et al. 2015b). Implementing P4P for care provided to a heterogeneous population with complex needs, characterised by a diverse market of providers, in which payments were primarily targeted towards treatment endpoints/outcomes, would appear to have been misconceived.

Future research could expand this analysis to consider the impacts on the scheme for another variable of interest in the substance misuse field, such as homelessness. Information on individuals' area of residence in HES includes "no fixed abode" as a value. The data therefore offer a resource for consideration of homelessness in this population, and more generally. Future studies might also want to consider other hospital services not considered herein such as outpatient and accident and emergency attendances, and use of primary care.

7 Discussion

This thesis set out to answer empirical questions relating to substance misuse for England between 2009-10 and 2015-16 by using various methods common to several fields such as epidemiology and social statistics, but especially common to econometrics. In particular, the objective was to generate novel evidence on: the volume and cost of substance misuse related hospital admissions and how these vary according to approaches common in the related literature; patterns of individuals' hospital admissions and costs relative to diagnoses of substance misuse in hospital; and how introduction of a P4P scheme for providers of substance misuse treatment services affected both in-treatment outcomes, and wider hospital admissions and costs for individuals with a history of admission to hospital with a diagnosis related to substance misuse. Collectively, these studies update the evidence base relating to substance misuse hospital admissions and associated costs with more recent information, and show introduction of a relatively novel P4P scheme affected both hospital admissions and costs - in addition to its direct effects on in-treatment outcomes for individuals treated by providers located in geographical areas that participated in the P4P scheme.

These studies are also linked by their focus on substance misuse and in their use of econometric methodology. They each use retrospective, observational administrative data – HES and NDTMS – linked with other publicly available data such as the NTPS. Chapter 2 informed the design of subsequent chapters – demonstrating the skewness of hospital activity and cost data, and the economic difference in use of primary or any ICD diagnoses in identifying substance misuse, leading to a comparison of methods for over-dispersed data and the use of primary and secondary ICD diagnoses in identifying a population of interest. Chapter 3 followed from the limitations of the approaches used in Chapter 2 for identifying hospital activity attributable to substance misuse. Chapters 4 and 5 accompany each other insofar as data on individuals in structured treatment for drug addiction are complemented by allowing for coverage of individuals who do not receive structured treatment, but who have a substance misuse related hospital admission between 2009-10 and 2015-16. The exploration of costing approaches in earlier chapters allowed for use of costs in analysis in Chapter 5.

The following discussion outlines the main objectives and findings of each study,

considers the strengths and limitations of this thesis, and sets out implications for policy and research.

7.1 Summary of thesis findings

Chapter 2 of this thesis explored alternate approaches that are used in studies that seek to estimate the cost of illness. Specifically, it focused on two main features by which COI studies vary and compared how these alternate approaches affect estimates of the disease specific costs in the case of substance misuse related hospital admissions. To do this, HES data with national coverage for England for 2009-10 to 2015-16 were linked to NTPS data and NHS Reference Costs for the same period. The first comparison implemented alternate approaches to estimating the direct costs of illness - known as top-down and bottom-up costing. The results indicated that top-down costing, particularly in the case of application of national unit costs to disease specific activity, can lead to estimates that are dramatically different to bottom-up costing which represents a replication of the methodology used to calculate hospital reimbursement. The second comparison illustrated the effects of using primary diagnoses only for identifying relevant hospital activity in estimating direct hospital costs compared with using any of the twenty diagnosis codes in HES. The results indicated that use of any diagnoses increases the included activity by a factor of around ten in the case of substance misuse; and affects the distributional properties of the costs included. A limitation of the latter approach for identifying hospital admissions and costs attributable to substance misuse is that it attributes all of the costs to substance misuse and fails to account for the impact of comorbid conditions. Exploration of the characteristics of the populations for activity included based on any or primary only diagnoses indicated that any diagnosis leads to inclusion of activity for a relatively older and more female population; and the LOS and mean cost is relatively and significantly higher for activity included based on any diagnosis of substance misuse. Finally, the distribution of activity according to which HRG admissions are assigned to varies considerably comparing these two groups of activity: admissions for primary diagnoses only are relatively homogenous – over 85% of all admissions are charged to four HRG codes, and over 99% of these admissions are non-elective. These differences highlight the importance of approaches which are able to identify the correct degree of attribution for particular conditions – there are diverse underlying distributions across costs and activity that reflect epidemiological difference.

Chapter 3 followed on from this by applying an event study to HES data for 2009-10 to 2015-16, defining an event as being admitted to hospital with any substance misuse related diagnosis. This chapter set out to explore this approach as an alternative to other methods (such as the matching cohorts with and without a particular condition) for identifying the impact of a particular condition on activity and costs. The study population was defined as any individual who had at least one such admission over the analysis period, and the outcomes of interest were the number of admissions per person per year (including separate analyses of emergency, drug related and non-drug related admissions only), the costs of these admissions, and the total LOS. The results of the chapter indicated that, relative to an occurrence of admission to hospital diagnosed as being related to substance misuse, individuals' hospital admissions increase steadily in the pre-event period, spike in the year of the event (in part due to the event definition), and decrease in the year after the event to a level substantially higher than in a pre-event period and reduce steadily thereafter but remain at a persistently higher level than before the occurrence. This result had a common pattern across the various outcomes, and the effects were found to be relatively more pronounced for non-elective admissions and LOS. Results were tested for sensitivity to the modelling approach and design of the study and were found to be consistent across robustness checks. Various potential explanations were set out for these findings: increasing use of hospital services prior to the definition of the event could indicate (undiagnosed) drug use leading to deteriorating health and increase in admission to hospital; and/or could indicate deteriorating health that in turn leads to an increased probability of substance misuse. Chapters 2 and 3 are complementary insofar as whilst the former looks at cost of admission where there is a diagnosis of substance misuse at a population-level, the latter constitutes an alternative approach which considers patterns of admissions and costs relative to diagnosis of substance misuse at the individual level for individuals who at some point are observed to have such a diagnosis.

Chapter 4 is an evaluation of a P4P policy 'pilot' that was introduced for providers of treatment services for individuals with problematic substance misuse in eight geographical areas in England in April 2012 using individual-level national administrative data from NDTMS. Policymakers intended this policy to increase the relative prioritization of discharging individuals from treatment abstinent from drugs by heavily linking provider reimbursement to achievement of this and related outcomes. The study applied difference-in-difference analysis to the first nine months of data from the pilot, comparing changes in two outcomes before and after the introduction of P4P: successful completion of treatment abstinent from drugs (long considered

an indicator of treatment effectiveness); and declining to continue with treatment (included to proxy for disengagement with treatment). The chapter found that the probability of the successful completion of treatment was differentially reduced in the intervention areas; that declining to continue was differentially higher; findings which were robust to a range of sensitivity checks. This was tentatively explained as being potentially caused by more risk averse discharge practices of providers located in pilot areas, whose income was increasingly dependent on individuals not relapsing in their abstinence from drugs such that they would subsequently return to structured treatment or the criminal justice system after discharge.

Finally, Chapter 5 extended assessment of the impact of this P4P pilot to not only the population in structured treatment for substance misuse, but to a wider population with a record of admission to hospital with a diagnosis of substance misuse. This allowed for complementary assessment in terms of capturing possible unintended consequences of the scheme on dimensions of access to treatment for addiction to drugs: acceptability and availability. The study applied difference-in-differences analysis to HES data before and after the pilot's introduction and found that individuals' hospital admissions were differentially increased in pilot areas. This effect was concentrated in the years during the pilot's operation, and the difference between pilot and non-pilot areas relented after its termination. The study findings were robust to a variety of sensitivity checks and modelling approaches. Findings were explained as being potential consequences of changes in access to addiction treatment services, given that other studies had found that the scheme had resulted in reduced engagement with services and increased waiting times in the pilot areas. The findings of both Chapter 5 and Chapter 4 are consistent with the results of other studies that have evaluated this specific P4P pilot and are also consistent with themes identified in Chapter 1 of this thesis in the wider health P4P literature.

7.2 Strengths and limitations

7.2.1 Strengths

This thesis uses administrative data with comprehensive national coverage over a long time period, providing considerable power to test for statistical differences in outcomes at an individual level. The data allow for regression analyses to adjust for confounding across demographic and socioeconomic characteristics known to affect a range of health outcomes.

The time period covered allowed for the identification of common time trends in the outcome variables. In the case of the event study in Chapter 3, this allowed for the removal of common effects such as changes to the national tariff that were observed in Chapter 2 that affect estimates of changes over time in individuals' hospital costs. These changes to tariff prices and the effect on estimates of individuals' costs were observed to occur at various points, but most prominently at the same time as the introduction of the P4P pilot in April 2012-13. These changes affect both the intervention and comparison groups in the evaluation of P4P in Chapter 5, but the distribution of admissions across HRGs can vary both across individuals and over time. The findings of Chapter 2 and 3 therefore support the use of difference-in-differences approaches thereafter.

All chapters using HES data sought to complement analysis of admission costs with analyses of LOS and counts of admissions – for several classifications of admission (emergency; with a diagnosis of substance misuse; etc). This allowed for consideration of whether results were sensitive to how resource use is measured. The comparisons made in Chapter 2 in terms of whether substance misuse related hospital activity is defined based on primary codes only or any ICD diagnosis of substance misuse informed use of cost data thereafter. The strength of this approach lies in its capacity for identification of activity and study populations not only where substance misuse is the presenting condition, but also where another component (such as cardiac arrest) is the presenting health condition – but substance misuse is an underlying cause of the admission (such as chronic cocaine use) and is therefore coded as such in hospital. Chapter 2 provided tentative indication that use of primary and secondary diagnoses potentially captures hospital activity relating to the chronic health effects of substance misuse in addition to acute health events relating to substance misuse. The limitations of this approach are explored below.

An additional strength of the approach used for identifying relevant activity and populations in terms of the ICD-10 codes included lies in its capacity to capture activity in which the substance misuse is considered diagnosable in hospital, but the substance of interest could not be identified. Other studies have characterized how such admissions are coded in hospital (Shah et al 2012), and the ICD-10 codes used herein are not only consistent with other approaches (Public Health Wales 2016; Home Office 2013), but also allow for inclusion of such activity – which is a strength in a policy context of increased concerns relating to legal highs and novel psychoactive substances.

7.2.2 Limitations

This thesis uses primary and secondary ICD diagnoses for defining hospital activity related to substance misuse. This is because use of primary diagnoses can 'greatly underestimate the rate of utilization that contributes to other diseases' (Ward et al. 2000). Persistent problematic drug use leads to chronic health problems. For example, the long-term health effects of heroin use include respiratory illnesses, weakening of the immune system, muscular weakening and paralysis. Frequent cocaine use can reduce blood flow in the gastrointestinal tract leading to ulcerations and tears, has significant cardiovascular effects and is linked with increased risk of stroke. However, whilst inclusion of hospital admissions of secondary diagnoses related to substance misuse allows for capturing this activity; assignment of costs on this basis is very likely to overstate the amount that costs are directly attributable to substance misuse. Cost assigned on this basis will in some cases be attributable to other co-morbid conditions unrelated to substance misuse.

The scope of the comparison of approaches to identifying activity attributable to substance misuse in Chapter 2 is therefore limited to illustrating scale differences between estimates that results from using either primary diagnoses only with using any diagnosis. Diagnosis codes in HES are ordered arbitrarily – except for the primary diagnosis. For alcohol, there is a 'attributable fraction' variable in HES that allows assessment of how much an admission is attributable to alcohol. For substance misuse, there is no equivalent variable; and there are limitations associated with the use of attributable fractions from published studies relating to insufficient confounding (as discussed). Chapter 3 therefore sought to explore alternative methods for identifying the impact of substance misuse on costs and hospital activity.

A distinct but related issue is that of the extent to which identification of relevant activity based on ICD diagnoses is confounded by diagnosis recording practices. This problem may vary across individuals, substances, diseases and over time. In particular, improvements in secondary diagnosis recording is a known phenomenon since 2006-07 (NHS Digital 2016b) and are likely to affect time trends during the period since then.

A final issue relating to the ICD codes used is that the definition of codes considered as related to substance misuse may have led to the inclusion of some activity unrelated to substance misuse – but related to drug interactions or polypharmacy. This is a particular concern at very young and older ages. This issue was seen in Chapter 2, and subsequent chapters exclude activity at very young ages for which problematic

substance misuse may be considered as either accidental or related to polypharmacy.

This PhD only considers hospital activity in the admitted patient care data in HES. This therefore excludes activity in A&E that prior to admission as an inpatient, or that does not result in an inpatient admission. It also excludes outpatients. Whilst outpatient and A&E attendance data are larger in terms of numbers of visits and attendances, inpatient admissions are typically more expensive - reflecting their focus on acute care.

Chapter 4 considered the population in structured treatment for substance misuse. A limitation of this thesis was that I was unable to link individuals between NDTMS and HES based on an identifier due to data permissions. Linkage would have permitted analysis on individuals with a recording of substance misuse problems according to treatment status.

The results of difference-in-differences analysis in this thesis are tested across various sensitivity models. First, comparisons are repeated for subgroups of comparator areas chosen for their similarity to the intervention areas both in terms of geography and area-level characteristics in terms of deprivation and severity of substance misuse. Second, comparisons were made to between areas based on information obtained relating to the adoption of P4P in some form, instead of between pilot and non-pilot areas. Chapter 4 used data from the first nine months of the P4P scheme in line with data access conditions; but subsequent studies have used data of structured treatment for the full two-year duration of the pilot scheme. These studies set out findings consistent with those contained in Chapter 4 of this thesis.

Chapter 5 complemented data from the first nine months of the scheme with data for covering the entire lifespan of the policy, and two years after this. Difference-in-differences analyses were tested for sensitivity across a range of approaches for overdispersed data, to the identifying assumptions used in the study design, and the assumption of parallel pre-trends required for obtaining unbiased and consistent estimates was tested formally. Further rigour was included in the form of estimation of lagged dependent variable models which are used to estimate policy effects in cases where pre-trends are not assumed to be parallel.

The event study in this thesis tested the sensitivity of results across a range of models for overdispersion, and the inclusion of unit-specific effects to account for time-invariant individual heterogeneity. It also tested the sensitivity of the results to the removal of data for all individuals whose event occurred in the first two years of

HES data; i.e. the assumption that the event was the first occurrence of substance misuse. Substance misuse has known problems relating to relapse and as such an ideal analysis would allow for evaluation of individuals' substance and health care use over a lifetime history. The significance and direction of the results were unaffected by the removal of these individuals.

Chapters 4 and 5 evaluated the effects of the same P4P pilot policy introduced in April 2012. This pilot scheme did not randomly select areas into the policy – 61 out of 141 geographical areas applied to participate in the scheme, from which eight were selected by policymakers in order to be 'nationally representative'. I attempted to test the robustness of the study findings to selection into treatment by both attempting to match intervention areas to similar comparison areas, and to compare pre-trends. However, areas may have differed by unobserved characteristics that potentially confound the analysis. A further complication relates to the fact that comparison areas that did not participate in the pilot scheme were permitted to include some form of P4P into their own commissioning arrangements. Chapter 4 tested the sensitivity of its results to an alternative definition of comparison groups into P4P and non-P4P areas rather than pilot and non-pilot areas.

7.3 Implications for research

Chapter 3 of this thesis applied an event study to hospital data. Event studies have been most commonly used in financial economics – especially for analyzing the effect of shocks (for example macroeconomic and market events) on economic variables such as stock prices. However, studies have begun to explore their use in other areas of research. Their application in this field allows for identification of trajectories relative to some health event or diagnosis. This provides information relevant to policymakers – who may potentially use information for structuring timings of interventions if they decide that trajectories can or should be modified. It also presents an alternative to existing approaches for identifying attributable costs based on matching cohorts.

Future research is nonetheless needed to complement Chapter 3 of this thesis in order to identify similar groups of admissions with and without a diagnosis of substance misuse in order to ascertain how 'much' of an admission can be attributed to substance misuse where such a recording is made in the secondary diagnoses only. Such analysis could be restricted to the most common HRG codes for which activities occur in these populations. This would then allow for a relatively comprehensive

comparison across two of the three methods outlined in Ward et al. (2000): that is - attribution of all costs to activity for primary diagnoses only (outlined in Chapter 2); and attribution of costs for any diagnosis of substance misuse based on comparison of matched cohorts with and without such a diagnosis. Comparison of these two methods with the approach used in Chapter 3 would allow for this comprehensive assessment and understanding of patterns of hospital admission and costs that are attributable to substance misuse.

There are also possible extensions to the specific work in the event study herein. In the substance misusing population, homelessness is a known problem that has a complex interaction with problematic substance misuse. HES data allow for identification of homelessness insofar as authority of residence can take the value 'no fixed abode'. Homelessness and substance misuse can both be defined as 'events', and future research may wish to use an event study approach to explore the interactions between these events and their probabilities and frequency.

7.4 Implications for policy

Chapter 2 examined patterns of hospital admission related to substance misuse using different definitions, and then explored different approaches to estimating their costs. Regardless of the approach, there has been a significant increase in substance misuse related hospital admissions – primary admissions only (posited in this thesis as representing acute events) are increasing more rapidly in percentage terms over the study period. This is consistent with published research on admissions and drug-related deaths (NHS Digital 2016; ONS 2016). This poses a considerable challenge for prevention and treatment policy.

There has been a reduction in the numbers of individuals in structured treatment for drugs misuse over recent years (PHE 2016). This is contemporaneous to increases in substance misuse related hospital admissions and drug related deaths. Structured treatment and harm reduction have well-established efficacy in reducing the damaging health consequences of substance misuse. A renewed challenge exists in improving access to structured treatment and in preventing problematic drug use more widely in society.

The use of P4P for providers of public services has become relatively more widespread over recent decades as policymakers have sought to refocus payments away from pas-

sive forms of reimbursement that failed to establish links between payment and activity or quality. Such schemes became especially popular in healthcare, and schemes for England such as the QOF in primary care and AQ/CQUIN in secondary care linked payment to 'best practice' processes for specific health conditions. The use of P4P in health care in general is contested and controversial, but there is at least some limited evidence that its adoption has had some limited benefits in cases like AQ, where rigorous clinical guidelines exist and relatively heterogenous pathways are well established.

Its extension to areas of public service such as probation and reoffending, addiction treatment services and 'welfare to work' means that government schemes accounted for at least £15 billion of public expenditure in 2015 (NAO 2015). In the case of the scheme evaluated in this thesis, there is clear evidence that the scheme had unforeseen consequences. There is evidence of poorer in-treatment outcomes, reduced engagement with treatment services, and differentially higher levels of substance misuse hospital admissions in pilot areas. These findings are consistent with lessons that are common in the P4P literature more widely, and common to schemes that were introduced at a similar time in England by policymakers. Experiences with these schemes has demonstrated that P4P is a technically challenging form of public contracting that is not necessarily suited to specific public services, and its adoption in specific fields has not been clearly justified in preference over other forms of contracting. Whilst it involves a considerable transfer of risk from commissioners to providers, this can lead to overlooking the risks that remain with payers – namely that providers fail to meet their objectives. Poorly designed schemes can have unintended consequences, and the scheme assessed in this thesis experienced considerable problems in its implementation (Disley et al. 2016). As such, whilst proponents of such schemes are inclined to speculate that P4P schemes offer value in the provision of public services, schemes can in practice be a false economy – their cost-effectiveness is in question. This may particularly be the case in the context of a problem as complex as problematic substance misuse. Helping drug-dependent individuals become abstinent is a long and complex process which frequently takes a range of attempts and approaches. A P4P scheme with a relatively narrow focus may have in practice been unhelpful insofar as it indirectly encouraged a focus akin to an 'acute care model', in contrast to the chronic disease management approach advocated in recent US studies (McLellan et al 2013; USDHHS 2016). The pilot scheme was discontinued after its two-year duration, after which I find reduced evidence of differentially higher hospital admissions. It may have been the case that a sub-population exists for whom the scheme would have proved effective and efficient; but a 'one-size fits

all' approach that focused on a relatively short-time horizon may have led to the scheme's unintended consequences.

8 References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1), 1-19.

Aldington, S., Harwood, M., Cox, B., Weatherall, M., Beckert, L., Hansell, A., ... & Beasley, R. (2008). Cannabis use and risk of lung cancer: a case–control study. European Respiratory Journal, 31(2), 280-286.

Alshamsan R, Majeed A, Ashworth M, Car, J & Millett, C (2010), Impact of pay for performance on inequalities in health care: systematic review, Journal of Health Services Research and Policy, 15(3): 178-84.

Amundsen EJ (2015), Drug-related causes of death: Socioeconomic and demographic characteristics of the deceased. Scandinavian journal of public health, 43(6), 571-579.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. The Review of Economics and Statistics, 47-57.

Ashenfelter, O. C., & Card, D. (1984). Using the longitudinal structure of earnings to estimate the effect of training programs.

Ashton CH (2001). Pharmacology and effects of cannabis: a brief review. The British Journal of Psychiatry, 178(2), 101-106.

Baldwin WA, Rosenfeld BA, Breslow, MJ, Buchman TG, Deutschman, CS & Moore, RD (1993). Substance abuse-related admissions to adult intensive care. Chest, 103(1), 21-25.

Basu, A., & Manning, W. G. (2009). Issues for the next generation of health care cost analyses. Medical care, S109-S114.

Baumberg, B. (2006). The global economic burden of alcohol: a review and some suggestions. Drug and alcohol review, 25(6), 537-551.

Berthiller, J., Straif, K., Boniol, M., Voirin, N., Benhaïm-Luzon, V., Ayoub, W. B., ... & Ayed, F. B. (2008). Cannabis smoking and risk of lung cancer in men: a

pooled analysis of three studies in Maghreb. Journal of Thoracic Oncology, 3(12), 1398-1403.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differ ences-in-differences estimates?. The Quarterly journal of economics, 119(1), 249-275.

Beynon, C. M., McVeigh, J. I. M., & Roe, B. (2007). Problematic drug use, ageing and older people: trends in the age of drug users in northwest England. Ageing & Society, 27(6), 799-810.

Binswanger, I. A., Kral, A. H., Bluthenthal, R. N., Rybold, D. J., & Edlin, B. R. (2000). High prevalence of abscesses and cellulitis among community-recruited injection drug users in San Francisco. Clinical Infectious Diseases, 30(3), 579-581.

Birnbaum HG, Kessler RC, Lowe SW, Secnik K, Greenberg PE, Leong SA. Costs of attention deficit-hyperactivity disorder (ADHD) in the US: excess costs of persons with ADHD and their family members in 2000. Curr Med Res Opin. 2005;21:195–206.

Bloom, B. S., Bruno, D. J., Maman, D. Y., & Jayadevappa, R. (2001). Usefulness of US cost-of-illness studies in healthcare decision making. Pharmacoeconomics, 19(2), 207-213.

Blundell, R., & Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. Journal of Human Resources, 44(3), 565-640.

Bonomo YA & Bowes G (2001), Putting harm reduction into an adolescent context. Journal of Paediatrics and Child Health, 37(1), 5-8.

Borusyak K, Jaravel X. (2017). Revisiting Event Study Designs. Working Paper. Available at: https://scholar.harvard.edu/borusyak/publications/revisiting-event-study-designs

Buckley, P. F., & Brown, E. S. (2006). Prevalence and consequences of dual diagnosis. The Journal of clinical psychiatry, 67(7), e01-e01.

Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. Journal of econometrics, 46(3), 347-364.

Cameron, A. C., & Trivedi, P. K. (2001). Essentials of count data regression. A companion to theoretical econometrics, 331.

Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (Vol. 53). Cambridge university press.

Campbell SM, Reeves D, Kontopantelis E, et al. (2007), Quality for primary care in England with the introduction of pay for performance, New England Journal of Medicine, 375:181-90.

Clabaugh, G., & Ward, M. (2008). Cost-of-Illness Studies in the United States: A Systematic Review of Methodologies Used for Direct Cost. Value in Health, 11(1), 13-21.

Cleary, S. M., Birch, S., Moshabela, M., & Schneider, H. (2012). Unequal access to ART: exploratory results from rural and urban case studies of ART use. Sex Transm Infect, 88(2), 141-146.

Commons, M., McGuire, T. G., & Riordan, M. H. (1997). Performance contracting for substance abuse treatment. Health services research, 32(5), 631.

Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. Archives of general psychiatry, 64(5), 566-576.

Crowe, A. V., Howse, M., Bell, G. M., & Henry, J. A. (2000). Substance abuse and the kidney. Qjm, 93(3), 147-152.

Culyer, A. J., & Wagstaff, A. (1993). Equity and equality in health and health care. Journal of health economics, 12(4), 431-457.

Darke S, Mills KL, Ross J, Teesson M (2011). Rates and correlates of mortality amongst heroin users: findings from the Australian Treatment Outcome Study (ATOS), 2001–2009. Drug and alcohol dependence, 115(3), 190-195.

Davies, L., Jones, A., Vamvakas, G., Dubourg, R. & Donmall, M. (2009). The Drug Treatment Outcomes Research Study (DTORS): Cost effectiveness analysis London: Home Office.

Day, E. (2018). Commentary on Jones et al. (2018): An inconvenient truth—complex problems require complex solutions. Addiction, 113(2), 287-288.

Deaton, A (1997). The Analysis of Household Surveys: a Microeconometric Approach to Development Policy. World Bank Publications.

Deaton, A., & Paxson, C. (2000). Growth and saving among individuals and households. Review of Economics and Statistics, 82(2), 212-225.

Degenhardt, L., Lynskey, M., & HALLY, W. (2000). Cohort trends in the age of initiation of drug use in Australia. Australian and New Zealand Journal of Public Health, 24(4), 421-426.

DH (2010), Invitation to Tender for PbRDR, London: DH.

DH (2012), A simple guide to Payment by Results. London: Department of Health. Available at: https://www.gov.uk/government/uploads/system/uploads/attachmen t_data/file/21315 0/PbR-Simple-Guide-FINAL.pdf

DH (2017), Clinical Guidelines on Drug Misuse and Dependence Update 2017 Independent Expert Working Group (2017). Drug misuse and dependence: UK guidelines on clinical management. London: Department of Health. Available at: https://as sets.publishing.service.gov.uk/government/uploads/system/uploads/attachm ent_d ata/file/673978/clinical_guidelines_2017.pdf

Disley, E., Donmall, M., Jones, A., Mason, T., [. . .] & Turnbull, P. (2017). Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme. Final Report. Available at http://research.bmh.manchester.ac.uk/epidemiology/NDEC/r esearch/publications/Pb RDR_Final_Report.pdf

Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. American Economic Review, 108(2), 308-52.

Doran T, Campbell S, Fullwood C, et al. (2010), Performance of small general practices under the UK's Quality and Outcomes Framework, British Journal of General Practice, 60(578):e335-44.

Doran, T., Kontopantelis, E., Valderas, J. M., Campbell, S., Roland, M., Salisbury, C., & Reeves, D. (2011). Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. Bmj, 342, d3590.

Drummond M. Cost-of-illness studies: a major headache? Pharmacoeconomics. 1992;2:1–4.

Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: A systematic review of systematic reviews. Health policy 2013;110(2):115-130.

Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. Journal of health economics, 17(5), 537-555.

EMCDDA (2010). Treatment and care for older drug users. EMCDDA: Spain.

Finkelstein EA, Fiebelkorn IC, Wang G. National medical spending attributable to overweight and obesity: how much, and who's paying? Health Aff (Millwood) 2003;Suppl Web Exclusives:W3-219--W3-226.

Fligiel, S. E., Roth, M. D., Kleerup, E. C., Barsky, S. H., Simmons, M. S., & Tashkin, D. P. (1997). Tracheobronchial histopathology in habitual smokers of cocaine, marijuana, and/or tobacco. Chest, 112(2), 319-326.

Fergusson DM, Boden JM & Horwood, LJ (2015). Psychosocial sequelae of cannabis use and implications for policy: findings from the Christchurch Health and Development Study. Social psychiatry and psychiatric epidemiology, 50(9), 1317-1326.

Flodgren G, Eccles M, Shepperd S, Scott A, Parmelli E & Beyer F (2011), An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes, Cochrane Database Systematic Review, 7. French, M. T., & Drummond, M. (2005). A research agenda for economic evaluation of substance abuse services. Journal of Substance Abuse Treatment, 29(2), 125-137.

French, M. T., Fang, H., & Balsa, A. I. (2011). Longitudinal analysis of changes in illicit drug use and health services utilization. Health services research, 46(3), 877-899.

French, M. T., Mauskopf, J. A., Teague, J. L., & Roland, E. J. (1996). Estimating the dollar value of health outcomes from drug-abuse interventions. Medical Care, 890-910.

French, M. T., Roebuck, M. C., McLellan, A. T., & Sindelar, J. L. (2000). Can the Treatment Services Review be used to estimate the costs of addiction and ancillary services?. Journal of Substance Abuse, 12(4), 341-361.

Freyaldenhoven, S., Hansen, C., & Shapiro, J. M. (2018). Pre-event trends in the panel event-study design (No. w24565). National Bureau of Economic Research.

Galea, S., & Vlahov, D. (2002). Social determinants and the health of drug users: socioeconomic status, homelessness, and incarceration. Public health reports, 117(Suppl 1), S135.

García-Altés, A., Ollé, J. M., Antoñanzas, F., & Colom, J. (2002). The social cost of illegal drug consumption in Spain. Addiction, 97(9), 1145-1153.

Gfroerer, J., Penne, M., Pemberton, M., & Folsom, R. (2003). Substance abuse treatment need among older adults in 2020: the impact of the aging baby-boom cohort. Drug and alcohol dependence, 69(2), 127-135.

Giordano, G. N., Ohlsson, H., Kendler, K. S., Winkleby, M. A., Sundquist, K., & Sundquist, J. (2014). Age, period and cohort trends in drug abuse hospitalizations within the total Swedish population (1975–2010). Drug and alcohol dependence, 134, 355-361.

Godfrey, C., Stewart, D., & Gossop, M. (2004). Economic analysis of costs and consequences of the treatment of drug misuse: 2-year outcome data from the National Treatment Outcome Research Study (NTORS). Addiction, 99(6), 697-707.

Gossop, M., & Moos, R. (2008). Substance misuse among older adults: a neglected but treatable problem. Addiction, 103(3), 347-348.

Greene, W. H. (1997). Econometric Analysis. US: Prentice Hall.

Greene, W. H. (2003). Econometric analysis. Pearson Education India.

Haapanen-Niemi, N., Miilunpalo, S., Vuori, I., Pasanen, M., & Oja, P. (1999). The impact of smoking, alcohol consumption, and physical activity on use of hospital services. American Journal of Public Health, 89(5), 691-698.

Haber, P. S., Demirkol, A., Lange, K., & Murnion, B. (2009). Management of injecting drug users admitted to hospital. The Lancet, 374(9697), 1284-1293.

Hamid H, El-Mallakh RS, Vandeveir K (2005). Substance abuse: medical and slang terminology. Southern medical journal, 98(3), 350-362.

Harwood, H. J., Fountain, D., & Livermore, G. (1999). Cost estimates for alcohol and drug abuse. Addiction, 94(5), 631-647.

Hashibe, M., Morgenstern, H., Cui, Y., Tashkin, D. P., Zhang, Z. F., Cozen, W., ... & Greenland, S. (2006). Marijuana use and the risk of lung and upper aerodigestive

tract cancers: results of a population-based case-control study. Cancer Epidemiology and Prevention Biomarkers, 15(10), 1829-1834.

Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Archives of general psychiatry, 64(7), 830-842.

Healey, A., Knapp, M., Astin, J., Gossop, M., Marsden, J., Stewart, D., ... & Godfrey, C. (1998). Economic burden of drug dependency: social costs incurred by drug users at intake to the national treatment outcome research study. The British journal of psychiatry, 173(2), 160-165.

Healey, A., Knapp, M., Marsden, J., Gossop, M., & Stewart, D. (2003). Criminal outcomes and costs of treatment services for injecting and non-injecting heroin users: evidence from a national prospective cohort survey. Journal of health services research & policy, 8(3), 134-141.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. The review of economic studies, 64(4), 605-654.

Heß, S. 2017. Randomization inference with Stata: A guide and software. Stata Journal 17: 630–651

Higate, P., Hughes, R., & Lart, R. (2006). Drugs: Policy And Politics: Policy and Politics. McGraw-Hill Education (UK).

HMG (2010). Drug Strategy 2010. Reducing Demand, Restricting Supply, Building Recovery: Supporting People to Live a Drug Free Life. HMG: London. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachm ent_data/file/98026/drug-strategy-2010.pdf

HMG (2010a). Open Public Services. White Paper. HMG: London. Available at: http://files.openpublicservices.cabinetoffice.gov.uk/OpenPublicServices-WhitePaper.pdf

HMG (2012). The Health and Social Care Act. HMG: London. Available at:. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/138257/A1.-Factsheet-Overview-240412.pdf

HMG (2013). Piloting Payment by Results for drug recovery: draft outcome definitions. HMG: London Available at: https://www.gov.uk/government/publications/piloting-payment-by-results-for-drugs-recovery-draft-outcome-definitions in May 2013.

HMG (2017). 2017 Drug Strategy. HMG: London Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachm ent_data/file/628148/Drug_strategy_2017.PDF

Hodgson, T. A., & Cohen, A. J. (1999). Medical care expenditures for diabetes, its chronic complications, and its comorbidities. Preventive Medicine, 29(3), 173-186.

Hoff, R. A., & Rosenheck, R. A. (1998). Long-term patterns of service use and cost among patients with both psychiatric and substance abuse disorders. Medical Care, 835-843.

Hogan, P., Dall, T., & Nikolov, P. (2003). Economic costs of diabetes in the US in 2002. Diabetes care, 26(3), 917.

Hollander, J. E. (1995). The management of cocaine-associated myocardial ischemia. New England Journal of Medicine, 333(19), 1267-1272.

Home Office (2013), Understanding organised crime: estimating the scale and the social and economic costs. Research Report 73. Home Office: London. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/24639 0/horr73.pdf

Home Office (2018), Drug Misuse: Findings from the 2017/18 Crime Survey for England and Wales. Statistical Bulletin 14/18. July 2018. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachm ent_data/file/729249/drug-misuse-2018-hosb1418.pdf

House of Commons (2017), Human and financial costs of drug addiction. Debate Pack Number CDP-0230. House of Commons Library: London. Available at: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CDP-2017-0230#fullreport

HSCIC (2016), Statistics on drugs misuse (England 2016). Available at: http://digital.nhs.uk/catalogue/PUB21159

IDPC (2018). Taking stock: a decade of drug policy. A civil society shadow report. Available at: http://fileserver.idpc.net/library/Shadow_Report_FINAL_ENGLISH.pdf

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. Journal of economic literature, 47(1), 5-86.

Jacobs, R., Chalkley, M., Aragón, M. J., Böhnke, J. R., Clark, M., & Moran, V. (2018). Funding approaches for mental health services: is there still a role for clustering? BJPsych advances, 24(6), 412-421.

Jo, C. (2014). Cost-of-illness studies: concepts, scopes, and methods. Clinical and molecular hepatology, 20(4), 327.

John, U., Rumpf, H. J., & Hapke, U. (1999). Estimating prevalence of alcohol abuse and dependence in one general hospital: an approach to reduce sample selection bias. Alcohol and Alcoholism, 34(5), 786-794.

Jones, A. M., & Rice, N. (2011). Econometric evaluation of health policies. The Oxford handbook of health economics. Oxford University Press, Oxford.

Jones, A. M., Lomas, J., & Rice, N. (2015). Healthcare cost regressions: going beyond the mean to estimate the full distribution. Health economics, 24(9), 1192-1212.

Jones, A., Donmall, M., Millar, T., Moody, A., Weston, S., Anderson, T., ... & D'Souza, J. (2009). The Drug Treatment Outcomes Research Study (DTORS): Final outcomes report. 3rd Edition. Research Report 24: Key Implications. Available at: https://s3.amazonaws.com/academia.edu.documents/46467983/horr24c.pdf?AWSAc cessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1542204464&Signature=vD MRl8NpbSxouvIpgVzaH99pUA0%3D&response-content-disposition=inline%3B%20filename%3DThe_Drug_Treatment_Outcomes_Research _Stu.pdf

Jones, A., Pierce, M., Sutton, M., Mason, T., & Millar, T. (2017). Does paying service providers by results improve recovery outcomes for drug misusers in treatment in England?. Addiction, 113(2), 279-286.

Kelly E, Stoye, G & Vera-Hernandez M. (2015), Public Hospital Spending in England: Evidence from National Health Service Administrative Records. IFS Working Paper W15/21 available at: https://www.ifs.org.uk/uploads/publications/wps/WP201521.pdf

Kennelly, B. (2017). How should cost-of-illness studies be interpreted?. The Lancet Psychiatry, 4(10), 735-736.

Kontopantelis, E., Springate, D., Reeves, D., Ashcroft, D. M., Valderas, J. M., & Doran, T. (2014). Withdrawing performance indicators: retrospective analysis of

general practice performance under UK Quality and Outcomes Framework. Bmj, 348, g330.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., & Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. Health economics, 25(12), 1514-1528.

Lancaster, T. (2000). The incidental parameter problem since 1948. Journal of econometrics, 95(2), 391-413.

Lange, DE & Schacter, B (1989). Prevalence of alcohol related admissions to general medical units. International Journal of Psychiatry, 19: 371-384.

Langendam MW, Van Brussel GH, Coutinho RA & Van Ameijden EJ (2001), The impact of harm-reduction-based methadone treatment on mortality among heroin users. American journal of public health, 91(5), 774.

Leslie KM (2008), Harm reduction: An approach to reducing risky health behaviours in adolescents. Paediatric Child Health, 13(1), 53-6.

Li W & Gunja N (2013). Illicit drug overdose: Prevalence and acute management. Australian family physician, 42(7), 481.

Li, X., Wong, W., Lamoureux, E.L. & Wong, T. Y. (2012). Are Linear Regression Techniques Appropriate for Analysis When the Dependent (Outcome) Variable Is Not Normally Distibuted? Investigative Ophthalmology & Visual Science, 53(6): 3082-3.

Li, X., Wong, W., Lamoureux, E.L. & Wong, T. Y. (2012). Are Linear Regression Techniques Appropriate for Analysis When the Dependent (Outcome) Variable Is Not Normally Distibuted? Investigative Ophthalmology & Visual Science, 53(6): 3082-3.

Liu, J. L. Y., Maniadakis, N., Gray, A., & Rayner, M. (2002). The economic burden of coronary heart disease in the UK. Heart, 88(6), 597-603.

Long, J.S., Freese, J. (2006) Regression Models for Categorical Dependent Variables Using Stata. 2nd ed. College Station, TX: Stata Press

Luce BR, Manning WG, Siegel JE, Lipscomb J. Estimating costs in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, editors. Cost-

effectiveness in health and medicine. New York: Oxford University Press; 1996. pp. 173–213.

MacGregor, S., & Thickett, A. (2011). Partnerships and communities in English drug policy: the challenge of deprivation. International Journal of Drug Policy, 22(6), 478-490.

Macleod J, Oakes R, Copello A, Crome I, Egger M, Hickman M, ... & Smith GD (2004). Psychological and social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies. The Lancet, 363(9421), 1579-1588.

Malehi, A. S., Pourmotahari, F., & Angali, K. A. (2015). Statistical models for the analysis of skewed healthcare cost data: a simulation study. Health economics review, 5(1), 11.

Manning, W. G., & Mullahy, J. (2001). Estimating log models: to transform or not to transform?. Journal of health economics, 20(4), 461-494.

March, J. C., Oviedo-Joekes, E., & Romero, M. (2006). Drugs and social exclusion in ten European cities. European addiction research, 12(1), 33-41.

Marik, P., & Mohedin, B. (1996). Alcohol-related admissions to an inner city hospital intensive care unit. Alcohol and alcoholism, 31(4), 393-396.

Mason, T., Lau, Y. S., & Sutton, M. (2016). Is the distribution of care quality provided under pay-for-performance equitable? Evidence from the Advancing Quality programme in England. International journal for equity in health, 15(1), 156.

Mason, T., Sutton, M., Whittaker, W., & Birch, S. (2015a). Exploring the limitations of age-based models for health care planning. Social Science & Medicine, 132, 11-19.

Mason, T., Sutton, M., Whittaker, W., McSweeney, T., Millar, T., Donmall, M., ... & Pierce, M. (2015b). The impact of paying treatment providers for outcomes: di fference-in-differences analysis of the 'payment by results for drugs recovery' pilot. Addiction, 110(7), 1120-1128.

Mattick, R. P., Breen, C., Kimber, J., & Davoli, M. (2009). Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. Cochrane database of systematic reviews, (3).

Mattick, R. P., Breen, C., Kimber, J., & Davoli, M. (2014). Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. Cochrane database of systematic reviews, (2).

Maynard, A., Street, A., & Hunter, R. (2011). Using 'payment by results' to fund the treatment of dependent drug users—proceed with care!. Addiction, 106(10), 1725-1729.

McCollister, K. E., French, M. T., Prendergast, M., Wexler, H., Sacks, S., & Hall, E. (2003). Is in-prison treatment enough? A cost-effectiveness analysis of prison-based treatment and aftercare services for substance-abusing offenders. Law & Policy, 25(1), 63-82.

McLellan, A. T., Kemp, J., Brooks, A., & Carise, D. (2008). Improving public addiction treatment through performance contracting: The Delaware experiment. Health policy, 87(3), 296-308.

Mehrotra A, Damberg CL, Sorbero, ME & Teleki SS (2009), Pay for performance in the hospital setting: what is the state of the evidence? American Journal of Medical Quality, 24(1), 19-28. Mendelson, A., Kondo, K., Damberg, C., Low, A., Motúapuaka, M., Freeman, M., ... & Kansagara, D. (2017). The effects of pay-for-performance programs on health, health care use, and processes of care: a systematic review. Annals of Internal Medicine, 166(5), 341-353. Mihaylova, B., Briggs, A., O'hagan, A., & Thompson, S. G. (2011). Review of statistical methods for analysing healthcare resources and costs. Health economics, 20(8), 897-916.

Monitor (2014), The maternity pathway payment system: Supplementary guidance. London: Monitor. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/34252 2/maternity_payment_pathway_system_supplementary_guidance.pdf

Morgenstern H, Kleinbaum DG, Kupper LL. (1980) Department of Epidemiology and Public Health, School of Medicine, Yale University, New Haven, CT. Int J Epidemiol. 9:97–104.

Mundlak, Y. (1978). On the pooling of time series and cross section data. Econometrica: journal of the Econometric Society, 69-85.

Neighbors, C. J., Zywiak, W. H., Stout, R. L., & Hoffmann, N. G. (2005). Psychobehavioral risk factors, substance treatment engagement and clinical outcomes

as predictors of emergency department use and medical hospitalization. Journal of studies on alcohol, 66(2), 295-304.

Newhouse, J. P. (1984). Cream skimming, asymmetric information, and a competitive insurance market. Journal of Health Economics, 3(1), 97-100.

Newhouse, J. P. (1994). Patients at risk: health reform and risk adjustment. Health Affairs, 13(1), 132-146.

NHS Digital (2016), HES Data Dictionary: Admitted Patient Care. Available at: ht tp://content.digital.nhs.uk/media/19425/APC-DD-Final-Doc2/pdf/DD-APC-V7.p df

NHS Digital (2017), Statistics on drugs misuse (England 2017). Available at: https: //digital.nhs.uk/catalogue/PUB23442

NHS Digital (2018). Background data quality report. Statistics on drugs misuse. Available at: https://files.digital.nhs.uk/publication/a/t/drug-misu-eng-2018-qual.p df

NHS Digital (2018). Statistics on drugs misuse (England 2018). Available at: https: //files.digital.nhs.uk/publication/c/k/drug-misu-eng-2018-rep.pdf

NHS England (2016). Refreshing the Formulae for CCG Allocations for allocations to Clinical Commissioning Groups from 2016-17 - Report on the methods and modelling. Leeds: NHS England. Available at: https://www.england.nhs.uk/wp-content/uploa ds/2016/04/3-rep-elland-all-sections.pdf

NHS Improvement (2018). Health costing standards for England. Acute: Costing methods. Publication code: CG 14/19. London: NHS Improvement. Available at: h ttps://improvement.nhs.uk/documents/4886/Acute_CMs_standards20.pdf

NICE (2007). Methadone and buprenorphine for the management of opioid dependence. Technology appraisal guidance. Available at: https://www.nice.org.uk /guidance/ta114/resources/methadone-and-buprenorphine-for-the-management-of-o pioid-dependence-pdf-82598072878789

NICE (2008). Drug Misuse: Psychosocial Interventions. Leicester (UK): British Psychological Society; 2008. (NICE Clinical Guidelines, No. 51.) 3, INTRODUCTION TO DRUG MISUSE. Available from: https://www.ncbi.nlm.nih.gov/books/NBK53 217/

NTA (2010). Commissioning for recovery. Drug treatment, reintegration and recovery in the community and prisons: a guide for drug partnerships. London: NTA.

NTA (2012). Medications in Recovery: reorientating drug dependence treatment. Available at: http://www.nta.nhs.uk/uploads/medications-in-recovery-mainreport3.pdf

NTA (2014), Why invest? How drug treatment and recovery services work for individuals, communities and society. Available at: https://www.scie-socialcareonline.org.uk/why-invest-how-drug-treatment-and-recovery-services-work-for-individuals-communities-and-society/r/a11G00000017w0HIAQ

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. Health Services and Outcomes Research Methodology, 16(1-2), 1-21.

Oliver, A. (2005). The English National Health Service: 1979-2005. Health economics, 14(S1).

Olsson, T. M., & Fridell, M. (2015). Women with comorbid substance dependence and psychiatric disorders in Sweden: a longitudinal study of hospital care utilization and costs. BMC health services research, 15(1), 224.

ONDCP (2004). The economic costs of drug abuse in the United States. Washington: Executive Office of the President Office of National Drug Control Policy. Available at: http://www.ncjrs.gov/ondcppubs/publications/pdf/economic_costs.pdf

ONS (2017). Deaths related to drug poisoning in England and Wales: 2016 registrations. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/d eaths/bulletins/deathsrelatedtodrugpoisoninginenglandandwales/2016registrations

ONS (2018). Deaths related to drug poisoning in England and Wales: 2017 registrations. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/d eaths/bulletins/deathsrelatedtodrugpoisoninginenglandandwales/2017registrations

Orford, J., Somers, M., Daniels, V. & Kirby, B. (1992). Drinking amongst medical patients: levels of risk and models of changes. British Journal of Addiction; 87: 1691-1702.

Palepu, A., Strathdee, S. A., Hogg, R. S., Anis, A. H., Rae, S., Cornelisse, P. G., ... & Schechter, M. T. (1999). The social determinants of emergency department and hospital use by injection drug users in Canada. Journal of Urban Health, 76(4), 409-418.

Parthasarathy, S., & Weisner, C. M. (2005). Five-year trajectories of health care utilization and cost in a drug and alcohol treatment sample. Drug and Alcohol Dependence, 80(2), 231-240.

PHE (2014). About the data: how the data is organised. Accessed at http://fingertips.phe.org.uk/topic/drugs-and-alcohol/about-data

PHE (2014a). NDTMS ViewIt. Accessed at: https://www.ndtms.net/viewit/Adult/ExecutiveSummary.aspx

PHE (2016) Adult substance misuse statistics from the National Drug Treatment Monitoring System (NDTMS). 1 April 2015 to 31 March 2016. Public Health England. Available at: http://www.nta.nhs. uk/uploads/adult-statistics-from-the-nationaldrug-treatment-monitoring-system-2015-2016[0].pdf

Pierce, M., Bird, S. M., Hickman, M., & Millar, T. (2015). National record linkage study of mortality for a large cohort of opioid users ascertained by drug treatment or criminal justice sources in England, 2005–2009. Drug and alcohol dependence, 146, 17-23.

Pierce, M., Bird, S. M., Hickman, M., Marsden, J., Dunn, G., Jones, A., & Millar, T. (2016). Impact of treatment for opioid dependence on fatal drug-related poisoning: a national cohort study in England. Addiction, 111(2), 298-308.

Piette, J. D., Barnett, P. G., & Moos, R. H. (1998). First-time admissions with alcohol-related medical problems: a 10-year follow-up of a national sample of alcoholic patients. Journal of studies on alcohol, 59(1), 89-96.

Poikolainen, K. (1982). Seasonality of alcohol-related hospital admissions has implications for prevention. Drug & Alcohol Dependence, 10(1), 65-69.

PSSRU (2011), Unit Costs of Health and Social Care 2011. PSSRU: Kent. Available at: http://www.pssru.ac.uk/pub/uc/uc2011/uc2011.pdf

Rao, R., & Roche, A. (2017). Substance misuse in older people.

Reece, A. S. (2007). Dentition of addiction in Queensland: poor dental status and major contributing drugs. Australian dental journal, 52(2), 144-149.

Reynaud, M., Gaudin-Colombel, A. F., & Le Pen, C. (2001). Two methods of estimating health costs linked to alcoholism in France (with a note on social costs). Alcohol and alcoholism, 36(1), 89-95.

Rice, D. P. (1967). Estimating the cost of illness. American Journal of Public Health and the Nations Health, 57(3), 424-440.

Rippeth, J. (2008). A Review of: Addicted to a challenge: An update on the neuropsychological effects of substance abuse disorders: Neuropsychology and substance use: State-of-the-art and future directions. Ari Kalechstein & Wilford G. van Gorp (Eds.).[Series: Studies on Neuropsychology, Neurology, and Cognition, Series Ed. LA Bieliauskas]. 2007. New York, NY: Psychology Press.

Rockhill B, Newman B, Weinberg C. (1998) Use and misuse of population attributable fractions. Am J Public Health. 88:15–19.

Roe S, Man L (2006). Home Office Statistical Bulletin. Drug misuse declared: findings from the 2005/06 British Crime Survey. http://www.homeoffice.gov.uk/rds/pdfs06/hosb1506.pdf.

Roxburgh, A., Degenhardt, L., & Breen, C. (2005). Drug use and risk behaviours among injecting drug users: a comparison between sex workers and non-sex workers in Sydney, Australia. Harm Reduction Journal, 2(1), 7.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. Health Services and Outcomes Research Methodology, 2(3-4), 169-188.

Ryan, A. M., & Doran, T. (2012). The effect of improving processes of care on patient outcomes: evidence from the United Kingdom's quality and outcomes framework. Medical care, 191-199.

Ryan, A. M., Burgess, J., & Dimick, J. B. (2014). Why we shouldn't be indifferent to specification in difference-in-differences analysis. Health services research.

Santora, P. B., & Hutton, H. E. (2008). Longitudinal trends in hospital admissions with co-occurring alcohol/drug diagnoses, 1994–2002. Journal of substance abuse treatment, 35(1), 1-12.

Schoenbaum, M., Zhang, W., & Sturm, R. (1998). Costs and utilization of substance abuse care in a privately insured population under managed care. Psychiatric Services, 49(12), 1573-1578.

Scott, A., Liu, M., & Yong, J. (2018). Financial incentives to encourage value-based health care. Medical Care Research and Review, 75(1), 3-32.

Scott, A., Sivey, P., Ouakrim, D. A., Willenberg, L., Naccarella, L., Furler, J., & Young, D. (2011). The effect of financial incentives on the quality of health care provided by primary care physicians. Cochrane database of systematic reviews, (9).

Shaffer, D. K., Hartman, J. L., & Listwan, S. J. (2009). Drug abusing women in the community: The impact of drug court involvement on recidivism. Journal of Drug Issues, 39(4), 803-827.

Shah, A. D., Wood, D. M., & Dargan, P. I. (2011). Survey of ICD-10 coding of hospital admissions in the UK due to recreational drug toxicity. QJM: An International Journal of Medicine, 104(9), 779-784.

Shen, Y., & Ellis, R. P. (2002). How profitable is risk selection? A comparison of four risk adjustment models. Health economics, 11(2), 165-174.

Simoni-Wastila, L., & Yang, H. K. (2006). Psychoactive drug abuse in older adults. The American journal of geriatric pharmacotherapy, 4(4), 380-394.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. International journal of public administration, 18(2-3), 277-310.

Solomon, L., Astemborski, J., Warren, D., Muñoz, A., Cohn, S., Vlahov, D., & Nelson, K. E. (1993). Differences in risk factors for human immunodeficiency virus type 1 seroconversion among male and female intravenous drug users. American Journal of Epidemiology, 137(8), 892-898.

Stein, M. D. (1994). Injected-drug use: complications and costs in the care of hospitalized HIV-infected patients. Journal of acquired immune deficiency syndromes, 7(5), 469-473.

Stein, M. D., & Sobota, M. (2001). Injection drug users: hospital care and charges. Drug and alcohol dependence, 64(1), 117-120.

Stein, M. D., O'Sullivan, P. S., Ellis, P., Perrin, H., & Wartenberg, A. (1993). Utilization of medical services by drug abusers in detoxification. Journal of Substance

Abuse, 5(2), 187-193.

Steventon, A., Tunkel, S., Blunt, I., & Bardsley, M. (2013). Effect of telephone health coaching (Birmingham OwnHealth) on hospital use and associated costs: cohort study with matched controls. Bmj, 347, f4585.

Street A. (2006), Reforma de la financición hospitalaria e introducción de mecanismos de elección en el sistema sanitario británico NHS. Revista Española de Economía de la Salud, 5(2) 97-102.

Sutton, M., Garfield-Birkbeck, S., Martin, G., Meacock, R., Morris, S., Sculpher, M., ... & Lilford, R. J. (2018). Economic analysis of service and delivery interventions in health care.

Sutton, M., Nikolova, S., Boaden, R., Lester, H., McDonald, R., & Roland, M. (2012). Reduced mortality with hospital pay for performance in England. New England Journal of Medicine, 367(19), 1821-1828.

Swensen AR, Birnbaum HG, Secnik K, Marynchenko M, Greenberg P, Claxton A. At tention-deficit/hyperactivity disorder: increased costs for patients and their families. J Am Acad Child Adolesc Psychiatry. 2003;42:1415–1423.

Tarricone, R. (2006). Cost-of-illness analysis: what room in health economics?. Health policy, 77(1), 51-63.

Taylor, C., Passmore, N., Kilbane, P., & Davies, R. (1986). Prospective study of alcohol-related admissions in an inner-city hospital. The Lancet, 328(8501), 265-268

United Nations Office on Drugs and Crime, World Drug Report 2017 (ISBN: 978-92-1-148291-1, eISBN: 978-92-1-060623-3, United Nations publication, Sales No. E.17.XI.6).

UNODC (2013). HIV Prevention, Treatment and Care in prisons and Other Closed Settings: A Comprehensive Package of Interventions. Vienna, Austria: United Nations Office on Drugs and Crime. Available at: https://www.unodc.org/documents/hiv-aids/HIV_comprehensive_package_prison_2013_eBook.pdf

UNODC (2017) World Drug Report. UN publication. Available at: https://www.unodc.org/wdr2017/field/Booklet_1_EXSUM.pdf

Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., &

Sermeus, W. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. BMC health services research, 10(1), 247.

Varmus H (2000). Disease-specific estimates of direct and indirect costs of illness and NIH support. Washington, DC: National Institutes of Health, November 1995, April 1997, February 2000.

Vasica G & Tennant CC (2002). Cocaine use and cardiovascular complications. Medical Journal of Australia, 177(5), 260-262.Cornford C & Close H (2016). The physical health of people who inject drugs: complexities, challenges, and continuity. British Journal of General Practice 2016; 66 (647): 286-287.

Veazie, P. J., Manning, W. G., & Kane, R. L. (2003). Improving risk adjustment for Medicare capitated reimbursement using nonlinear models. Medical Care, 741-752.

Whittaker, W., Anselmi, L., Kristensen, S. R., Lau, Y. S., Bailey, S., Bower, P., ... & Hodgson, D. (2016). Associations between extending access to primary care and emergency department visits: a difference-in-differences analysis. PLoS medicine, 13(9), e1002113.

WHO (2006). Lexicon of Alcohol and Drug Terms Published by the World Health Organization. Available at: http://www.who.int/substance_abuse/terminology /who_lexicon/en/

WHO (2009). Technical Guide for countries to set targets for universal access to HIV prevention, treatment and care for injecting drug users. Available at: https://www. who.int/hiv/pub/idu/targets_universal_access/en/

WHO (2016). International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)-WHO Version for 2016. Available at https://icd. who.int/browse10/2016/en#/

Wickizer, T. M. (2013). State-level estimates of the economic costs of alcohol and drug abuse. Journal of health care finance, 39(3), 71-84.

Wiley, L. K., Shah, A., Xu, H., & Bush, W. S. (2013). ICD-9 tobacco use codes are effective identifiers of smoking status. Journal of the American Medical Informatics Association, 20(4), 652-658.

Wooldridge (2009) Introductory Econometrics: A modern approach. Fourth Edition. Canada: South-Western

Wooldridge, J M (2010). Correlated random effects models with unbalanced panels. Working Paper. Available at: http://econ.msu.edu/faculty/wooldridge/docs/cre1_r4.pdf

Wooldridge, J. M. (2012). Introductory Econometrics: A Modern Approach, (Boston: Cengage Learning).

Zhang, L. R., Morgenstern, H., Greenland, S., Chang, S. C., Lazarus, P., Teare, M. D., ... & Brhane, Y. (2015). Cannabis smoking and lung cancer risk: Pooled analysis in the International Lung Cancer Consortium. International journal of cancer, 136(4), 894-903.

Zhang, M., Fortney, J. C., Tilford, J. M., & Rost, K. M. (2000). An application of the inverse hyperbolic sine transformation—a note. Health Services and Outcomes Research Methodology, 1(2), 165

9 Appendix

Figure A4a: Coefficient plots comparing event time estimates across models (annual cost of hospital admissions)

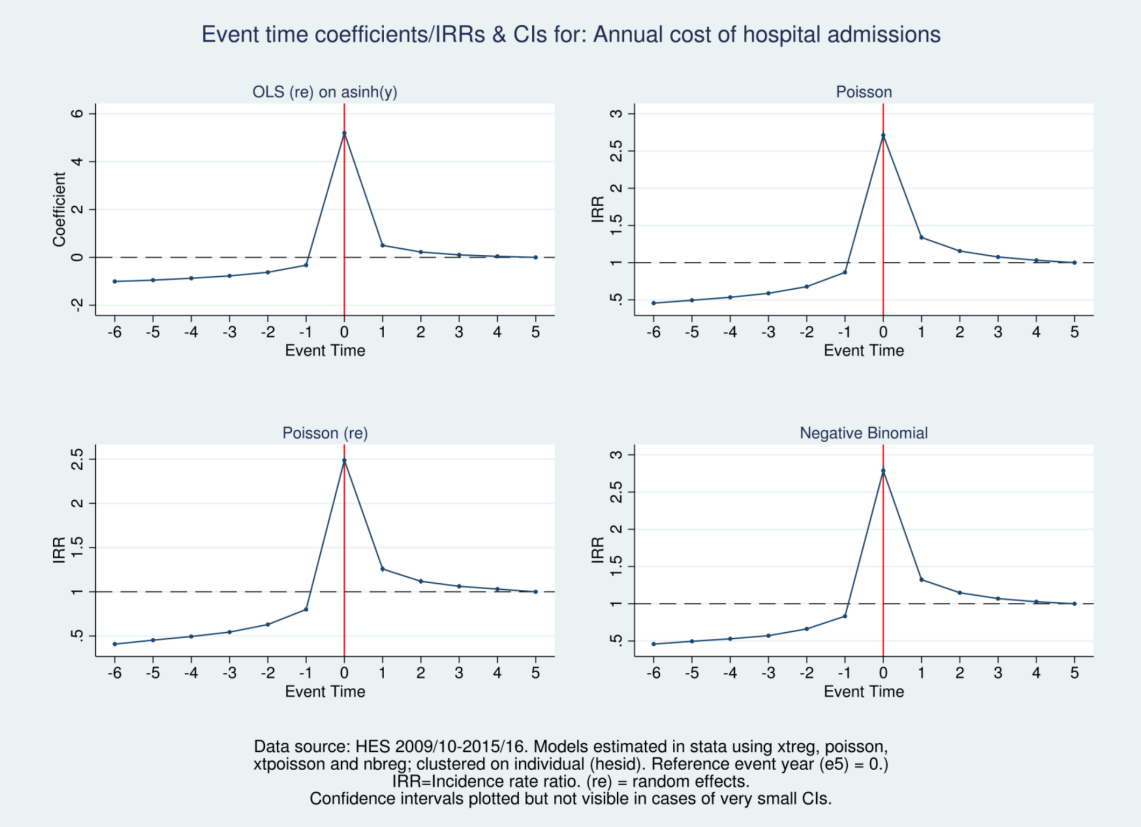


Table A1.1: Domain 1 in the National Outcomes Framework (part 1)

	Outcome domain 1: Free from drug(s) of dependence	Eligibility Criteria
1.1	Drug and/or alcohol use significantly improved (i.e. reduced their consumption by statistically significant levels) for all presenting substances at any two Treatment Outcomes Profile (TOP) review (or one review and one treatment exit TOP where applicable) within the last 12 months	Clients 18 or over who cited as a problematic substance any of opiates, crack cocaine, alcohol, cannabis or amphetamines in their latest treatment journey
	No more than 20% of the total value assigned to this outcome domain can be placed on this initial outcome. Where a client exits treatment before 6 months then the appropriate exit TOP will be sufficient without any prior review TOP	Where an individual presents with only substances not recorded on TOP (currently 1%) then a change of one standard deviation will be used for the information recorded under 'other substance'. Only one substance will be able to be measured in this way.
1.2	Abstinent from all presenting substances at any two TOP reviews (or one review and one treatment exit TOP where applicable) where this was in the last 12 months. Where a client exits treatment before 6 months then the appropriate exit TOP will be sufficient without a prior review TOP	Clients 18 or over who cited as a problematic substance any of opiates, crack cocaine, alcohol, cannabis or amphetamines in their latest treatment journey - still in structured treatment at the end of the specified payment period. Same treatment of non-TOP substances as above.

Table A1.2: Domain 1 in the National Outcomes Framework (part 2)

	Outcome domain 1: Free from drug(s) of dependence	Eligibility Criteria
1.3	Planned exit from the treatment component of the recovery journey, free from drug(s) - including alcohol - of dependence	Clients aged 18 or over, not in structured treatment in past 21 days. Assessed as having a dependency on drugs/alcohol which would benefit from structured treatment
1.4	Discharged from treatment successfully (free from drugs(s) of dependence) and does not re-present in either the treatment system or in the criminal justice system (taken onto the DIP/prison caseload) in the following 12 months	An individual does not show up on NDTMS having started a new treatment journey, nor are they shown as being subsequently taken onto DIP/prison caseload in the 12 months following their successful completion of structured treatment

Table A1.3: Domain 2 in the National Outcomes Framework

	Outcome domain 2: Offending	Eligibility Criteria
2	Reduction in average offending of cohort compared against baseline, calculated and paid quarterly. Cohort to be made up of those individuals in treatment and discharged in the previous 6 months	All clients aged 18 or over at time of initial engagement with provider regardless of previous offending history

Table A1.4: Domain 3 in the National Outcomes Framework

	Outcome domain 3: Health and wellbeing outcomes	Eligibility Criteria
3.1	Of those injecting at the start of treatment, those who reported 0 days injecting on any two reviews TOPs (or one review and one treatment exit TOP where applicable) within the last 12 months	Clients 18 or over who indicated that they were currently injecting at the start of treatment (as either indicated at their start TOP, or had an injecting status of 'currently injecting', or a route of administration of 'injecting' in their latest treatment journey)
	Where a client exits treatment before 6 months then the appropriate exit TOP will be sufficient without a prior review TOP	Providers will also want to consider testing for other key blood borne viruses such as Hep C and HIV

Table A4.1: Model Comparison: full estimates (annual cost of admissions)

	OLS	OLS (re)	Poisson	Poisson (re)	Neg. Binomial
Event time					
(reference=+5)					
-6	-.939***	-1.005***	.456***	.408***	.458***
-5	-.899***	-.951***	.495***	.453***	.495***
-4	-.830***	-.871***	.534***	.494***	.529***
-3	-.738***	-.773***	.588***	.544***	.571***
-2	-.589***	-.624***	.677***	.629***	.662***
-1	-.287***	-.326***	.869***	.799***	.833***
Event=0	5.242***	5.203***	2.713***	2.489***	2.788***
1	.527***	.501***	1.339***	1.258***	1.323***
2	.234***	.224***	1.156***	1.119***	1.148***
3	.106***	.105***	1.076***	1.062***	1.069***
4	.041**	.044**	1.032**	1.031**	1.027*
Year					
2009-10	-.019*	-0.010	.913***	.908***	.896***
2010-11	.019**	.023***	.974***	.959***	.971**
2011-12	.031***	.032***	.967***	.968***	.932***
2012-13	-.054***	-.052***	.710***	.716***	.724***
2013-14	-.017*	-0.013	.821***	.828***	.821***
2014-15	-.052***	-.040***	.733***	.738***	.717***
Age in 0910	.017***	.018***	1.031***	1.025***	1.027***
(squared)	.001***	.001***	.999***	1.000	0.999
Male	-.435***	-.427***	.984*	.938***	.854***
IMD 1	-.475***	-.472***	.788***	.767***	.739***
IMD 2	-.262***	-.262***	.887***	.874***	.862***
IMD 3	-.129***	-.129***	.926***	.924***	.924***
IMD 4	-.058***	-.058***	0.972	.968*	.965*
IMD 6	.052***	.052***	1.023	1.031*	1.039*
IMD 7	.106***	.106***	1.043**	1.045***	1.043**
IMD 8	.109***	.109***	1.042**	1.047**	1.051**
IMD 9	.067***	.068***	1.008	1.003	1.007
IMD 10	-.128***	-.128***	.923***	.910***	.897***
Black	.079***	.080***	1.469***	1.498***	1.543***
Asian	-.133***	-.131***	1.093**	1.038	1.031
Other	-.049**	-.051**	1.088***	1.088***	1.071***
Mixed	.086***	.086***	1.187***	1.187***	1.167***
Constant	2.014***	2.013***	507.311***	627.114***	583.759***
lnalpha (dispersion parameter)				1.388***	16.684***
N			3,233,056		
R-sq	0.265	0.265			
AIC	17,244,316		$1.41*10^{10}$	$6.91*10^{9}$	29,628,507
BIC	17,244,758		$1.41*10^{10}$	$6.91*10^{9}$	29,628,961

*p<0.05; **p<0.01; ***p<0.001; OLS estimated on the asinh(y); Poisson and Negative Binomial show incident rate ratios.

Table A4.2: Model Comparison, full estimates (Total LOS per year)

	OLS	OLS (re)	Poisson	Poisson (re)	Neg. Binomial
-6	-.279***	-.323***	.356***	.326***	.348***
-5	-.257***	-.298***	.410***	.381***	.398***
-4	-.230***	-.270***	.476***	.440***	.466***
-3	-.205***	-.245***	.529***	.489***	.509***
-2	-.165***	-.203***	.617***	.571***	.591***
-1	-.072***	-.112***	.816***	.751***	.765***
Event=0	1.137***	1.099***	3.148***	2.893***	3.101***
1	.174***	.143***	1.532***	1.433***	1.494***
2	.075***	.055***	1.236***	1.189***	1.222***
3	.032***	.021***	1.141***	1.119***	1.126***
4	.011*	.008*	1.060**	1.060**	1.039
2009-10	0.004	.005*	0.998	0.995	1.020
2010-11	-.009***	-.009***	1.012	1.005	1.056***
2011-12	-.006**	-.006*	1.013	1.021*	1.037**
2012-13	.006**	.006**	.978**	.982*	0.991
2013-14	0.004	0.004	.965***	.974**	.977*
2014-15	-.014***	-.015***	.827***	.839***	.810***
Age in 0910	-.001**	-.001**	1.022***	1.019***	1.022***
(squared)	.001***	.001***	1.001***	1.000***	1.000***
Male	-.045***	-.043***	1.063***	1.061***	.980*
IMD 1	-.153***	-.153***	.743***	.721***	.682***
IMD 2	-.091***	-.091***	.830***	.827***	.815***
IMD 3	-.053***	-.054***	.879***	.877***	.868***
IMD 4	-.022***	-.023***	.968*	0.971	0.982
IMD 6	.015**	.015**	0.992	0.994	1.012
IMD 7	.029***	.029***	1.006	0.994	1.004
IMD 8	.036***	.037***	1.017	1.008	1.028
IMD 9	.028***	.029***	0.973	.954**	.959*
IMD 10	-.027***	-.025***	.831***	.817***	.794***
Black	.184***	.185***	2.182***	2.412***	2.448***
Asian	.015**	.015**	1.223***	1.284***	1.281***
Other	.045***	.043***	1.285***	1.344***	1.359***
Mixed	.092***	.092***	1.708***	1.770***	1.799***
Constant	.322***	.349***	1.493***	1.653***	1.543***
lnalpha (dispersion parameter)				2.892***	11.695***
N			3,233,056		
R-sq	0.156	0.156			
AIC	10,264,974		65,915,410	31,375,016	9,524,749.9
BIC	10,265,415		65,915,852	31,375,471	9,525,204.5

*p<0.05; **p<0.01; ***p<0.001; Poisson and Negative Binomial show incident rate ratios.

Table A4.3: Model Comparison, full estimates (Number of hospital admissions per year)

	OLS	OLS (re)	Poisson	Poisson (re)	Neg. Binomial
-6	-.166***	-.188***	.557***	.508***	.549***
-5	-.159***	-.177***	.579***	.543***	.574***
-4	-.147***	-.162***	.614***	.579***	.604***
-3	-.131***	-.144***	.665***	.628***	.649***
-2	-.105***	-.118***	.743***	.704***	.723***
-1	-.051***	-.065***	.899***	.848***	.869***
Event=0	.864***	.850***	2.714***	2.559***	2.803***
1	.100***	.091***	1.269***	1.221***	1.262***
2	.044***	.039***	1.118***	1.098***	1.111***
3	.019***	.018***	1.049**	1.046***	1.042**
4	.007*	.007**	1.016	1.017	1.011
2009-10	-.004**	-0.002	0.977	0.979	.974*
2010-11	0.001	0.002	1.008	0.999	0.998
2011-12	.005***	.006***	1.015***	1.018***	1.014**
2012-13	.005***	.005***	1.016***	1.019***	1.021***
2013-14	.004**	.005***	1.018**	1.022***	1.019**
2014-15	.004*	.006**	1.017*	1.016*	1.012
Age in 0910	.004***	.004***	1.025***	1.022***	1.023***
(squared)	.000***	.001***	.999***	.999***	.999**
Male	-.088***	-.085***	.871***	.825***	.798***
IMD 1	-.086***	-.085***	.822***	.810***	.789***
IMD 2	-.047***	-.047***	.921***	.910***	.906***
IMD 3	-.023***	-.023***	.951**	.951**	.949**
IMD 4	-.009**	-.009**	0.979	0.979	0.979
IMD 6	.009***	.009***	1.019	1.023	1.027
IMD 7	.019***	.019***	1.026	1.029	1.036*
IMD 8	.020***	.021***	1.039*	1.047**	1.050**
IMD 9	.013***	.013***	1.017	1.015	1.020
IMD 10	-.025***	-.025***	.956*	.949**	.947*
Black	.011**	.012**	1.361***	1.329***	1.383***
Asian	-.019***	-.019***	1.112**	1.064*	1.076*
Other	-.012**	-.012***	1.027	1.009	1.007
Mixed	0.000	0.001	0.995	0.985	0.988
Constant	.296***	.297***	.381***	.431***	.404***
lnalpha (dispersion parameter)				.791***	1.816***
N			3,233,056		
R-sq	0.221	0.221			
AIC	6,528,100.9		10,818,748.0	7,951,664.5	7,999,096.9
BIC	6,528,542.5		10,819,190.0	7,952,119.2	7,999,551.5

*p<0.05; **p<0.01; ***p<0.001; Poisson and Negative Binomial show incident rate ratios.

Table A4.4: Model Comparison, full estimates (Number of emergency admissions per year)

	OLS	OLS (re)	Poisson	Poisson (re)	Neg. Binomial
-6	-.168***	-.186***	.379***	.362***	.376***
-5	-.161***	-.176***	.409***	.394***	.406***
-4	-.149***	-.163***	.445***	.431***	.440***
-3	-.133***	-.145***	.501***	.485***	.494***
-2	-.108***	-.119***	.591***	.573***	.582***
-1	-.056***	-.069***	.780***	.753***	.764***
Event=0	.812***	.799***	3.512***	3.388***	3.596***
1	.082***	.073***	1.339***	1.302***	1.329***
2	.034***	.029***	1.146***	1.129***	1.137***
3	.013***	.012***	1.060***	1.056***	1.053***
4	0.003	0.003	1.018	1.019	1.015
2009-10	.011***	.013***	1.038***	1.034***	1.039***
2010-11	.013***	.014***	1.040***	1.029***	1.037***
2011-12	.008***	.009***	1.027***	1.026***	1.028***
2012-13	.003**	.003**	1.012***	1.013***	1.015***
2013-14	-0.001	0.001	1.000	1.002	0.999
2014-15	-0.002	-0.001	0.995	0.996	.988*
Age in 0910	.001***	.001***	1.016***	1.015***	1.017***
(squared)	.001***	.001***	0.999	1.000	0.999
Male	-.016***	-.015***	.962***	.945***	.935***
IMD 1	-.077***	-.077***	.761***	.762***	.725***
IMD 2	-.046***	-.046***	.850***	.853***	.829***
IMD 3	-.024***	-.024***	.918***	.923***	.909***
IMD 4	-.009***	-.009***	.976*	.979*	.976*
IMD 6	.009***	.009***	1.031**	1.029**	1.038**
IMD 7	.019***	.019***	1.068***	1.065***	1.079***
IMD 8	.022***	.023***	1.080***	1.079***	1.098***
IMD 9	.020***	.021***	1.079***	1.074***	1.095***
IMD 10	-.007**	-.006**	.968***	.962***	.965**
Black	-.014***	-.014***	1.035	1.047	1.071*
Asian	-.020***	-.020***	.923***	.915***	.910***
Other	-0.001	-0.001	1.009	1.004	1.008
Mixed	-.020***	-.020***	.907***	.908***	.910***
Constant	.192***	.197***	.269***	.287***	.259***
lnalpha (dispersion parameter)				.726***	1.452***
N			3,233,056		
R-sq	0.261	0.261			
AIC	5,225,048.5		6,765,854.2	5,703,301.2	5,975,822.6
BIC	5,225,490.1		6,766,295.8	5,703,755.8	5,976,277.2

*p<0.05; **p<0.01; ***p<0.001; Poisson and Negative Binomial show incident rate ratios.

Table A4.5: Model Comparison, full estimates (non-drug vs. drug only)

	Non-drug admissions		Drug-related admissions	
	Poisson (re)	Neg. Binomial	Poisson (re)	Neg. Binomial
-6	.605***	.686***		
-5	.649***	.717***		
-4	.694***	.754***		
-3	.752***	.809***		
-2	.841***	.901***		
-1	1.008	1.079**		
Event=0	1.404***	1.515***	8.468***	8.451***
1	1.181***	1.248***	1.393***	1.391***
2	1.092***	1.122***	1.121***	1.118***
3	1.051***	1.055**	1.035	1.032
4	1.027*	1.027	0.973	0.970
2009-10	0.976	0.975	.989***	.387***
2010-11	0.992	0.996	0.999	1.010**
2011-12	1.008	1.008	1.019***	1.019***
2012-13	1.028***	1.025***	1.001	1.002
2013-14	1.045***	1.033***	.984***	.985**
2014-15	1.061***	1.044***	.957***	.959***
Age in 0910	1.028***	1.026***	1.015***	1.016***
(squared)	.999*	0.999	.999***	.999***
Male	.736***	.736***	1.039***	1.040***
IMD 1	.788***	.785***	.856***	.851***
IMD 2	.903***	.908***	.904***	.900***
IMD 3	.953*	.955*	.935***	.933***
IMD 4	0.979	0.979	.976*	.975*
IMD 6	1.026	1.028	1.017*	1.018*
IMD 7	1.025	1.029	1.043***	1.045***
IMD 8	1.043	1.043	1.061***	1.063***
IMD 9	0.991	0.997	1.081***	1.084***
IMD 10	.911***	.913***	1.041***	1.042***
Black	1.487***	1.500***	.936***	.933***
Asian	1.139***	1.141***	.837***	.833***
Other	1.031	1.029	.936***	.934***
Mixed	0.980	0.986	0.977	0.976
Constant	.291***	.287***	.110***	.109***
lnalpha	1.609***	3.457***	.131***	.038***
N		3,233,056		
AIC	6,795,188.7	6,719,799.6	2,315,741.9	2,349,134.8
BIC	6,795,643.3	6,720,254.2	2,316,196.5	2,349,589.4

*p<0.05; **p<0.01; ***p<0.001; Poisson and Negative Binomial show incident rate ratios.

Table A4.6: Sensitivity analyses; cost of admissions

	Directly costed activity only		Excluding hesid if event pre-11/12	
	Poisson (re)	Neg. Binomial	Poisson (re)	Neg. Binomial
-6	.502***	.519***	.518***	.523***
-5	.568***	.559***	.569***	.562***
-4	.625***	.595***	.621***	.599***
-3	.698***	.638***	.681***	.645***
-2	.803***	.732***	.783***	.745***
-1	1.032	.918***	0.989	.924***
Event=0	3.347***	3.269***	3.094***	3.082***
1	1.530***	1.409***	1.423***	1.389***
2	1.274***	1.199***	1.178***	1.165***
3	1.147***	1.091***	1.042***	1.053***
4	1.074***	1.043**		
2009-10	.972*	.934***	.893***	.877***
2010-11	0.992	0.983	.962***	.962**
2011-12	1.007	.966***	.952***	.921***
2012-13	.771***	.777***	.716***	.732***
2013-14	.839***	.847***	.835***	.845***
2014-15	.748***	.753***	.737***	.743***
Age in 0910	1.028***	1.032***	1.026***	1.032***
(squared)	.999***	.999***	.999**	.999***
Male	.879***	.801***	.922***	.838***
IMD Dec 1	.744***	.726***	.782***	.759***
IMD Dec 2	.858***	.856***	.879***	.874***
IMD Dec 3	.922***	.925***	.931***	.934***
IMD Dec 4	.962*	.959*	.964*	.961*
IMD Dec 6	1.020	1.029	1.023	1.029
IMD Dec 7	1.035*	1.032	1.037*	1.035
IMD Dec 8	1.051**	1.054*	1.046**	1.050*
IMD Dec 9	1.003	1.008	1.002	1.002
IMD Dec10	.906***	.904***	.906***	.892***
Black	1.547***	1.557***	1.534***	1.585***
Asian	1.032	1.028	1.029	1.029
Other	1.087***	1.052*	1.104***	1.089***
Mixed	1.170***	1.138***	1.174***	1.154***
Constant	410.737***	401.173***	518.307***	500.267***
lnalpha	2.216***	20.430***	1.356***	17.589***
N	3,072,499		2,642,626	
AIC	$5.75*10^9$	24,592,279	$5.64*10^9$	23,431,722
BIC	$5.75*10^9$	24,592,732	$5.64*10^9$	23,432,157

*p<0.05; **p<0.01; ***p<0.001; Poisson and Negative Binomial show incident rate ratios.

Table A4.7: Marginal effects (Sensitivity analyses)

Event Time	Directly costed activity only						Excluding hesid if event pre-11/12					
	Poisson (fe)			Neg. Binomial			Poisson (fe)			Neg. Binomial		
	ME	95% CI		ME	95% CI		ME	95% CI		ME	95% CI	
-6	£456.60	£424.89	£488.32	£464.87	£427.42	£502.32	£594.44	£555.11	£633.77	£596.03	£552.06	£640.01
-5	£516.89	£488.01	£545.77	£500.90	£468.55	£533.26	£653.16	£617.85	£688.47	£640.54	£602.27	£678.82
-4	£568.84	£542.34	£595.34	£532.80	£503.76	£561.84	£713.70	£682.22	£745.18	£683.76	£649.54	£717.98
-3	£635.10	£609.78	£660.43	£571.52	£544.07	£598.98	£781.70	£753.03	£810.37	£735.49	£703.37	£767.60
-2	£731.88	£706.92	£756.85	£655.13	£627.81	£682.45	£899.87	£872.32	£927.41	£849.80	£817.90	£881.70
-1	£939.45	£909.64	£969.25	£822.26	£789.02	£855.51	£1,136.02	£1,103.43	£1,168.60	£1,053.42	£1,014.89	£1,091.95
0	£3,047.61	£2,961.37	£3,133.86	£2,926.61	£2,823.72	£3,029.51	£3,554.21	£3,464.56	£3,643.86	£3,512.67	£3,401.98	£3,623.36
1	£1,393.52	£1,350.27	£1,436.77	£1,262.05	£1,215.03	£1,309.07	£1,634.37	£1,589.40	£1,679.35	£1,582.87	£1,529.62	£1,636.12
2	£1,160.16	£1,122.30	£1,198.03	£1,073.53	£1,032.51	£1,114.56	£1,353.06	£1,312.69	£1,393.42	£1,328.16	£1,281.48	£1,374.83
3	£1,044.38	£1,008.02	£1,080.74	£976.63	£937.86	£1,015.41	£1,197.33	£1,158.33	£1,236.32	£1,200.24	£1,154.79	£1,245.68
4	£977.71	£942.68	£1,012.73	£933.44	£893.89	£972.98	£1,148.65	£1,106.82	£1,190.47	£1,139.91	£1,091.16	£1,188.65
5	£910.52	£873.93	£947.11	£895.32	£853.73	£936.92						

Marginal effect showing predicted value of y at year=2012-13; IMD decile=5; Age in 0910=35; Male=1; assuming (re)=0.

Table A4.8: Comparing differences in predicted costs between sensitivity models and full model

Event Time	Full model	Directly costed activity only			Excluding hesid if event pre 2011-12		
	Predicted cost (ME)	Predicted cost (ME)	Difference	(as %)	Predicted cost (ME)	Difference	(as %)
-6	£580.84	£464.87	£115.97	19.97%	£596.03	-£15.19	-2.55%
-5	£627.77	£500.90	£126.87	20.21%	£640.54	-£12.77	-1.99%
-4	£671.24	£532.80	£138.44	20.62%	£683.76	-£12.52	-1.83%
-3	£724.14	£571.52	£152.62	21.08%	£735.49	-£11.35	-1.54%
-2	£839.62	£655.13	£184.49	21.97%	£849.80	-£10.18	-1.20%
-1	£1,055.92	£822.26	£233.66	22.13%	£1,053.42	£2.50	0.24%
0	£3,534.49	£2,926.61	£607.88	17.20%	£3,512.67	£21.82	0.62%
1	£1,677.69	£1,262.05	£415.64	24.77%	£1,582.87	£94.82	5.99%
2	£1,455.31	£1,073.53	£381.78	26.23%	£1,328.16	£127.15	9.57%
3	£1,355.44	£976.63	£378.81	27.95%	£1,200.24	£155.20	12.93%
4	£1,301.83	£933.44	£368.39	28.30%	£1,139.91	£161.92	14.20%
5	£1,267.86	£895.32	£372.54	29.38%	omitted		

Full model shown in Table 3.8. Comparison between marginal effects predicting costs at year=2012-13; IMD decile=5; Age in 0910=35; Male=1; assuming (re)=0.

Table A5.1: Domain 1 in the National Outcomes Framework

	Outcome domain 1: Free from drug(s) of dependence	Eligibility Criteria	Data Source/Lag
1.1	Drug and/or alcohol use significantly improved (i.e. reduced their consumption by statistically significant levels) for all presenting substances at any two Treatment Outcomes Profile (TOP) review (or one review and one treatment exit TOP where applicable) within the last 12 months	Clients 18 or over who cited as a problematic substance any of opiates, crack cocaine, alcohol, cannabis or amphetamines in their latest treatment journey	NDTMS (4 weeks)
	No more than 20% of the total value assigned to this outcome domain can be placed on this initial outcome. Where a client exits treatment before 6 months then the appropriate exit TOP will be sufficient without any prior review TOP	Where an individual presents with only substances not recorded on TOP (currently 1%) then a change of one standard deviation will be sued for the information recorded under 'other substance'. Only one substance will be able to be measured in this way.	

*The Treatment Outcomes Profile (TOP) is a tool that allows clinicians and key workers to keep track of the progress of individuals through their treatment journeys. It consists of a 20 item survey on substance use, injecting risk behaviour, housing, employment, crime, health and quality of life. It is usually performed by an individual's key worker

Table A5.2: Domain 1 in the National Outcomes Framework

1.2	Abstinent from all presenting substances at any two TOP reviews (or one review and one treatment exit TOP where applicable) where this was in the last 12 months. Where a client exits treatment before 6 months then the appropriate exit TOP will be sufficient without a prior review TOP	Clients 18 or over who cited as a problematic substance any of opiates, crack cocaine, alcohol, cannabis or amphetamines in their latest treatment journey - still in structured treatment at the end of the specified payment period. Same treatment of non-TOP substances as above.	NDTMS (4 weeks)
1.3	Planned exit from the treatment component of the recovery journey, free from drug(s) - including alcohol - of dependence	Clients aged 18 or over, not in structured treatment in past 21 days. Assessed as having a dependency on drugs/alcohol which would benefit from structured treatment	NDTMS (4 weeks)
1.4	Discharged from treatment successfully (free from drugs(s) of dependence) and does not re-present in either the treatment system or in the criminal justice system (taken onto the DIP/prison caseload) in the following 12 months	An individual does not show up on NDTMS having started a new treatment journey, nor are they shown as being subsequently taken onto DIP/prison caseload in the 12 months following their successful completion of structured treatment	NDTMS/DIRWEB* (4 weeks)

*DIRWEB was a tool used for reporting of entry into the criminal justice system of individuals which could then be incorporated into NDTMS. It has now been discontinued and Public Health England perform this function

Figure A4b: Coefficient plots comparing event time estimates across models (Total LOS per year)

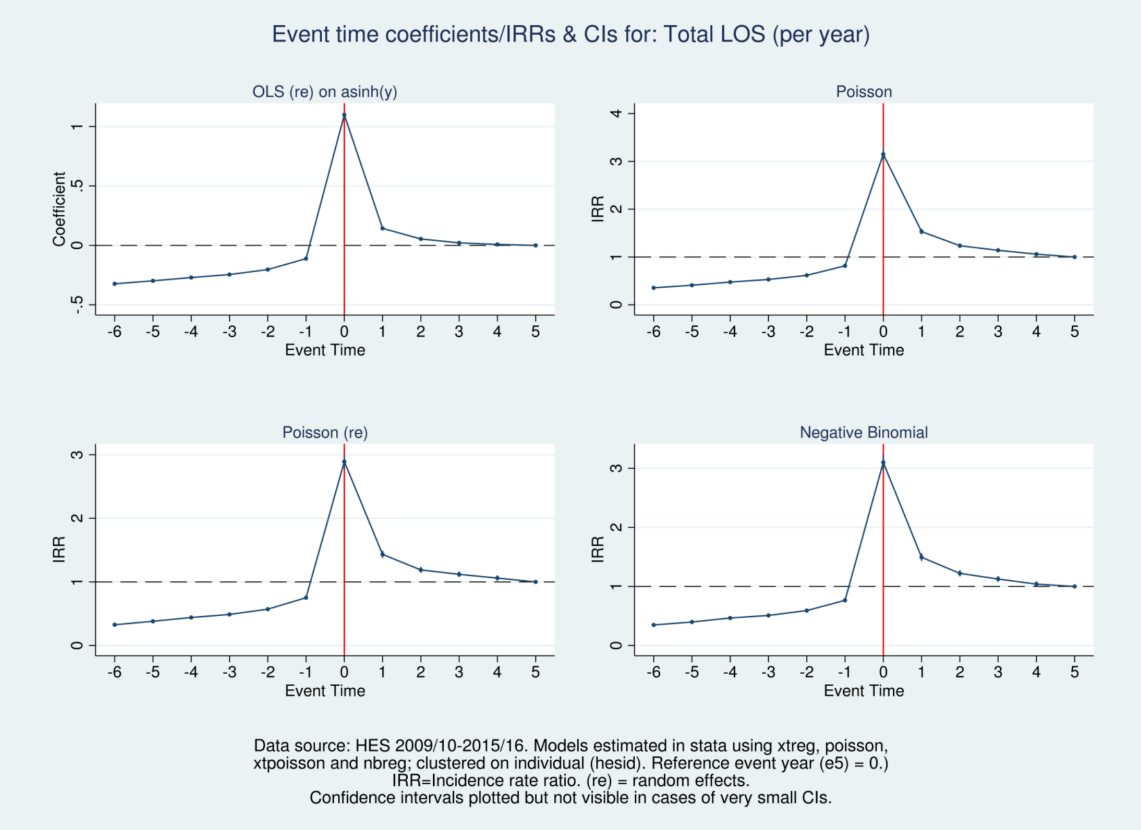


Table A5.3: Difference-in-difference estimates for primary and secondary analyses of completion

	Effect of Pilot Scheme					Number of	
Completion	Marginal Effect	Odds Ratio	Std. Err.	[95% Conf.	Interval]	Individuals	DATs
Main Analysis	-0.013	0.859	0.038	[0.788	0.937]	346,300	149
Secondary Analysis I	-0.010	0.895	0.045	[0.812	0.988]	108,306	50
Secondary Analysis II	-0.016	0.834	0.038	[0.764	0.912]	224,688	98

Table A5.4: Difference-in-difference estimates for primary and secondary analyses of declining

	Effect of Pilot Scheme					Number of	
Declined to Remain in Treatment	Marginal Effect	Odds Ratio	Std. Err.	[95% Conf.	Interval]	Individuals	DATs
Main Analysis	0.006	2.934	0.505	[2.094	4.113]	346,300	149
Secondary Analysis I	0.006	3.014	0.586	[2.059	4.411]	108,306	50
Secondary Analysis II	0.005	2.827	0.5	[1.999	3.998]	224,688	98

Table A6.1: DiD and LDV estimates of treatment effects - admission variables (intention to treat)

		All admissions			Emergency admissions			Drug-related admissions			Non-drug related		
		OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences													
Treatment effect	DiD	0.018***	1.043**	1.044**	0.013***	1.046***	1.043***	0.007**	1.042**	1.041**	0.015***	1.044*	1.044*
	[95% CI]	[0.012; 0.024]	[1.017; 1.07]	[1.016; 1.073]	[0.008; 0.018]	[1.024; 1.068]	[1.021; 1.066]	[0.003; 0.01]	[1.016; 1.068]	[1.015; 1.067]	[0.009; 0.02]	[1.01; 1.078]	[1.007; 1.082]
	AIC		10,674,606	10,177,262		8,169,779	7,966,474		5,098,709	4,921,840		8,352,349	8,215,341
	R-sq.	0.03			0.019			0.004			0.045		
	N		3,893,165			3,893,165			3,893,165			3,893,165	
Test of parallel trends	Year		0.026***			0.018***			0.009***			0.019***	
	Year x Pilot		0.001			-0.003			0.001			0.002	
	Pr > chi-sq.		0.849			0.209			0.916			0.406	
Lagged dependent variable													
Treatment effect	Intervention=1	0.008**	1.019	1.01	0.008***	1.055***	1.026***	0.001	1.024*	1.007	0.008*	1.019	1.013
	[95% CI]	[0.003; 0.012]	[0.99; 1.049]	[0.999; 1.021]	[0.004; 0.011]	[1.036; 1.075]	[1.015; 1.038]	[-0.001; 0.004]	[1.005; 1.042]	[0.995; 1.02]	[0.003; 0.012]	[0.979; 1.059]	[0.999; 1.027]
	AIC		7,326,379	5,372,016		5,318,961	4,159,661		2,937,370	2,673,626		6,251,723	4,277,947
	R-sq.	0.144			0.23			0.077			0.171		
	N		2,185,067			2,185,067			2,185,067			2,185,067	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.2: DiD and LDV estimates of treatment effects - admission variables (treatment effect on treated (non-movers))

		All admissions			Emergency admissions			Drug-related admissions			Non-drug related		
		OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences													
Treatment effects	DiD	0.018***	1.042**	1.041*	0.013***	1.049***	1.046***	0.005**	1.031*	1.029*	0.016***	1.046*	1.043*
	[95% CI]	[0.011; 0.024]	[1.012; 1.073]	[1.008; 1.074]	[0.007; 0.019]	[1.027; 1.072]	[1.023; 1.069]	[0.001; 0.009]	[1.004; 1.058]	[1.003; 1.057]	[0.01; 0.022]	[1.007; 1.086]	[1.001; 1.089]
	AIC		8,532,625	8,129,197		6,424,780	6,258,027		4,052,085	3,931,168		6,584,769	6,489,667
	R-sq.	0.039			0.026			0.003			0.056		
	N		3,276,435			3,276,435			3,276,435				
Test of parallel trends	Year		0.025***			0.016***			0.009***			0.017***	
	Year*Pilot		0			-0.002			-0.001			0.003	
	Pr > chi-sq.		0.701			0.319			0.687			0.191	
Lagged dependent variable													
Treatment effects	Intervention=1	0.013***	1.026	1.02*	0.011***	1.059***	1.037***	0.002	1.018*	1.009	0.014***	1.027	1.029***
	[95% CI]	[0.008; 0.018]	[0.994; 1.058]	[1.008; 1.033]	[0.007; 0.014]	[1.039; 1.079]	[1.025; 1.049]	[-0.001; 0.004]	[1.001; 1.036]	[0.996; 1.022]	[0.009; 0.019]	[0.984; 1.073]	[1.013; 1.046]
	AIC		5,761,740	4,293,859		3,995,609	3,278,192		2,295,352	2,148,172		4,899,219	3,374,311
	R-sq.	0.142			0.226			0.059			0.174		
	N		1,836,925			1,836,925			1,836,925			1,836,925	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.3: DiD and LDV estimates of treatment effects - admission variables (intention to treat, matched comparators)

		All admissions			Emergency admissions			Drug-related admissions			Non-drug related		
		OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences													
Treatment effect	DiD	0.018***	1.051***	1.044**	0.011***	1.039***	1.036**	0.007**	1.042**	1.039**	0.014***	1.054**	1.043*
	[95% CI]	[0.011; 0.024]	[1.022; 1.079]	[1.014; 1.075]	[0.006; 0.017]	[1.016; 1.063]	[1.012; 1.059]	[0.003; 0.011]	[1.014; 1.069]	[1.012; 1.067]	[0.008; 0.019]	[1.018; 1.091]	[1.004; 1.038]
	AIC		3,944,602	3,773,372		3,032,973	2,950,844		1,866,504	1,804,769		3,104,722	3,066,207
	R-sq.	0.031			0.019			0.003			0.045		
	N		1,452,303			1,452,303			1,452,303			1,452,303	
Test of parallel trends	Year		0.028***			0.018***			0.009***			0.021***	
	Year*Pilot		-0.001			-0.002			-6.32E-07			0.001	
	Pr > chi-sq.		0.725			0.36			0.999			0.866	
Lagged dependent variable													
Treatment effect	Intervention=1	0.006*	1.035*	1.01	0.008***	1.054***	1.033***	0.002	1.004	1.009	0.006*	1.044*	1.01
	[95% CI]	[0.001; 0.011]	[1.006; 1.066]	[0.999; 1.021]	[0.005; 0.012]	[1.035; 1.074]	[1.021; 1.046]	[-0.001; 0.004]	[0.975; 1.033]	[0.996; 1.023]	[0.001; 0.010]	[1.005; 1.084]	[0.996; 1.024]
	AIC		2,667,719	1,992,683		1,913,654	1,544,872		1,056,980	982,196		2,276,406	1,597,509
	R-sq.	0.152			0.232			0.079			0.181		
	N		814,540			814,540			814,540			814,540	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.4: DiD & LDV estimates of treatment effects - admission variables (treatment effect on treated (non-movers), matched comparators)

		All admissions			Emergency admissions			Drug-related admissions			Non-drug related		
		OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB	OLS	Poisson	NB
						Difference-in-differences							
Treatment effects	DiD	0.017***	1.049**	1.041*	0.011***	1.042***	1.039**	0.005*	1.03*	1.029*	0.014***	1.055**	1.042
	[95% CI]	[0.009; 0.024]	[1.017; 1.081]	[1.007; 1.075]	[0.005; 0.017]	[1.018; 1.067]	[1.015; 1.065]	[0.001; 0.009]	[1.002; 1.059]	[1.001; 1.057]	[0.008; 0.02]	[1.014; 1.098]	[0.998; 1.089]
	AIC		3,201,070	3,062,610		2,427,881	2,360,901		1,506,211	1,464,464		2,489,335	2,464,941
	R-sq.	0.038			0.025			0.002			0.055		
	N		1,236,001			1,236,001			1,236,001			1,236,001	
Test of parallel trends	Year		0.027***			0.017***			0.009***			0.02***	
	Year*Pilot		-0.002			-0.002			-0.001			0.001	
	Pr > chi-sq.		0.566			0.317			0.603			0.936	
						Lagged dependent variable							
Treatment effects	Intervention=1	0.009**	1.029	1.018**	0.009***	1.029	1.034***	0.002	1.016	1.01	0.009**	1.037	1.021*
	[95% CI]	[0.003; 0.014]	[0.998; 1.061]	[1.006; 1.031]	[0.005; 0.013]	[0.984; 1.076]	[1.02; 1.047]	[-0.001; 0.004]	[0.992; 1.041]	[0.996; 1.024]	[0.004; 0.014]	[0.995; 1.08]	[1.005; 1.038]
	AIC		2,125,250	1,617,988		1,441,370	1,239,096		834,315	800,438		1,807,525	1,282,056
	R-sq.	0.147			0.231			0.06			0.178		
	N		692,351			692,351			692,351			692,351	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Figure A4c: Coefficient plots comparing event time estimates across models (Number of hospital admissions per year)

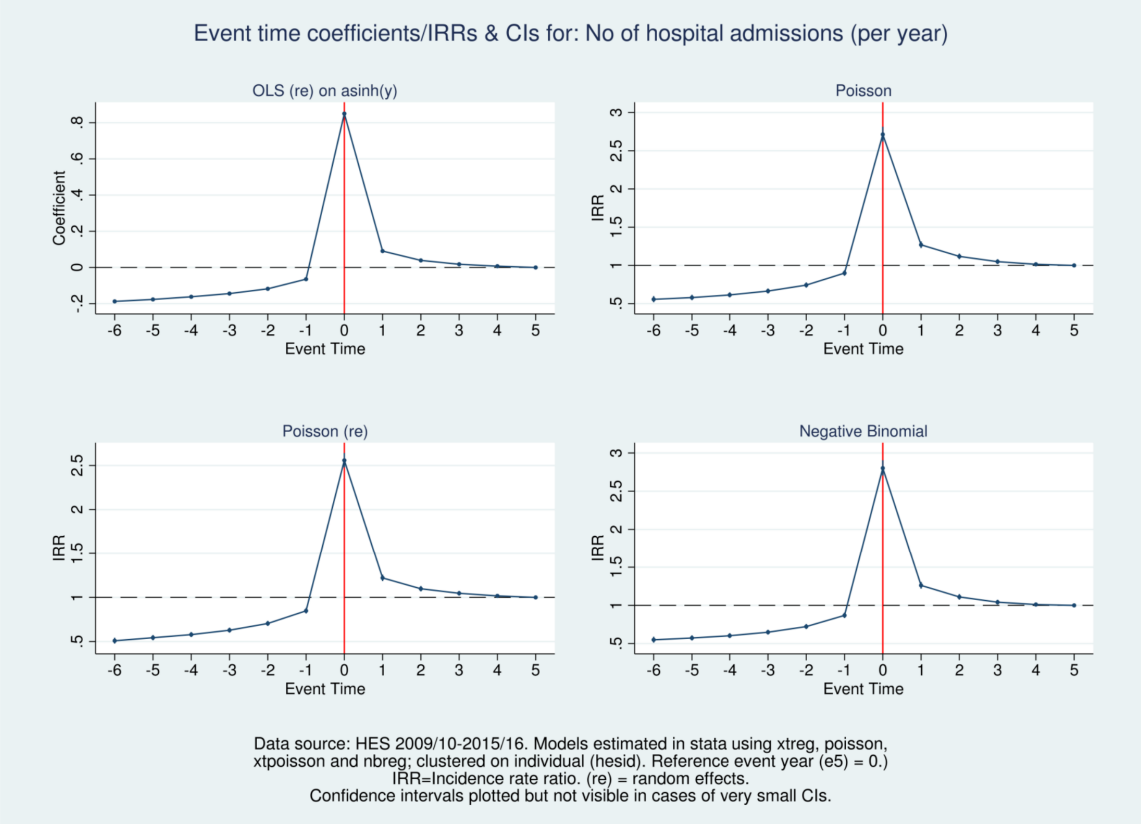


Figure A4d: Coefficient plots comparing event time estimates across models (Number of emergency admissions per year)

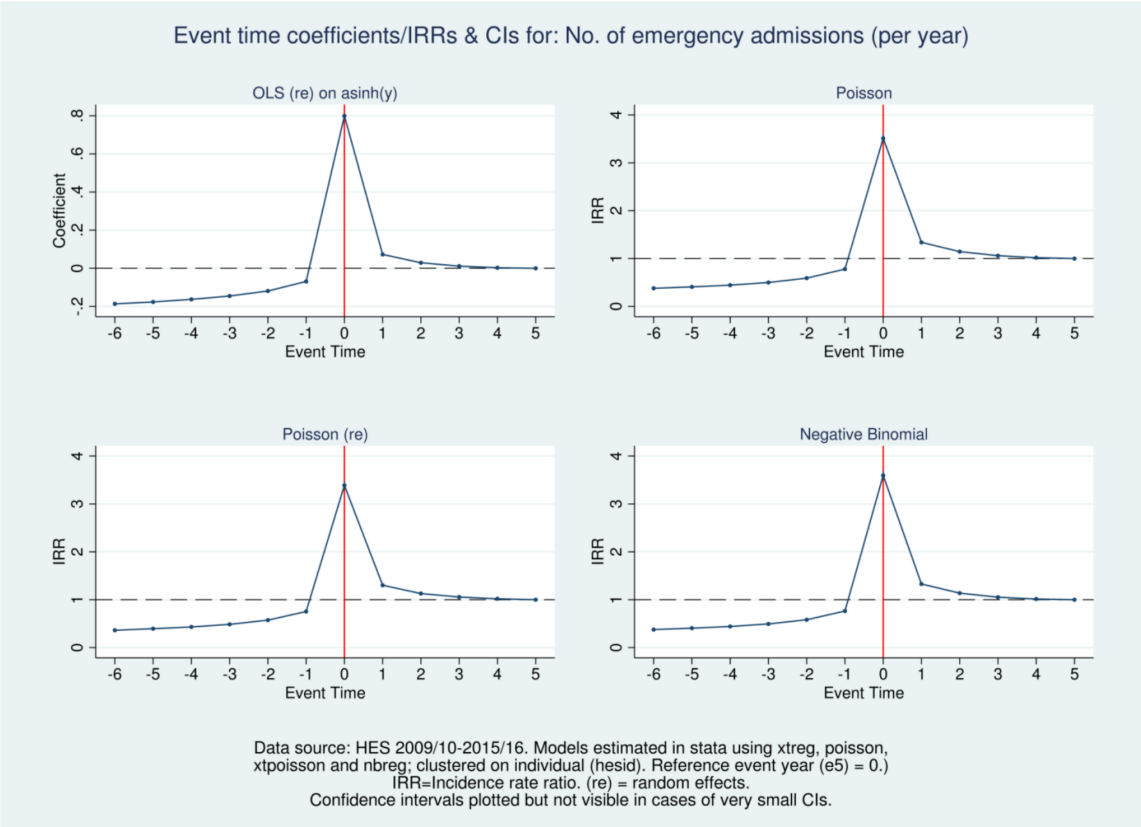


Figure A4e: Coefficient plots comparing event time estimates across models (drug vs. non-drug admissions)

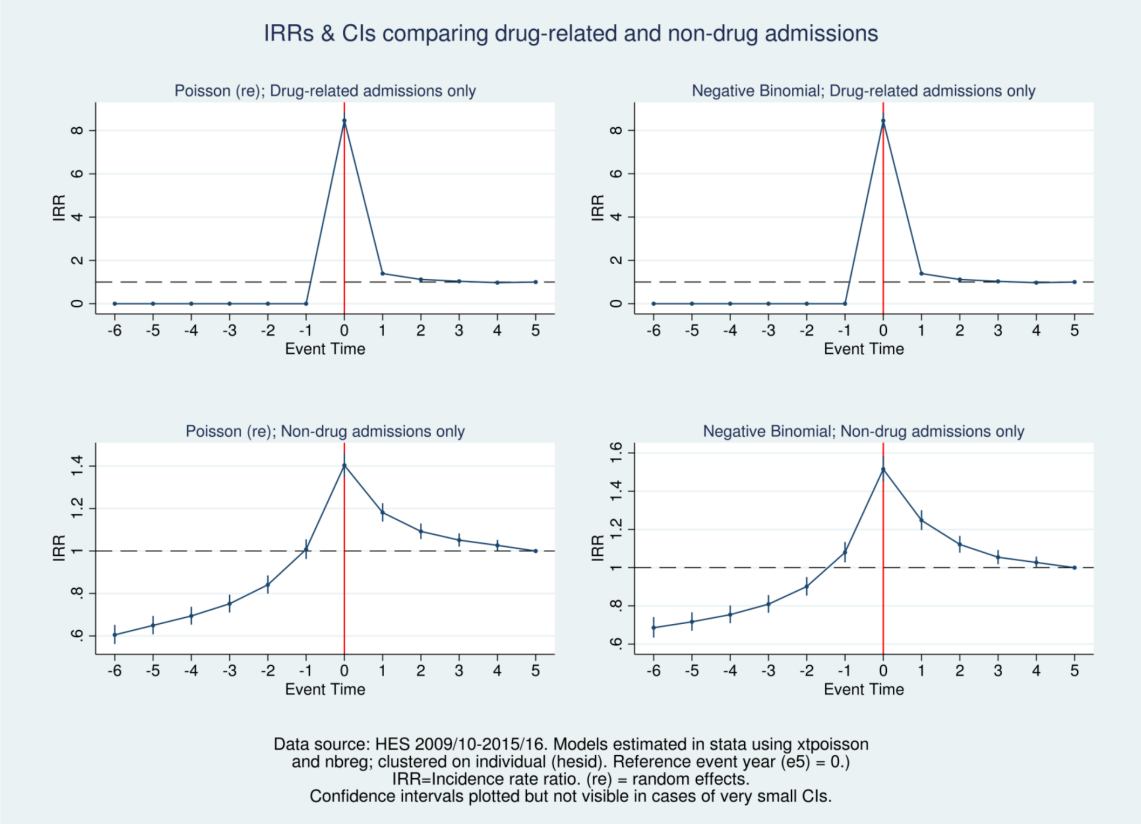


Figure A4f: Coefficient plots illustrating estimates on event dummies from sensitivity analyses (Y=cost)

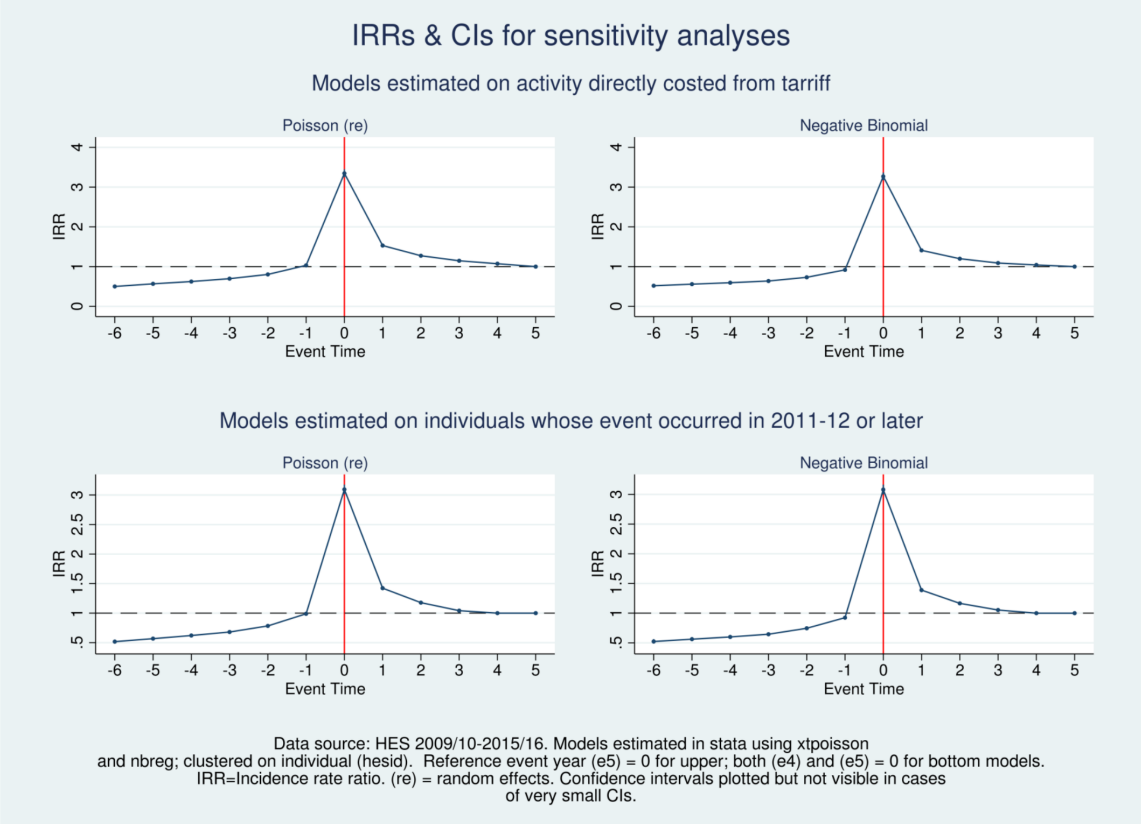


Table A6.5: DiD and LDV estimates of treatment effects - costs and LOS (intention to treat)

		Cost of admissions			LOS		
		OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences							
Treatment effect	DiD	0.085***	1.062***	1.056**	0.02***	1.083***	1.085***
	[95% CI]	[0.051; 0.119]	[1.029; 1.096]	[1.024; 1.089]	[0.009; 0.030]	[1.040; 1.128]	[1.039; 1.132]
	AIC		9.93E+09	36,514,980		44,702,455	11,883,702
	R-sq.	0.029			0.039		
	N		3,893,165			3,893,165	
Test of parallel trends	Year		0.136***			0.042***	
	Year*Pilot		-0.004			0.002	
	Pr > chi-sq.		0.725			0.571	
Lagged dependent variable							
Treatment effect	DiD	0.039*	0.987	0.995	-0.01*	0.96	0.997
	[95% CI]	[0.015; 0.064]	[0.964; 1.01]	[0.981; 1.01]	[-0.018; -0.002]	[0.887; 1.039]	[0.964; 1.03]
	AIC		9.49E+09	20,730,013		41,643,139	6,663,933
	R-sq.	0.08			0.156		
	N		2,185,067			2,185,067	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.6: DiD and LDV estimates of treatment effects - costs and LOS (treatment effect on treated (non-movers))

		Cost of admissions			LOS		
		OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences							
Treatment effect	DiD	0.087***	1.074***	1.063**	0.021***	1.099***	1.092***
	[95% CI]	[0.051; 0.123]	[1.035; 1.114]	[1.024; 1.102]	[0.010; 0.031]	[1.05; 1.149]	[1.04; 1.146]
	AIC		7.85E+09	29,146,724		33,798,309	9,366,634
	R-sq.	0.036			0.049		
	N		3,276,435			3,276,435	
Test of parallel trends	Year		0.132***			0.037***	
	Year*Pilot		-0.006			0.005	
	Pr > chi-sq.		0.701			0.251	
Lagged dependent variable							
Treatment effect	DiD	0.071***	0.999	1.008	-0.006	1.012	0.999
	[95% CI]	[0.045; 0.097]	[0.975; 1.024]	[0.992; 1.025]	[-0.014; 0.002]	[0.978; 1.047]	[0.963; 1.037]
	AIC		7.42E+09	16,513,722		31,271,151	5,248,957
	R-sq.	0.082			0.163		
	N		1,836,925			1,836,925	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.7: DiD and LDV estimates of treatment effects - costs and LOS (intention to treat, matched comparators)

		Cost of admissions			LOS		
		OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences							
Treatment effects	DiD	0.088***	1.069***	1.054***	0.014**	1.055*	1.058*
	[95% CI]	[0.052; 0.124]	[1.035; 1.106]	[1.021; 1.089]	[0.004; 0.025]	[1.011; 1.101]	[1.011; 1.107]
	AIC		3.66E+09	13,622,052		16,455,902	4,392,238
	R-sq.	0.029			0.042		
	N		1,452,303			1,452,303	
Test of parallel trends	Year		0.144***			0.047***	
	Year*Pilot		-0.012			-0.002	
	Pr > chi-sq.		0.425			0.585	
Lagged dependent variable							
Treatment effects	Treatment=1	0.033*	0.001	1.002	-0.015***	0.979	0.977
	[95% CI]	[0.007; 0.059]	[0.976; 1.026]	[0.987; 1.019]	[-0.022; -0.005]	[0.94; 1.013]	[0.943; 1.012]
	AIC		3.44E+09	7,731,832		15,315,566	2,468,426
	R-sq.	0.089			0.162		
	N		814,540			814,540	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for Poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.

Table A6.8 : DiD and LDV estimates of treatment effects - costs and LOS (treatment effect on treated (non-movers), matched comparators)

		Cost of admissions			LOS		
		OLS	Poisson	NB	OLS	Poisson	NB
Difference-in-differences							
Treatment effects	DiD	0.087***	1.077***	1.059**	0.014*	1.075**	1.075**
	[95% CI]	[0.049; 0.126]	[1.036; 1.119]	[1.019; 1.099]	[0.003; 0.025]	[1.025; 1.128]	[1.021; 1.132]
	AIC		2.95E+09	11,061,802		12,753,192	3,527,075
	R-sq.	0.036			0.051		
	N		1,236,001			1,236,001	
Test of parallel trends	Year		0.146***			0.044***	
	Year*Pilot		-0.018			-0.001	
	Pr > chi-sq.		0.264			0.77	
Lagged dependent variable							
Treatment effects	Intervention=1	0.041**	0.973	0.999	-0.012**	0.975	0.959*
	[95% CI]	[0.013; 0.069]	[0.943; 1.004]	[0.982; 1.017]	[-0.021; -0.004]	[0.935; 1.017]	[0.923; 0.997]
	AIC		2.74E+09	6,265,964		11,816,917	1,980,838
	R-sq.	0.089			0.166		
	N		692,351			692,351	

Coefficients displayed for OLS models of the asinh(y); incidence rate ratios for poisson and negative binomial. Test of parallel trends estimated on the full model in the pre-period (2009-10 to 2011-12); LDV estimated for post years only (2012-13 to 2015-16). OLS and Poisson include individual random effect, all models clustered on hesid. *=p<0.05; **=p<0.01; ***p=<0.001.