

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Investigating the Economic Impact of Different Patterns of Medication Monitoring in General Practices

Vladislav Berdunov

BA

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

October 2015

Abstract

Medication monitoring is used in primary care to guide treatment and protect the patient from adverse drug events (ADEs). Early detection of ADEs may prevent their development and avoid serious or permanent effects to patient health. Published guidelines for the conduct of medication monitoring in UK primary care recommend monitoring in patients with long-term prescription of certain cardiovascular medications, including angiotensin-converting enzyme inhibitors (ACEI), loop diuretics and amiodarone. Although much evidence exists on the practice of monitoring of these medications in primary care, few studies have considered the effect of carrying out monitoring at different frequencies during long-term therapy on the risk of ADEs. Similarly, evidence of the economic effect of regular monitoring of these medications is sparse. As a consequence, policy-makers rely primarily on evidence from expert opinion as a basis for recommendations and compliance to guidelines in practice is poor. This programme of research aimed to gain an understanding of the nature of monitoring of ACEI, loop diuretics and amiodarone in primary care in England, and investigate the effectiveness and cost-effectiveness of monitoring in accordance with current guidelines.

Using electronic health records from the Clinical Practice Research Datalink (CPRD) and Hospital Episode Statistics (HES), a cohort study was carried out in order to quantify the effect of monitoring the study medications at different frequencies during the course of long-term therapy. Weights based on propensity scores were used to control for confounding arising from non-random assignment into alternative monitoring regimens in retrospectively observed data. The time-to-event analysis found that conducting thyroid function tests (TFT) in amiodarone therapy below the recommended interval of once in 6 months was associated with more than a two-fold increase in the hazard rate of hypothyroidism (HR 2.79; 95% CI 1.42,5.47), thyrotoxicity (HR 2.57; 1.40,4.74) and major adverse cardiovascular events (HR 2.13; 1.31,3.45), when comparing against a group of patients monitored according to guidelines. Conversely, conducting urea & electrolyte monitoring in ACEI therapy below recommended

frequency was associated with a reduced rate of hyperkalaemia treated in primary care (HR 0.57; 0.42,0.77) and hospital admission (HR 0.13;0.02,0.95). In addition, monitoring urea & electrolytes during loop diuretic therapy below recommended frequency was associated with a lower hazard rate of hypokalaemia or hyponatraemia treated in primary care (HR 0.17; 0.06,0.45). These findings support current recommendations on frequency of regular monitoring in the case of amiodarone TFT monitoring, but not in the case of ACEI or loop diuretics.

The results of the time-to-event analysis of amiodarone monitoring were used to populate a decision-analytic model designed in order to estimate the cost-effectiveness of different strategies of monitoring medication, compared to recommended practice. This analysis demonstrated that conducting TFT at recommended frequency yielded modest cost savings per patient (£129, compared to less frequent monitoring and £192 compared to more frequent monitoring option) and utility gains per patient (0.0245 quality-adjusted life-years (QALYs) compared to less frequent monitoring and 0.0543 QALYs compared to more frequent monitoring option). The probabilistic model estimated that the recommended frequency strategy had a 97% probability of being a cost saving option compared to the two alternatives. These findings support the current policy of encouraging 6-monthly monitoring of TFT in amiodarone therapy.

This programme of research has demonstrated that medication monitoring is potentially effective and cost-effective in amiodarone therapy, but did not find regular monitoring of ACEI or loop diuretic therapy to be effective. Observational research using routinely collected electronic health records can be used to gauge both the clinical and cost-effectiveness of medication monitoring in order to guide practice in this area and improve the safety of medications in primary care.

Acknowledgements

I would like to thank:

My supervisors, Prof. Anthony Avery and Prof. Rachel Elliott for the countless hours of supervision and discussion over the past four years which have guided me through the process of completing this project.

The NIHR School for Primary School Research for providing the funding, training and support which enabled me to carry out this project.

Friends and colleagues in the Division of Social Research in Medicines and Health at the School of Pharmacy, as well as the Division of Primary Care, School of Medicine, University of Nottingham, for all your help and support during my PhD journey.

My parents, Dr. Nikolay Berdunov and Mrs. Alina Berdunova, my sister Miss Lisa Berdunova, for all your love, support and patience.

Miss Lorna Trimnell for your love and encouragement throughout the past three years of my life.

Output arising from this research

Conference presentations

Berdunov, V., Avery. A., Elliott, R. A. (2015). "Cost effectiveness analysis of alternative strategies of monitoring amiodarone-related thyroid toxicity in UK primary care."
Poster presentation at the International Society for Pharmacoeconomics and Outcomes Research conference in Milan, Italy, 9-11th November 2015

Berdunov, V., Avery. A., Elliott, R. A. (2015). "Cost effectiveness of regular monitoring of long-term pharmacotherapy: a case study of amiodarone."
Podium presentation at the Society for Academic Primary Care conference in Oxford, 8-10th July 2015

Berdunov, V., Avery. A., Elliott, R. A. (2014). "An analysis of regular monitoring of high-risk medications in UK primary care."
Elevator pitch presentation at the Society for Academic Primary Care conference in Edinburgh, 9-11th July 2014

Contents

Abstract i
Acknowledgements iii
Output arising from this research iv
List of tables xi
List of figures xiv
Abbreviations xvi
Chapter 1 1
Introduction 1
1.1 Introductory remarks 1
1.2 Background to medication errors 2
1.2.1 Medication safety 2
1.2.2 Definitions used in literature 3
1.2.3 Errors in healthcare 5
1.2.4 Medication errors 7
1.2.5 Definitions used in this thesis: 9
1.2.6 Why study medication safety in primary care? 10
1.3 Background to medication monitoring 11
1.3.1 When is medication monitoring used? 11
1.3.2 Phases of medication monitoring 12
1.3.3 Choice of measurement 13
1.3.4 Timing of monitoring 14
1.3.5 Monitoring errors 16
1.4 Aims of programme of research 16
1.5 Why use observational research to study medication monitoring? 18
1.6 Why carry out economic analyses of medication monitoring? 19
1.7 Outline of thesis 21
1.8 Potential impact of the research described in this thesis 23
Chapter 2 24
2.1 Introduction and aims of the chapter 24
2.2 Background of healthcare safety 24
2.3 Literature review of adverse drug events 27
2.3.1 Summary of selected studies 27
2.3.2 Systematic reviews 27

2.3.3 Summary of individual studies 28
2.3.4 ADEs linked to cardiovascular medications. 29
ACE inhibitors and ARBs 29
Diuretics 32
Amiodarone 32
2.3.5 Critique of methods used in studies of ADE in healthcare 33
2.4 Costs of ADE 39
2.4.1 Hospital-based studies 41
2.4.2 Community care-based studies 41
2.4.3 Other costs: productivity losses 44
2.5 Discussion 47
Chapter 3: 52
Literature review: medication monitoring for the prevention of ADEs 52
3.1 Introduction to the chapter 52
3.2 Aims and objectives of chapter 52
3.3 Background of medication monitoring 52
3.4 Review of guidelines for follow-up monitoring of ACEI, loop diuretics and amiodarone. 55
3.5 Review of studies of completion of monitoring 60
3.5.1 Aims and objectives 60
3.5.2 Review 61
3.5.3 Summary of studies 62
3.5.4 Comment on study design and methods 74
3.6 Review of studies evaluating interventions to improve medication monitoring 76
3.6.1 Objectives of review 76
3.6.2 Search strategy 76
3.6.3 Summary of studies of information technology interventions 77
3.6.4 Studies which evaluated the cost-effectiveness of interventions to improve monitoring 79
3.7 Discussion 81
3.7.1 What we know 81
3.7.2 What we don't know 86
3.7.3 Design aspects 88
3.7.4 Concluding remarks 89
Chapter 4: 91
Methods for data extraction, manipulation and identification of exposures, outcomes and covariates for the analysis 91
4.1 Introduction to chapter 91
4.2 Aims of analysis of the effect of monitoring frequency on risk of ADEs 91

4.3 Rationale for the study. 92
4.4 Comparators 93
4.5 Background of large databases of EHRs 94
4.5.1 Medication safety research using large databases of electronic patient records 94
4.5.2 Advantages and limitations of the use of large databases of EHRs in the context of this study. 95
4.6 Description of data sources used in this thesis 97
4.6.1 General information 97
4.6.2 Representativeness and validation 98
4.6.3 Included data 100
4.6.4 Description of data contained in HES 101
4.6.5 Coding systems 102
4.7 Study design, inclusion criteria and extraction methods 103
4.7.1 Study approval 103
4.7.2 Study design 103
4.7.3 Study medications 103
ACE inhibitors 104
Loop diuretics 104
Amiodarone 104
4.7.4 Data sources 104
4.7.5 Selection of study cohort 104
4.7.6 Initial data cleaning 105
4.8 Identifying exposure 106
4.8.1 Identifying exposure to the study medication 107
4.8.2 Identifying monitoring regimen 111
ACEI and loop diuretics 111
Amiodarone 112
4.9 Identifying outcomes 112
4.9.1 General information 112
4.9.2 ACE inhibitors 112
4.9.3 Loop diuretics: 113
4.9.4 Amiodarone 114
4.10 Comment on methods for identification of exposure and outcomes from retrospective data 114
4.11 Identifying covariates 116
4.11.1 General information 116
4.11.2 Demographic factors 116
4.11.3 Multimorbidity 116

4.11.4 Polypharmacy and co-prescribed cardiovascular medications 117
4.11.5 Specific conditions 117
4.11.6 Other factors 117
4.12 Discussion 118
Chapter 5 121
Statistical methods 121
5.1 Introduction to the chapter 121
5.2 Chapter objectives 121
5.3 Set up of the analysis 122
5.3.1 Aims and objectives of analysis 122
5.3.2 Patient inclusion and cohort selection 122
5.3.3 Exposure 122
5.3.4 Outcomes 123
5.3.5 Patient covariates 124
5.4 Estimating treatment effect from observational data 125
5.4.1 Rationale for the choice of observational cohort design 125
5.5 Methods for controlling non-random assignment 127
5.5.1 Statistical regression analysis 128
5.5.2 Propensity scores 129
5.6 Methods for estimation of causal effect of monitoring 134
5.6.1 Survival analysis: rationale 134
5.6.2 Preparing dataset for time-to-event analysis 134
5.6.3 Specifying the time-to-event regression model 135
5.7 Discussion 137
Chapter 6 139
Results of the time-to-event analysis of monitoring and ADEs linked to the study medications 139
6.1 Introduction to the chapter 139
6.2 ACE inhibitors: full-sample analysis 139
6.2.1 Descriptive analysis 139
6.2.2 Estimation of propensity scores 142
6.2.3 Time-to-event analysis 144
6.3 ACE inhibitors: subset analysis of HES-linked practices 145
6.3.1 Descriptive analysis 145
6.3.2 Estimation of propensity scores 147
6.3.3 Time-to-event analysis 149
6.4 Loop diuretics: full cohort 150
6.4.1 Descriptive analysis 150

6.4.2 Estimation of propensity scores 152
6.4.3 Time-to-event analysis 155
6.5 Loop diuretics: subset analysis of HES-linked practices 155
6.5.1 Descriptive analysis 155
6.5.2 Estimation of propensity scores 157
6.5.3 Time-to-event analysis 160
6.6 Amiodarone cohort 160
6.6.1 Descriptive analysis 160
6.6.2 Estimation of propensity scores 163
6.6.3 Time-to-event analysis 164
6.7 Interval of common support 166
6.8 Analysis of monitoring over calendar-years in the study 168
6.9 Observed ADEs during the study period 169
6.10 Additional analysis 1: an investigation of outcomes in the "no monitoring group" in the ACEI cohort 169
6.11 Additional analysis 2: an investigation of ADEs potentially driven by frequent monitoring in the ACEI cohort 171
6.12 Additional analysis 3: an investigation of potential change in assigned monitoring regimen during the exposure period 172
6.13 Additional analysis 4: separating the time periods used to identify monitoring exposure and ADE outcomes using CPRD data 173
6.14 Additional analysis 5: performance of alternative methods of controlling for non-random assignment to monitoring groups 175
6.15 Discussion 179
6.15.1 Discussion of results of time-to-event analysis 179
6.15.2 Discussion of additional analyses to test methodological assumptions 183
Inclusion of patients with no monitoring 183
Abnormal laboratory test results driven by intensive monitoring 185
Reduction in frequency of monitoring during follow-up 186
Alternative method of identifying exposure and outcomes 186
Comparing different methods of adjusting for confounding from non-random treatment assignment 187
Chapter 7 190
Methods and results of the cost-effectiveness analysis of amiodarone monitoring 190
7.1 Introduction to the chapter 190
7.2 Aims and objectives of the cost-effectiveness analysis 190
7.3 Methods 191
7.3.1 Study approval 191
7.3.2 Model population 191

7.3.3 Setting and perspective 192
7.3.4 Model comparators. 192
7.3.5 Model horizon and discounting 192
7.3.6 Measure of effectiveness 193
7.3.7 Decision-analytic model 194
Model structure 194
State transitions 197
Treatment pathway 200
Probability of death 201
7.3.8 Outcomes 204
7.3.9 Resource use and cost 205
7.3.10 Analysis design 210
Base case analysis 210
Deterministic sensitivity analysis 211
Probabilistic sensitivity analysis 212
Distributional assumptions for transition probabilities 213
Distributional assumptions for costs 213
Distributional assumptions for utilities 214
7.4 Results 214
7.4.1 Base case analysis 214
7.4.2 Deterministic sensitivity analysis 215
Assumptions regarding effect of monitoring on HR of ADEs 216
Effect of uncertainty in cost parameters 217
Effect of uncertainty in utility parameters 218
Effect of uncertainty in transition probabilities 218
7.4.3 Probabilistic sensitivity analysis. 219
7.5 Discussion 221
Chapter 8 229
Discussion 229
8.1 Introduction to the chapter 229
8.2 Summary of findings 229
8.3 Strengths and limitations of the study 237
8.4 How this study fits into and contributes to the existing evidence base 245
8.5 Implications for policy 248
8.6 Implications for research 250
8.7 Concluding remarks 252
References 254
Appendix 273

List of tables

Table 1.1	Key definitions used in the thesis	8
Table 2.1	Summary of the results of studies of hospital-acquired ADE	30
Table 2.2	Summary of the results of studies of community-acquired ADE	31
Table 2.3	Description of study design in studies of incidence of ADE	34
Table 2.4	Summary of studies of cost of community-acquired ADE	46
Table 3.1	Summary of information included in guidelines for follow-up monitoring	59
Table 3.2	Summary of studies of initiation monitoring	72
Table 3.3	Summary of studies of follow-up monitoring	73
Table 3.4	Summary of studies evaluating interventions to improve monitoring	80
Table 6.1	Summary of patient characteristics, ACEI cohort	140
Table 6.2	Coefficient estimates of the probit model, ACEI full dataset	143
Table 6.3	Hazard ratios in Weibull models, ACEI full cohort	145
Table 6.4	Summary of patient characteristics in ACEI cohort, HES-linked practices	146
Table 6.5	Coefficient estimates of the probit model, ACEI HES-linked dataset	148
Table 6.6	Hazard ratios in Weibull models. ACEI HES-linked practices	150
Table 6.7	Summary of patient characteristics, loop diuretic full cohort	151
Table 6.8	Coefficient estimates of the probit model, loop diuretics full cohort	153
Table 6.9	Hazard ratios in Weibull models, loop diuretic full cohort	155
Table 6.10	Summary of patient characteristics in loop diuretic cohort, HES-linked practices	156
Table 6.11	Coefficient estimates of the probit model, loop diuretics HES-linked dataset	158

Table 6.12	Hazard ratios in Weibull model, loop diuretics HES-linked practices	160
Table 6.13	Summary of patient characteristics, amiodarone cohort	161
Table 6.14	Coefficient estimates in probit model, amiodarone cohort	163
Table 6.15	Coefficient estimates in Weibull models, amiodarone cohort	165
Table 6.16	Exclusion of patients with propensity scores outside common support	168
Table 6.17	Summary of observed ADEs across medication-specific cohorts	169
Table 6.18	Number of ADEs in each comparison group, ACEI cohort	169
Table 6.19	Results of sensitivity analysis of inclusion of no monitoring group	171
Table 6.20	Unadjusted and adjusted relative risk of hyperkalaemia and AKI, by monitoring frequency	172
Table 6.21	Breakdown of patients by monitoring frequency	173
Table 6.22	Baseline patient characteristics	173
Table 6.23	Descriptive statistics of patients who completed follow-up	174
Table 6.24	Results of time-to-event analysis of ACEI monitoring, alternative exposure method	174
Table 6.25	Results of time-to-event analysis of ACEI monitoring, by adjustment method	176
Table 6.26	Mean values of covariates across comparison groups, split by adjustment method	177
Table 7.1	Definitions of ADE used in this programme of research	194
Table 7.2	Hazard ratios from the time-to-event analysis	194
Table 7.3	Description of sources of data used to populate state transition probabilities	196
Table 7.4	Shape parameter of hazard function of AIH and AIT	198
Table 7.5	Transition probabilities from euthyroid to AIH, AIT and MACE, by strategy and cycle	199
Table 7.6	Breakdown of death probabilities depending on condition	203
Table 7.7	Age-adjusted probability of death, split by ADE state	204

Table 7.8	Utility weights	205
Table 7.9	Unit costs of care in the model	205
Table 7.10	Unit costs of medications	206
Table 7.11	Composition of "no symptoms" and "untreated AIH" state cost	207
Table 7.12	Cost in untreated AIT state	207
Table 7.13	Cost of medically treated AIH	208
Table 7.14	Cost of medically treated AIT	208
Table 7.15	Cost of surgically treated AIT	209
Table 7.16	Cost of MACE state	210
Table 7.17	Cost post-treatment of AIH and AIT	210
Table 7.18	Results of deterministic analysis	214
Table 7.19	Parameter ranges used in the DSA	215
Table 7.20	Results of probabilistic sensitivity analysis	219
Table A1	Summary of patient numbers per medication-specific cohort	273
Table A2	Breakdown of variables in HES raw data	273
Table A3	List of study medications and corresponding CPRD product codes	274
Table A4	Laboratory tests used to define exposure to monitoring, corresponding CPRD medical and Read codes	276
Table A5	Description of variables used in the study	279
Table A6	List of outcomes assessed in the study	279
Table A7	Criteria for the calculation of the multimorbidity score	284

List of figures

Fig. 1.1	Relationship between the main concepts in medication safety	8
Fig. 1.2	Outline of content of thesis	22
Fig. 4.1	Raw plain text format of prescription data from CPRD	101
Fig. 4.2	Cleaned and formatted prescription data following input into Stata 11	101
Fig. 4.3	Description of cohort selection based on continuation of therapy in the first 9 months for all medication cohorts	108
Fig. 4.4	Description of the approach to medication persistence	110
Fig. 4.5	Alternative method of identifying monitoring exposure and ADE outcomes from CPRD	115
Fig. 6.1	ACEI: proportion of patients appropriately monitored, by practice	141
Fig. 6.2	Loop diuretics: proportion of patients appropriately monitored, by practice	152
Fig. 6.3	Amiodarone: proportion of patients appropriately monitored, by practice	162
Fig. 6.4	Distribution of propensity scores, ACEI full cohort	166
Fig. 6.5	Distribution of propensity scores, ACEI subset of HES-linked practices	166
Fig. 6.6	Distribution of propensity scores, loop diuretics full cohort	167
Fig. 6.7	Distribution of propensity scores, loop diuretics HES dataset	167
Fig. 6.8	Distribution of propensity scores, amiodarone cohort	167
Fig. 6.9	Analysis over calendar-years in the study	168
Fig. 6.10	Mean age across comparison groups, split by adjustment method	177
Fig. 6.11	Proportion of patients with ≥1 comorbid condition, split by adjustment method	177
Fig. 6.12	Proportion of patients with ≥4 co-prescriptions, split by adjustment method	178
Fig. 7.1	Model diagram	197
Fig. 7.2	Expected values of cost and utility accumulated in each strategy	214
Fig. 7.3	One-way sensitivity analysis of estimated effectiveness of TFT monitoring	216

Fig. 7.4	One-way sensitivity analyses of cost parameters	217
Fig. 7.5	One-way sensitivity analyses of utility parameters	218
Fig. 7.6	One-way sensitivity analyses of transition probability parameters	219
Fig. 7.7	Cost-effectiveness plane: recommended vs more frequent monitoring	220
Fig. 7.8	Cost-effectiveness plane: more frequent vs. less frequent monitoring	220
Fig. 7.9	Cost-effectiveness acceptability curve, 5 year horizon	221

Abbreviations

ACEI	Angiotensin-converting enzyme inhibitor
ADE	Adverse drug event
ADR	Adverse drug reaction
AE	Adverse event
AIH	Hypothyroidism
AIT	Thyrotoxicity
AKI	Acute kidney injury
ARB	Angiotensin-II receptor blocker
BMJ	British Medical Journal
Bn	Billion
BNF	British National Formulary
CCB	Calcium channel blocker
CDSS	Computerised decision support system
CKD	Chronic kidney disease
CPOE	Computerised provider order entry
CPRD	Clinical Practice Research Datalink
DM	Deutschmark
DoH	Department of Health
EGFR	Estimated glomerular filtration rate
HER	Electronic health record
GBP	Great British Pound
GP	General practitioner
HR	Hazard ratio
HRG	Healthcare Resource Group
ICD	International classification of diseases
INR	International normalised ratio
IPTW	Inverse probability of treatment weighting
ISAC	Independent Scientific Advisory Committee
MeSH	Medical Subjects Heading
MMWS	Mean weighting through stratification
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OPCS	Office of Population Censuses and Surveys
OR	Odds ratio
OR	Odds ratio
pADE	preventable adverse drug event
PINCER	Pharmacist-led information technology intervention for medication errors
PPV	Positive predictive value
QALY	Quality-adjusted life-year
QOF	Quality and Outcomes Framework
RCT	Randomised controlled trial
RR	Risk ratio

SD	Standard deviation
SE	Standard error
TFT	Thyroid function test
TSH	Thyroid stimulating hormone
U&E	Urea & electrolytes
UK	United Kingdom
US/USA	United States/United States of America
USD	United States Dollar
WHO	World Health Organisation

Chapter 1

Introduction

1.1 Introductory remarks

This thesis presents a programme of research which explored medication monitoring for the prevention of adverse drug events (ADEs) in United Kingdom (UK) general practice. This work comprises several independent but linked analyses which aimed to investigate the effectiveness and cost-effectiveness of regular monitoring of specific medications associated with preventable harm in patients in the UK primary care setting. This thesis aims to bring together current evidence from published literature on medication monitoring and the results of empirical analyses conducted to inform the current policy and practice of regular monitoring of a subset of high-risk medications, make informed judgments regarding the value of monitoring strategies prescribed in current practice guidelines and make recommendations regarding the direction of policy and future research in the area of medication monitoring.

This chapter contains a general introduction to the key concepts used in medication safety research, ADEs in healthcare and medication monitoring. This is followed by an outline of the aims of research, brief description of the methods used to achieve these aims and a summary of each chapter of this thesis. Finally, a comment is made on how the outcomes of this research may impact future research and recommendations for future policy.

1.2 Background to medication errors

1.2.1 Medication safety

Pharmacotherapy is the most common therapeutic method, with around 760 million medications dispensed in the UK community setting in 2006 (Department of Health 2008). The management of medication therapy encompasses several distinct clinical processes, which include prescribing, dispensing, administration, monitoring, as well as education and training of patients and healthcare professionals in the use of medicines (Department of Health 2000).

All medical care interventions inherently carry a certain amount of risk of unintended harm. However, whilst a certain amount of risk during medical treatment is tolerable when appropriately weighed against the benefits of therapy, complications of care arising from error or negligence are preventable and deserve particular attention. Adverse events (AEs) are distinguished from all other types of harm occurring during care by the fact that the injury is caused by medical management, rather than by the underlying disease or medical condition of the patient (Department of Health 2003). Additionally, adverse events can be defined as preventable or non-preventable. The former are occasionally referred to as "negligent" if they are caused by a medical error or substandard care (Gray 2003). Given that a large proportion of medical care is carried out using medications, it is natural that a large proportion of unintended harm in healthcare is linked to medications. Large observational studies in the US reported that between 2.9% to 3.7% of hospital episodes resulted in adverse effects of therapy which led to severe harm to the patient, with medications contributing to around 1 in 5 adverse events (Leape, Bates et al. 1995, Thomas and Brennan 2000). Around the same time as the publication of the above large-scale epidemiological studies of AEs, investigations of the cost of drug-related harm reported that ADEs extend hospital length of stay by 1.7-2.2 days at a cost of $2013-3244 (1997 USD) per patient (Bates, Spell et al. 1997, Classen, Pestotnik et al. 1997). UK-based estimates of the impact of ADE-related hospital admissions reported a cost of £1824 per admission, which, when extrapolated to population level, resulted in an excess cost burden to the National Health Service (NHS) of £466 million per annum (Pirmohamed, James et al. 2004). These findings diverted public attention to issues of healthcare and medication safety in the UK, catalysing research and policy efforts in this area. This was highlighted by the publication of "An Organisation with a Memory" in 2000, which shed light on the issues of preventable adverse events and errors across all healthcare sectors in the UK and called for an overhaul in the NHS to promote a new, system-oriented approach to

healthcare safety (Department of Health 2000). Amongst the findings reported by the Department of Health (DoH) in this report was an estimate of 314,000 adverse events which led to harm overall, with potentially as many as 60,000 adverse events leading to disability or death yearly in the UK, which was derived by extrapolating the results of previous studies of medication safety to the UK setting.

At the time of publication of "An Organisation with a Memory", UK-based data from the Medicines Control Agency listed 10,000 serious adverse drug events reported by healthcare professionals in the year 2000 (Department of Health 2000). However, it was recognised by the authors that the true incidence of ADEs in the UK is unknown due to poor levels of reporting of adverse events in routine care settings. The importance of medication harm was additionally emphasised in the 2001 DoH report on healthcare safety, which set an aim of reducing the number of serious medication errors by 40% (Department of Health 2001). The National Patient Safety Agency was set up with the specific aim of achieving this target in 2001.

1.2.2 Definitions used in literature

A number of concepts have been developed in the domain of medication safety in order to classify medication-related adverse events. The principal term used in pharmacology to represent accidental medication-related harm is "adverse drug reaction" (ADR). According to the World Health Organisation (WHO) definition an adverse drug reaction is an unintended effect of a medication which occurs at doses used in humans for prophylaxis, diagnosis or therapy of disease, or for modification of physiological function (WHO 1972). Within the nomenclature of pharmacovigilance, generally two types of adverse drug reaction can be distinguished: Type A reactions, which are dose-dependent and predictable and type B reactions, which are not dose dependent and usually unexpected. Type B reactions are also known as idiosyncratic drug reactions and are usually harder to prevent. This definition thus assumes that an ADR is an effect of the drug occurring during normal pharmacotherapy. However, this definition does not account for medication-related adverse events which occur due to error. Alternative definitions are found in the drug safety literature, such as one by Aronson et al who defined an ADR as an "appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product" (Aronson and Ferner 2005). However, this definition assumes that ADRs may occur under abnormal circumstances, including errors, which may not be helpful

in separating preventable and non-preventable events. Another definition by Edwards et al assumed the definition by Aronson et al, adding that an ADR additionally "…predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product." This alternative form assumes that an ADR is not to be treated as an isolated event and assigns responsibility to the healthcare professionals in charge of medication management to take appropriate action in case of its occurrence (Edwards and Aronson 2000).

The more encompassing use of the term ADR has led to reduced consistency of terminology across medication safety studies. A clear distinction must be made between harmful effects of medication occurring under normal circumstances and ones which occur as a result of poor quality care. Such a distinction is necessary in order to identify those events which have been caused by errors in order to prevent their occurrence in the future. For instance, a harmful effect which occurred at a dose higher than normal due to an error in prescribing or administration is not considered an ADR under the classic WHO definition or its variants used in literature. The term "adverse drug event" (ADE), defined as an injury resulting from medical intervention related to a drug, was introduced into the patient safety literature to fill this terminology gap. ADEs are composed of both ADRs (under the original WHO definition) and medication-related harm caused by an error, which are assumed to be preventable and are thus also known as preventable ADEs (pADEs) (Bates, Cullen et al. 1995). The term ADE was eagerly taken up by authors in the domain of medication-related harm as it reconciled research in pharmacovigilance, which mostly concerns itself with events occurring under normal care (ADRs) and medication safety, which studies medication errors. The remaining issue is the inconsistent use of the term ADE: in some studies it has been equated to ADRs (Classen, Pestotnik et al. 1997), whilst in others it was synonymous to medication errors (Andrews, Stocking et al. 1997), which complicates cross-study comparisons.

1.2.3 Errors in healthcare

Healthcare is a highly complex, dynamic system which involves an interplay of a multitude of factors and relationships, whose outcomes and consequences are uncertain (Lipsitz 2012). As in any other complex system which relies to a large degree on human decision-making, errors and lapses are likely to occur. A number of developments in recent years have increased the propensity for errors to occur in healthcare. The ageing population inevitably produces an increased demand for healthcare, in particular for treatment of chronic disease. An epidemiological study of multimorbidity in Scotland reported in 2012 that individuals classed as multimorbid (≥2 concomitant chronic conditions) made up a majority of the population aged 65 years and older (Barnett, Mercer et al. 2012). Multiple conditions require multiple parallel treatments, which produces an increased need for individualised care and increases the complexity of treatment. This inevitably translates into concomitant use of multiple medications, which is known as polypharmacy. Fifty-one percent of older patients (≥65 years old) across several European countries, including the UK, were co-prescribed 6 medications or more in a retrospective study carried out in 2005 (Fialova, Topinkova et al. 2005). This increases the likelihood that an error is made during one of the stages of medication management.

There are multiple ways of classifying human error: one approach is to identify whether it consists of wrong planning (planning error) or wrong execution of a correct plan (slip or lapse). Lapses are occasions in which the plan for the course of action was correct but the execution was wrong due to inattention or negligence, whilst planning errors are flawed from their conception, such as incorrect choice of drug due to insufficient knowledge of its potential adverse effects or insufficient information regarding the patient. Prescribing of 1g instead of 1mg of lorazepam or prescribing phenylephrine instead of fentanyl due to inattention are examples of lapses, whilst prescribing haloperidol in an elderly patient due to inadequate drug knowledge or potassium chloride in the presence of renal failure due to insufficient patient information can be considered errors in planning (Leape, Bates et al. 1995). Additionally, errors may be classified as errors of omission and commission: the former represent the failure to undertake a necessary action, such as adequate dose transcription; the latter represent inappropriate action on behalf of the healthcare professional, such as prescribing a wrong drug.

The theories of the cause of human error in professional settings include the person and system approaches. The patient-centred approach focuses on the

individuals who are responsible for errors, also known as active failures. Such failures in the context of medication therapy may include prescribing an inappropriate drug, wrong dosage, inattention during medication checks at the dispensing stage. The person-centred approach assumes that these errors occur due to carelessness and negligence arising from forgetfulness and lack of attention. The usual response to these "active failures" in practice involves focusing on individual culpability. The system approach recognises that human errors can happen under normal circumstances and views failures as the consequence of the shortcomings of the system, such as insufficient resources, poor management and inadequate organisational culture, which leads to low levels of motivation, lack of training and an atmosphere of mistrust. It has been recognised that the patient-centred approach is not sufficient in the healthcare sector due to the complex interplay of factors which determine outcomes in healthcare, including heterogeneity of therapies, workload and resources and multiple levels of decision-making (Reason 2000). Placing responsibility for the failings of the healthcare system on individuals inhibits reporting of adverse events by healthcare professionals who fear sanctions or dismissal. This encourages a culture of mistrust and prevents active learning from past mistakes, increasing the likelihood of the same errors reoccurring in the future. This is supported by evidence that hospitals with a relatively high culture of blame and fear of mistakes amongst staff experience a higher incidence of adverse events (Singer, Lin et al. 2009). The systems approach is geared towards issues in the structure of the organisation, such as inadequate education and training, excess workload or inadequate equipment, which precipitate latent failures and errors which occur systematically, irrespective of the people who execute the tasks on the ground.

1.2.4 Medication errors

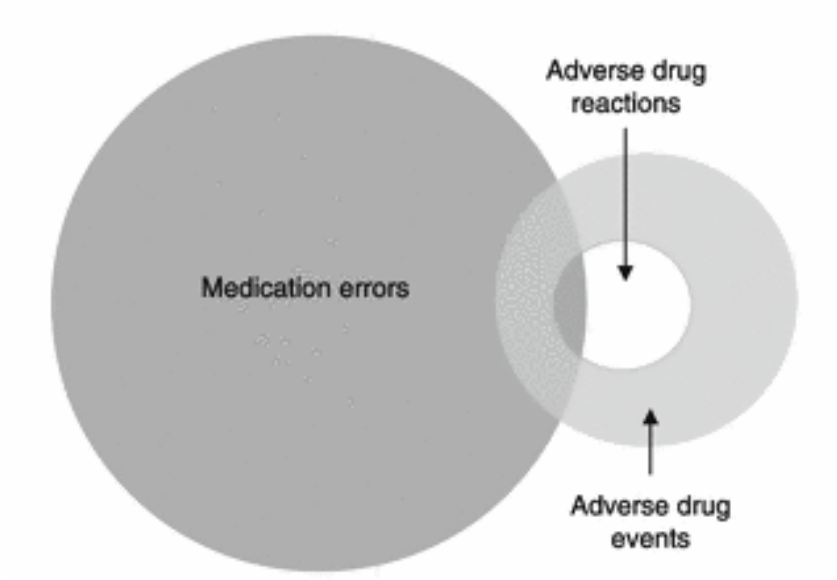
A medication error is defined as a preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of a healthcare professional, patient or consumer (Ferner RE 1999). An error may occur at any stage of the medication management process. A prescribing error may include wrong dose, inappropriate drug (such as benzodiazepine to treat delirium in an elderly patient) or failing to take into account hazardous drug-drug interactions (such as prescription of warfarin in combination with NSAID). A dispensing error may include the failure by a pharmacist of stopping an order for a hazardous prescription or dose when dispensing a drug to the patient. Drug administration errors may include administering a dose at the wrong time or missing a dose, incorrect dilution or wrong route of administration (e.g. intravenous rather than intramuscular injection). Failure to monitor a prescribed medicine in a way which would be considered acceptable in routine practice, such as not carrying out weekly monitoring of INR at the initiation of warfarin therapy, constitutes a monitoring error (Alldred, Standage et al. 2008).

A large proportion of ADEs in healthcare are caused by medication errors, providing a strong case for interventions to improve medication safety. Large cohort studies have shown that 28% of hospital-acquired (70/247) (Bates, Spell et al. 1997) and community-acquired (421/1523) (Gurwitz, Field et al. 2003) ADEs are preventable. The most common medication errors at the root of preventable ADEs occur at the prescribing (58.4%) and monitoring (60.8%) stages of medication management (Gurwitz, Field et al. 2003). Concentrating policy efforts on these areas may yield the best results in terms of reduction of the incidence of preventable ADEs.

Medication errors are much more ubiquitous in healthcare than ADEs. Prior evidence suggests that as many as 1 in 20 medication orders involve an error, with just 1 in 100 errors leading to actual patient harm (Bates, Boyle et al. 1995). The vast majority of medication errors do not have consequences for patient health (Bates 1996), although all medication errors have the potential to cause harm. This suggests that in order to study medication errors and their effect on patient outcomes, a large sample size would be required in order to identify rare pADEs, in combination with reliable methods for identifying errors and ADEs from observed patient records.

Fig. 1.1, reproduced from a diagram in a study by Krahenbuhl-Melcher (Krahenbuhl-Melcher, Schlienger et al. 2007) illustrates well the concepts explained above and the relationships between them:

Fig. 1.1. Relationship between the main concepts in medication safety



In summary, medication errors have the potential to produce adverse drug events, but in practice only a small proportion of errors progress to patient harm. Not all adverse drug events are caused by errors. Adverse drug reactions are a subset of adverse drug events and in the majority of cases of ADR are non-avoidable.

1.2.5 Definitions used in this thesis:

Table 1.1: Key definitions used in the thesis

Term	Definition	Source
Adverse drug event (ADE)	Injury resulting from medical intervention related to a drug	(Bates, Cullen et al. 1995)
Adverse drug reaction (ADR)	Unintended effect of a medication which occurs at doses used in humans for prophylaxis, diagnosis or therapy of disease, or for modification of physiological function	(WHO 1972)
Error	Failure to complete a planned action as intended or the use of an incorrect plan of action to achieve a given aim	(Kohn)
Medication error	Preventable event which may cause or lead to inappropriate medication use or patient harm while the medication is in the control of a healthcare professional, patient or consumer	(Ferner RE 1999)
Monitoring error	Failure to monitor a prescribed medicine in a way which would be considered acceptable in routine practice	(Alldred, Standage et al. 2008)

Table 1.1 denotes the key terms in medication safety and their definitions as used in this thesis. Adverse drug reaction was defined as an "unintended effect of a medication which occurs at doses used in humans for prophylaxis, diagnosis or therapy of disease, or for modification of physiological function", as per the WHO definition described in Section 1.2.2. An ADE was defined as an "injury resulting from a medical intervention related to a drug" (Bates, Cullen et al. 1995). The choice of definitions of ADR and ADE was justified by the desire to clearly separate those events which happen under normal therapeutic circumstances and those which were potentially caused by errors in the medication management process. Whilst it is impossible to preclude that an error may have precipitated an ADR under the WHO definition, the risk of harm may be assumed to be considerably lower if a drug is used under normal circumstances, compared to a situation where a drug is used at the wrong dose, circumstances or with

inadequate follow-up. The term ADE encompasses both ADRs and events caused by errors, which are referred to as preventable ADEs (pADEs).

An error was defined as a "failure to complete a planned action as intended or the use of an incorrect plan of action to achieve a given aim" (Kohn). A medication error was defined as a "failure in the treatment process involving a medication which leads to, or has potential to lead to, harm to the patient" (Ferner RE 1999). A monitoring error was defined using the formulation by Alldred et al (Alldred, Standage et al. 2008): failure to monitor a prescribed medicine in a way which would be considered acceptable in routine practice.

1.2.6 Why study medication safety in primary care?

The medication safety literature has focused on ADEs in the secondary care setting. A systematic review has shown that 6.7% of patients experience an ADE during a hospital stay, with 0.32% suffering a fatal ADE (Lazarou, Pomeranz et al. 1998). A number of factors facilitate drug safety research in the hospital setting: ease of recruitment and follow-up, opportunity to closely observe patients on a regular basis in a controlled setting, access to patient charts, availability of information from other staff members which may identify drug-related issues in patients. Pharmacoepidemiological research in the primary care setting is limited by loss to follow-up and fewer opportunities to consistently identify patient exposure and outcomes. However, with 4 in 5 medications prescribed in the primary care setting, there is a strong case for conducting more research on medication safety in the community. Moreover, there is evidence that community-acquired ADEs occur frequently, with a systematic review reporting a median incidence rate of 14.9 ADEs per 1000 patient-months, which translates into an ADE once in 5.6 years of therapy on average per patient (Thomsen, Winterstein et al. 2007), although a large amount of cross-study variation in estimates was reported (range 4-91 ADEs per 1000 patient-months), reflecting lack of consensus on the definitions of ADE and methods of identifying ADEs in the literature. These issues were discussed in more detail in the context of the current evidence on the incidence of ADEs in healthcare in Chapter 2 of this thesis.

1.3 Background to medication monitoring.

An effective course of therapy is one which achieves its intended health benefits in a patient whilst minimising the occurrence of unintended adverse effects. A process which captures the effects of therapy and guides treatment in order to improve the effectiveness of therapy is often useful in improving outcomes. For example, a course of antihypertensive therapy can be made more effective by measuring BP at periodic intervals and using this BP to adjust drug dosage; an anticoagulation treatment with warfarin could be made more effective and safe by minimising the risk of thrombosis or haemorrhage through periodic measurement of INR. In this sense, monitoring can be defined as the periodic measurement that guides the management of a chronic or recurrent condition (Glasziou, Irwig et al. 2005).

1.3.1 When is medication monitoring used?

Monitoring is a tool which can be used to guide treatment in order to achieve therapeutic goals and to simultaneously protect the patient from ADEs (Pirmohamed and Ferner 2003). Therapeutic outcomes may be improved through better treatment selection and titration, whilst the safety of treatment may be improved through measurement of a surrogate value with an established correlation with the risk of known ADEs (Glasziou, Irwig et al. 2005). Monitoring, when used appropriately, can identify changes related to ADEs before they have caused serious or permanent effects or withdrawal of therapy (Coleman, McDowell et al. 2010). This is a particularly useful property in cases where mild adverse events may be inevitable or extremely hard to prevent, but may develop into severe adverse events with a significant impact on the patient's health if the ADE is not identified at an early stage.

In many cases the harm potentially arising from foregone drug therapy outweighs the risk of harm from ADEs linked to the drug, thus it may be more beneficial to initiate the treatment with a high-risk medication in combination with regular monitoring, rather than foregoing treatment altogether (Steinman, Handler et al. 2011). This is supported by evidence that underuse of medications has a significant effect on patient mortality (Soumerai, McLaughlin et al. 1997, Havranek, Abrams et al. 1998). Monitoring allows the general practitioner (GP) to balance the risks and benefits throughout the course of therapy, making adjustments to dosage to maintain both its efficacy and safety (Nobili, Garattini et al. 2011). Monitoring of drug therapy may be carried out by the GP or by the patient with periodic oversight of the process by a healthcare professional. Self-

monitoring has been shown to be effective at improving outcomes and preventing adverse events from therapy in asthma (Gibson, Powell et al. 2003) and patients with chronic obstructive pulmonary disease (Dale, Connor et al. 2003). In the case of blood test monitoring, however, the process is usually carried out in the clinic. Some medications have a narrow therapeutic range and are toxic at high doses. For example, digoxin has been used in the treatment of congestive heart failure but requires careful titration to achieve efficacy and may require frequent therapeutic level monitoring to ensure safety (Matzuk, Shlomchik et al. 1991, Canas, Tanasijevic et al. 1999). In certain cases specific infrastructure is required to undertake monitoring. Anticoagulation clinics are more effective than standard primary care at preventing significant bleeding events from antithrombotic medications such as warfarin (Chiquette, Amato et al. 1998, Menzin, Boulanger et al. 2005).

1.3.2 Phases of medication monitoring

Monitoring of drug therapy can be subdivided into several distinct phases: pre-treatment, initial response, maintenance, re-establishing control and post-treatment (Glasziou, Irwig et al. 2005). Pre-treatment monitoring aims to measure baseline values of surrogate markers which are directly related to the subsequent safety of therapy, whilst testing shortly after initiation is used to measure response to therapy and to identify idiosyncratic adverse reactions. Pre-treatment and initial monitoring is explicitly recommended for some medications due to their safety profile, such as phenytoin, carbamazepine, phenobarbital, and valproic acid (Schoenenberger, Tanasijevic et al. 1995, Beach, Faich et al. 1998). Baseline and initiation monitoring is also recommended for certain cardiovascular medications, such as renin-angiotensin-aldosterone system (RAAS) agents (e.g. angiotensin-converting enzyme (ACE) inhibitors and angiotensin-II receptor blockers (ARBs)), amiodarone and loop diuretics in heart failure (North West 2002, British Thyroid Association 2006, London and South East Medicines Information Service SWMIS 2008). Maintenance monitoring, which is also referred to as regular or follow-up monitoring in the medication safety literature, is usually carried out at a lower frequency than initiation monitoring and is designed to a) ensure that therapeutic goals are achieved by periodically measuring the patient's response to the drug and b) identify ADEs before they have caused severe harm. For example, periodic testing of serum potassium and creatinine is undertaken during treatment with RAAS agents in order to identify hyperkalaemia and renal impairment and potentially avert severe harm (Raebel,

McClure et al. 2007). The frequency of testing in maintenance monitoring may change to reflect the condition of the patient or in order to re-establish control following failure to achieve therapeutic targets or an ADE. Post-treatment monitoring is carried out following cessation of therapy in cases where potential adverse effects of a medication may persist for a certain amount of time following cessation of therapy.

1.3.3 Choice of measurement

Monitoring should be used to target an intermediate outcome which is closely correlated with the onset of the ADE. A monitoring strategy should ideally help identify ADEs in a timely manner whilst minimising the signal-to-noise ratio (Glasziou 2008). However, short-term within-patient variation is high when initial monitoring is carried out at relatively short intervals. Regular monitoring over a longer interval can differentiate signals from random variation much more effectively and can thus reassure the physician that the drug therapy achieves adequate levels of efficacy and safety. Regular monitoring for the prevention of ADEs during established drug therapy may target the drug concentration (known as therapeutic drug monitoring (TDM)) or another clinical marker which is closely related to the risk of ADE. TDM aims to maintain the drug concentration between a lower and upper bound required to achieve maximum benefit at a given level of risk of ADE which is deemed acceptable, which is known as the therapeutic range. Statistical process control is used to identify systematic deviations from this range and control is re-established via dosage adjustment (Coleman, Ferner et al. 2006). TDM is routinely used to manage the risk of toxicity of drugs such as lithium, phenytoin and digoxin (Reynolds and Aronson 1993). Other medications require the measurement of a surrogate marker: examples include monitoring of thyroid function via thyroid-stimulating hormone (TSH) levels in amiodarone therapy or platelet count for the prevention of methotrexate-induced thrombocytopenia.

1.3.4 Timing of monitoring

An effective monitoring strategy involves periodic testing at a frequency which allows the observation of an abnormal result as early as possible, giving enough time to distinguish the signal as genuine and for corrective action to be taken to avoid harm. The performance of the monitoring strategy is therefore directly linked to the length of time that typically elapses between the change in the surrogate measure and actual harm (Coleman, Ferner et al. 2006). The longer this "signal interval", the more time there is to take action and the less frequent the monitoring needs to be. In some cases, monitoring too frequently may produce spurious results due to significant short-term variation and measurement error, which may lead to erroneous treatment decisions if such results are misinterpreted as a change in the underlying response to treatment or risk of ADE (Bell, Irwig et al. 2008). For this reason, excessive frequency of monitoring may lead to suboptimal treatment outcomes (Schoenenberger, Tanasijevic et al. 1995). If used appropriately, frequent monitoring may be effective at preventing rapid-onset ADEs: weekly monitoring of full blood count during clozapine therapy has been shown to dramatically reduce the incidence of agranulocytosis and associated mortality (Honigfeld 1996).

In medications which pose the risk of rapid-onset ADE this interval is very short, requiring intensive monitoring. For example, international normalised ratio (INR) may be measured daily in order to avert severe haemorrhaging from warfarin until the therapeutic range has been reached (Hirsh 1995). Monitoring in primary care at very short intervals may be unfeasible if the distance to the clinic is long, patient has limited mobility or the laboratory is unable to generate results fast enough. Symptom-based monitoring and self-monitoring may be more effective in such cases. For example, peak expiratory flow monitoring of paediatric asthma in patients with symptoms has been shown to produce improved asthma severity scores and reduced healthcare utilisation compared to daily monitoring of all patients in a randomised study (Yoos, Kitzman et al. 2002). Alternatively, close monitoring may be carried out in the inpatient hospital setting, albeit at much greater cost compared with monitoring in primary care due to the need for inpatient admission.

In certain cases regular monitoring may be ineffective due to the lack of a marker which identifies the onset of a rapidly-developing ADE, such as idiosyncratic liver damage caused by statins. Previously published guidelines recommended liver function tests (LFTs) once in 12 months during the course of therapy with statins. However, such long intervals are unlikely to prevent hepatic damage due to its

unpredictability and rapid onset. Monitoring frequently enough to allow corrective action to be taken is unlikely to be justified due to the low risk of hepatic ADEs in statin therapy. This was recognised in a recent NICE guideline, which determined that LFT should be measured at initiation, 3 months and 12 months, but no further regular monitoring of LFT should be conducted (National Institute for Health and Care Excellence 2014).

Monitoring is occasionally of limited use for preventing slow-onset ADEs, such as corticosteroid-induced osteoporosis. Around 2 in 5 patients on long-term corticosteroid therapy suffer observable bone loss and the standard care involves osteoporosis prevention therapy in patients prescribed a high dosage of the drug and/or in the presence of risk factors, such as co-prescription with another drug linked to bone loss, which is more effective at prevention of the ADE than regular monitoring (Hougardy, Peterson et al. 2000).

Regular monitoring may be ineffective if the ADE is irreversible, such as chloramphenicol-induced aplastic anaemia (Feder, Osier et al. 1981, Horler AR 1981). As a result, little opportunity for corrective action exists and discontinuation of therapy is the only option once the ADE is detected. In this case appropriate risk profiling prior to the start of therapy in order to avoid prescribing the drug to high-risk patients may be preferable to regular monitoring of all patients prescribed these medications.

In summary, medication monitoring may improve the safety of therapy with certain medications which represent a risk of preventable harm to the patient. However, a monitoring strategy may only be effective under a specific set of conditions. These include, but are not restricted to, the presence of a measurable marker which is closely linked to the ADE which it is aimed at preventing, availability of a test which may identify changes in this marker characterised by a high signal-to-noise ratio, preventability and/or ameliorability of the target ADE and sufficient time to act upon a developing ADE before severe harm occurs. The economic rationale for the application of monitoring in the absence of benefit in terms of reduced ADEs is weak as the cost of carrying out the tests will not be justified in terms of improved patient health (or avoided health loss) or reduction in the cost of treatment of ADEs.

Monitoring may not be considered as a stand-alone strategy which, on its own, will improve medication safety. In order to get the most out of regular monitoring, it must be incorporated into an ongoing, systematic medication

management process which adequately manages risk and responds to changes in the patient's treatment needs.

1.3.5 Monitoring errors

The monitoring of drug therapy is one of the stages of the medication management process where an error may lead to ADE and potential harm to the patient. Monitoring errors have previously been found to be one of the leading causes of community-acquired ADEs leading to hospital admission, contributing to between 21-42% of hospitalisations based on a systematic review of ADE studies (Howard, Avery et al. 2007). Another systematic review found that monitoring errors were the second most common reason for all community-acquired ADEs (36%, based on Gurwitz et al (Gurwitz, Field et al. 2003)). The proportion was even higher for severe events requiring hospital admission (45%, range 22-70%) (Thomsen, Winterstein et al. 2007).

Failures in monitoring may occur after a required test has been carried out if the physician does not respond to abnormal test results in a timely manner, causing a delay in treatment. Previous studies investigating this issue have found that 27% of abnormal laboratory test results were acted upon 5 or more hours following communication of results by laboratory to the hospital (Kuperman, Boyle et al. 1998). Another US-based study reported that 5.5% of ADEs may have been prevented through improved electronic reporting of laboratory tests, reducing the delay between obtaining an abnormal test result and taking corrective action (Bates, O'Neil et al. 1994).

1.4 Aims of programme of research

Regular monitoring during the course of therapy with ACEI, loop diuretics and amiodarone is designed to identify ADEs at an early stage, prevent harm to patient health and optimise the resources spent in the management of medication therapy. However, the hypothesised reduction in drug-related harm resulting from regular monitoring in these particular high-risk medications has never been rigorously tested. Several studies reported poor compliance to the frequency of testing prescribed within guidelines (these studies are described in more depth in Chapter 3), but there is a lack of evidence of the effect of monitoring errors on patient health, cost from the perspective of the NHS and patient utility.

The effect of monitoring frequency on the probability of occurrence of ADE needs to be assessed by observing current practice involving a number of processes of care: prescription records in order to identify patients who require regular monitoring, laboratory tests in order to observe which monitoring regimen they have been assigned to by their GP and diagnoses of ADEs occurring during the period of therapy. In order to assess the economic rationale for regular monitoring, information on the cost of monitoring, cost of treatment of ADEs and the impact of ADEs on utility need to be obtained. The overarching aim for this programme of research was to gain an understanding of the current practice of monitoring, the effect of monitoring frequency on the risk of ADE and the economic rationale for monitoring involving a subset of cardiovascular medications which pose a risk of ADE to patients and are listed in guidelines for regular monitoring in UK primary care.

The following specific aims were set for this programme of research:

Aim 1: Investigate the compliance with guidelines for regular monitoring of serum potassium, sodium and creatinine during ACE inhibitor and loop diuretic therapy and thyroid function during amiodarone therapy between 2003 and 2012 in primary care in England.

Aim 2: Investigate the effect of regular monitoring of ACE inhibitors, loop diuretics and amiodarone in primary care on the risk of specific ADEs linked to these medications during the study period.

Aim 3: Evaluate the cost-effectiveness of different frequencies of regular monitoring of thyroid function in amiodarone therapy.

1.5 Why use observational research to study medication monitoring?

The most basic definition of patient safety is the freedom from accidental injury (Kohn 1999). Research in patient safety concerns itself with estimating the incidence of adverse events in healthcare, investigating their causes and evaluating interventions to prevent such events from occurring in the future. The domain of medication safety focuses specifically on patient safety in the use of medications.

This study aimed to systematically assess the current practice of monitoring, the effect of regular monitoring on the risk profile of the medications included in the study and the economic effect of monitoring in UK primary care. The best way of achieving these aims was by observing what happens in practice and deducing causal relationships through inferential methods. A quantitative, data-driven approach was deemed to be the best suited method to achieve these aims.

With 98% of the population registered with a GP in England and ubiquitous use of electronic health records (EHRs), the UK primary care system provides an ideal setting for pharmacoepidemiological research (Strom 2006). The availability of systematically recorded, validated source of nationwide prescription, laboratory test and diagnostic data collected in large research databases enables the investigator to apply analytical and statistical methods on real-life data to assess causal relationships between medications and ADEs and evaluate interventions to improve medication safety, such as monitoring. This study will use two large databases: the Clinical Practice Research Datalink (CPRD) and Hospital Episode Statistics (HES) to achieve the research aims set out in Section 1.4. A description of how this data was extracted and used in the analyses can be found in Chapter 4 of this thesis.

An important limitation of studies using observational data is potential confounding arising from non-random selection. Direct comparisons of groups of patients on different treatments may produce biased results as treatment assignment may be dependent on other factors. For instance, patients may be monitored more frequently by GPs if they are deemed to be at higher risk of suffering an ADE due to age, health, presence of certain comorbidities or co-prescribed treatments which may predispose them to drug-related harm. In order to isolate the causal effect of monitoring on the risk of ADE and study outcomes, the investigator needs to make sure that the patient groups are comparable. A range of statistical and analytical tools exist which may be used to ensure this.

This issue in the context of the analyses described in this thesis and methods for dealing with confounding are described in more detail in Chapter 5 of this thesis.

1.6 Why carry out economic analyses of medication monitoring?

Apart from potentially causing patient morbidity and mortality, adverse drug events impose economic consequences on providers, patients and society as a whole. Additional treatment to mitigate the effects of drug-related harm involves the use of scarce healthcare resources. This includes physician time, excess hospital stays, outpatient care, medications and other costs which need to be borne by either the patient or a third-party payer. This represents an opportunity cost as the resources devoted in response to ADEs may have been spent more effectively on other goods or services to maximise individual utility (in cases when the cost is borne by the patient) or invested in other parts of the healthcare system to maximise societal welfare (in cases of a state payer system). In the context of UK healthcare, this opportunity cost is borne by the NHS, and, ultimately, the taxpayer who funds it. Additional costs may be incurred by the patient through lost earnings due to time taken off work or family due to the need for additional informal care. Costs borne by society potentially include loss in productivity and additional costs linked to disability later in life. The loss of health linked to an ADE may lead to a reduction in life expectancy and number of years spent in good health. These are strong economic arguments for intervening in order to improve the safety of medication use.

The hypothesised benefits of monitoring as a strategy to prevent ADEs are reduced cost of treatment and reduced losses of life-years and utility. Monitoring naturally entails a cost of carrying out laboratory tests and physician time in order to interpret test results, act upon abnormal results and consult the patient. The current state of knowledge on the efficacy of monitoring and its economic effect is poor, and it is insufficient to claim with any certainty that the cost of monitoring at regular intervals during therapy is justified in terms of the aforementioned economic benefits. Only two prior studies have assessed the cost-effectiveness of strategies to improve monitoring of ACE inhibitors, diuretics and amiodarone; however, no prior research has been conducted to establish the value for money of regular monitoring as a current "best practice" option.

The most commonly used methods in economic evaluation of interventions in healthcare include cost-benefit analysis (CBA) and cost-effectiveness analysis

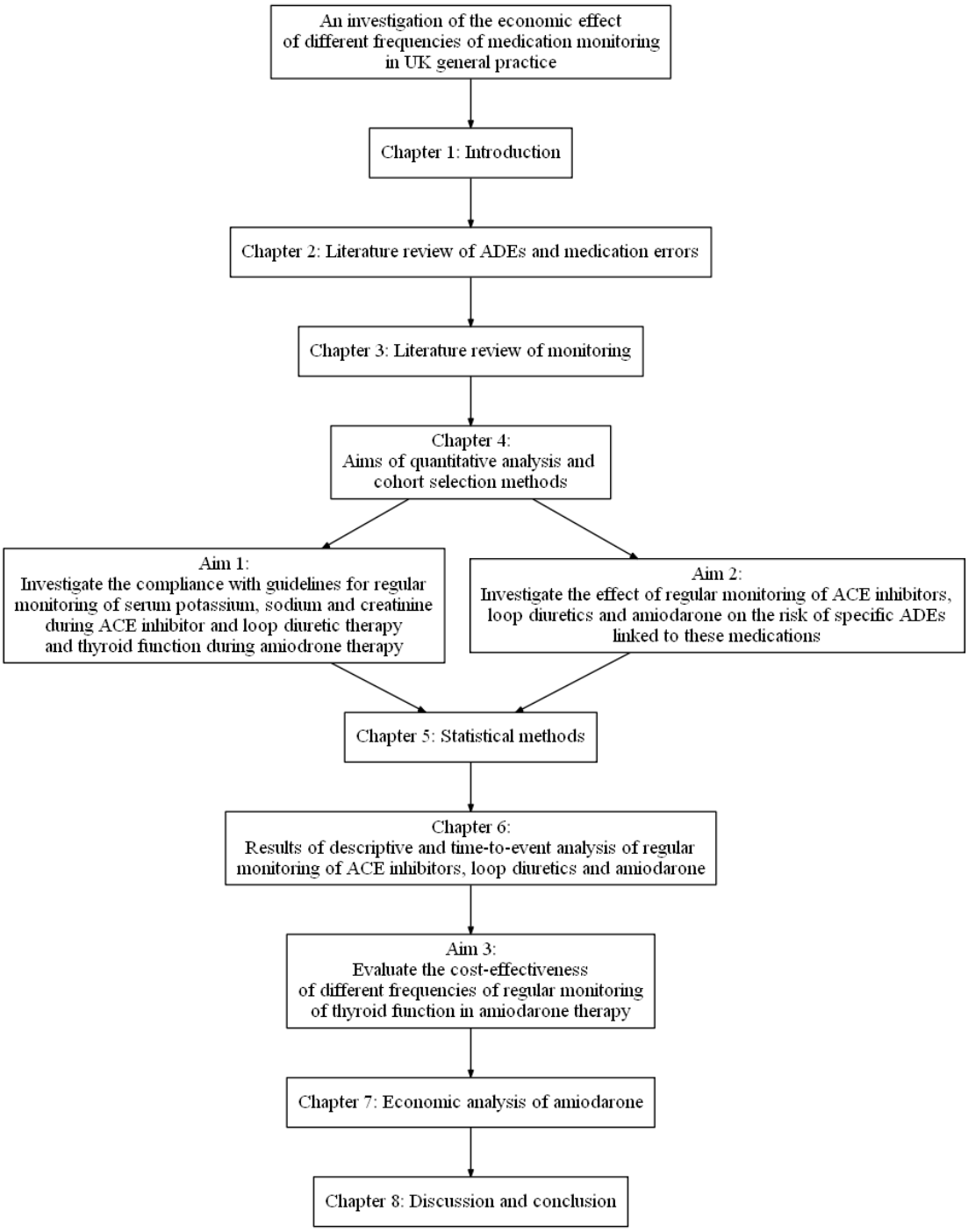
(CEA). Cost-benefit analysis involves the valuation of the benefits and costs of each policy option in monetary terms and selecting the option with the highest net monetary benefit (NMB) to society. Difficulty in attributing a monetary value to health outcomes has led to the development of CEA as an alternative method, which involves comparison in terms of cost per unit of outcome achieved. This requires the investigator to quantify the outcomes of each alternative intervention in terms of a common outcome, such as life-years gained, number of medication errors reduced or a surrogate clinical outcome. Cost-utility analysis (CUA), as a subset of CEA, incorporates individual preferences over health outcomes by using a measure of individual utility, such as quality-adjusted life-years (QALYs) (Drummond 1997). Cost-effectiveness analysis is the method of choice in many economic evaluations as it can be easily incorporated into practical decision analysis. It relies on accurate estimates of cost and utility, which may not necessarily be available in under-researched areas of healthcare, such as medication monitoring. In such cases one-way and probabilistic sensitivity analyses need to be carried out in order to estimate the impact of uncertainty in cost and utility generated by each alternative strategy on the choice of optimal strategy.

A potential limitation of economic evaluation in medication safety research is the fact that standard methods for economic evaluation involving health-maximisation may be insensitive to interventions in healthcare safety, where significant non-health patient and societal benefits may be present. These may include reduction in uncertainty and dread associated with ADEs and increased trust in the healthcare system, and are difficult to quantify and incorporate into a standard CEA analysis involving a decision problem (Steuten and Buxton 2010). The methods of CEA employed in this study and a detailed discussion of the applicability of these methods in the context of this research are discussed in Chapter 7.

1.7 Outline of thesis

This thesis, split into 8 chapters, presents a programme of original research carried out in order to achieve the aims of research set out in Section 1.4. Chapters 2 and 3 include reviews of literature on ADEs, medication errors and monitoring aimed at understanding the current evidence in this area and identifying the main gaps in the knowledge of the practice of monitoring and its effect on medication safety. Chapter 4 describes the methods of observational research using electronic health records (EHRs) from the Clinical Practice Research Datalink and Hospital Episode Statistics databases to study monitoring and ADEs in UK primary care. Chapter 5 outlines the design of the statistical analysis and methods for dealing with confounding. Chapter 6 includes the results of the descriptive analysis of monitoring of the study medications using CPRD laboratory test records between 2003 and 2012 and time-to-event analysis designed to evaluate the effectiveness of different frequencies of regular monitoring in terms of the rate of specific ADEs linked to the use of these medications. Chapter 7 describes the methods and results of an economic analysis of amiodarone monitoring using model-based cost-effectiveness analysis. Chapter 8 combines the findings from all the studies included in this thesis, draws conclusions from the results and informs recommendations for future direction of policy and research in medication monitoring in primary care.

Fig. 1.2. Outline of content of thesis



1.8 Potential impact of the research described in this thesis

The first stage of the research outlined in this thesis involves an investigation of the current practice of monitoring and compliance with guidelines for the monitoring of ACE inhibitors, diuretics and amiodarone, fitting into the current body of evidence on the practice of monitoring involving these commonly prescribed medications in primary care. The subsequent analyses in this programme of research advance the knowledge in the area of monitoring of high-risk medications in the following directions:

i) The data on monitoring derived from primary care EHRs was linked to diagnostic records on CPRD and HES databases in order to investigate the effectiveness of different frequencies of monitoring long-term therapy with ACE inhibitors, loop diuretics and amiodarone, which has hitherto not been explored in the UK context

ii) In the case of amiodarone monitoring, this study contributed to the evidence base on the cost-effectiveness of alternative strategies of monitoring this drug, thus assessing the economic rationale for regularly monitoring patients on amiodarone in UK primary care

As a whole, this programme of research provides much-needed evidence on the potential value of regular monitoring of three medications which are commonly associated with preventable iatrogenic harm in the UK. In light of the low quality of evidence supporting current recommendations for regular monitoring of the medications included in this study, the new evidence arising from the research outlined in this thesis has the potential to contribute to the decision-making process of policy-makers in English healthcare and the formation of a new set of improved, evidence-based guidelines for the conduct of regular monitoring of this subset of high-risk medications in UK primary care.

Given the paucity of previously conducted research on both the effectiveness and cost-effectiveness of regular monitoring in the UK context, this study will contribute towards the development of methodology in this under-researched area. In particular, this study will tackle the challenges involving the use of non-randomised retrospectively collected data to carry out an analysis of monitoring of specific medications and the use of this data source to populate a model-based cost-effectiveness analysis, which will provide a useful insight into the strengths, limitations and potential opportunities of this approach for future investigations in medication safety.

Chapter 2

Literature review of adverse drug events and medication errors in healthcare

2.1 Introduction and aims of the chapter

This chapter describes the issue of adverse events and errors in healthcare, with a particular focus on medication-related harm. The purpose of this chapter is to summarise the key aspects of medication safety and review the current evidence relevant to this project. The aims of the chapter are as follows:

1. Review the literature on medication safety and evidence on the incidence of adverse events in the UK and other countries.
2. Review studies which have examined the incidence of adverse events linked to cardiovascular medications.
3. Examine the economic effect of medication errors.
4. Discuss the findings presented in this chapter to place this project in the context of the current evidence base. Outline the rationale for conducting further research on monitoring of medications in UK primary care.

2.2 Background of healthcare safety

The turn of the 21st century has seen a significant shift in the attitude of policy-makers towards the problem of failures in healthcare safety both in the UK and abroad. One of the contributing factors to this shift has been the publication of an observational study of harm occurring in US hospitals involving chart review of over 30,000 patients, which was conducted as part of the Harvard Medical Practice Study in 1991. The study found that adverse events resulting in severe harm occurred in 3.7% of hospitalised patients and ADEs were the most common event, contributing to 1 in 5 cases (Leape, Brennan et al. 1991). The policy reaction to this evidence was the publication of extensive reports on the problem of preventable harm in healthcare, such as "To err is human" (Kohn 1999) and "Organisation with a memory" (Department of Health 2000). These reports had the purpose of setting in motion a reconfiguration of organisational cultures

within the respective healthcare systems in order to improve safety of medical care. In particular, "Organisation with a memory" highlighted the major issue of iatrogenic harm in NHS hospitals, citing the work of Vincent et al who found that around 10% of inpatient episodes in NHS hospitals involved harmful adverse events, with around half preventable in nature. At the time of publication, this finding was in line with reports from literature on iatrogenic harm affecting 4%-17% of patients in hospital (Vincent, Taylor-Adams et al. 1998). This figure was extrapolated to a nationwide estimate of 850,000 adverse events in the UK hospital setting yearly by the authors of the report, with around half of the events judged to be preventable in nature, although it is not clear whether the results of the original study are generalizable to an entire sector of healthcare. The study additionally attempted to extrapolate the 1991 results of the Harvard Medical Practice Study (Leape, Brennan et al. 1991) to the UK setting, arriving at an estimate of 314,000 adverse events which lead to harm overall, with potentially as many as 60,000 adverse events leading to disability or death yearly (Department of Health 2000).

The recognition of the scale of the problem of errors and adverse events in healthcare has prompted policy shifts in other countries. The US Institute of Medicine has called for an overhaul of the approach to dealing with medical errors following the publication of "To err is human" by Kohn et al (Kohn 1999), which estimated that between 44,000 and 98,000 patients die as a result of medical error in US hospitals each year through extrapolation of the Harvard Medical Practice Study results described above and a similar study carried out in the US states of Utah and Colorado a decade later (Thomas and Brennan 2000). This follow-up study was carried out to investigate whether healthcare safety in hospitals has changed since the publication of the Harvard study. The authors investigated 15,000 randomly sampled hospital admissions and found an incidence rate of adverse events of 2.9%, with approximately 30% of the events occurring due to negligence. ADEs contributed to 1 in 5 events, with 53% caused by error and 17.4% of ADEs involved cardiovascular medications. The link of individual adverse events to errors was determined by physicians and root causes determined from patient records. By extrapolating their findings to the US population, the authors concluded that medical error in hospitals was a leading cause of mortality with an estimated 25,000 deaths per year in the year 2000, ahead of well-publicised causes such as car accidents and breast cancer. However, such extrapolations of observed incidence to population-wide estimates require careful interpretation as the characteristics and quality of care provided

within the facilities included in the study may not necessarily be representative of standard care delivered to the population.

In the years following the publication of "To err is human" and "Organisation with a memory" adverse drug events, and adverse events in general, have received increased attention in healthcare research. This increased the number and quality of studies of the incidence, preventability and causes of ADEs in different sectors of healthcare.

Previously conducted reviews of literature on adverse events in healthcare have shown a large variation in reported estimates of incidence across studies. A review of studies of AEs across all clinical settings between 1995 and 2007 found a range of 0.1 to 65.4% for the proportion of hospital episodes with an AE, with variation in estimates negatively related to study size (Lessing, Schmitz et al. 2010). This is intuitive, as estimated incidence from larger studies is more likely to be closer to the true incidence. No other factor, including study type, setting or method of collection explained the variation, suggesting that no single study design or identification method was clearly advantageous in this research area.

In order to examine the evidence relating to the incidence, nature and causal factors of ADEs in the context of this study, a literature review was performed in order to identify relevant studies on the subject published from 1990 to 2014. The review focused on studies of ADEs and medication errors over the past 25 years in a chronological order. The review was structured into sections according to study setting: first, it examined studies of ADEs occurring in hospitals, which was the most common study setting found in the review. Following this, the review focused on ADEs acquired in primary care, given that most prescribing of medications occurs outside the hospital in the community. This part of the review focused on both those ADEs which were acquired and treated in primary care and community-acquired ADEs which led to an emergency department visit or hospitalisation.

2.3 Literature review of adverse drug events

The search was conducted using a Medical Subjects Heading (MeSH) search builder by applying the following terms: 'adverse drug events'; 'adverse drug reactions; 'adverse drug effects'; 'medication errors'. For the analysis of ADEs involving cardiovascular medications included in this study, additional terms included 'angiotensin converting enzyme inhibitor'; 'loop diuretic'; 'thiazide diuretic'; 'amiodarone'; 'cardiovascular'. For the analysis of economic and costing studies of ADE, the following were added: 'cost'; 'economic'; 'cost-effectiveness'; 'budget impact'. The terms were used alone or in combination in order to find studies from relevant areas. A manual search of abstracts was carried out and for studies which were deemed relevant the full text was obtained in order to extract information on study results and methods used to ascertain incidence, characteristics and cost of ADEs. Systematic and meta-analyses were not included as individual studies and were instead summarised at the beginning of the review.

2.3.1 Summary of selected studies

The review identified 20 studies of the incidence of ADE which satisfied the inclusion criteria. This included 4 systematic reviews, 8 retrospective cohort studies, 6 which employed a prospective cohort design and 2 case-control studies published between 1992 and 2010. The review identified 9 studies of hospital-acquired ADE and 11 studies of community-acquired ADE. The review also identified 7 studies of the cost of ADE. A summary of the studies included in the review can be found in Table 2.1.

2.3.2 Systematic reviews

A systematic review in the hospital setting between 1990 and 2005 showed a range of estimates between 0.2% and 60.7% for the proportion of hospital episodes with an ADE (Krahenbuhl-Melcher, Schlienger et al. 2007). The authors attributed the wide variation in published estimates to differences in methods used to identify ADEs, although significant variation remained within studies grouped by identification methods: studies which used spontaneous reporting typically found much lower estimates of incidence (median 2.2%, range 0.2-7%), compared to chart review (median 13%, range 0.59-60.7%).

The largest meta-analysis of prospective studies of hospital-acquired ADRs and hospitalisations resulting from ADRs published between 1966 and 1996 reported that hospital-acquired ADRs occurred in 2.1% of episodes (Lazarou, Pomeranz et al. 1998). The proportion of hospital admissions caused by an ADR was 4.7%. The meta-analysis included 39 studies, applied a rigorous search criteria which included only prospective studies specifically involving ADRs, in order to provide a reliable estimate of drug-related harm in US secondary care. However, the authors did not discriminate between non-preventable ADRs and those caused by error, which is necessary in order to determine whether medication safety may be improved though better quality care. The authors noted that variation in methodology of data collection and identification of ADRs across individual studies impacted on the internal validity of their meta-analysis.

Other reviews focused on hospital admissions caused by community-acquired ADEs. A systematic review of studies of ADE-related hospital admission found a median of 3.7% of admissions are caused by an ADE (range 1.4% to 15.4%) (Howard, Avery et al. 2007). The medications which were most commonly linked to ADE-related admissions were NSAIDs, diuretics and antiplatelets. Prescribing and monitoring errors were the cause of 31% and 22% of admissions, respectively.

A meta-analysis by Beijer and colleagues reported that between 0.2% and 41% of hospitalisations were caused by ADEs, with a weighted average of 4.8%. The authors noted a clear negative relationship between study size and the estimated proportion of ADE-related hospitalisations. The link was attributed to differences in methods of identifying the causal relationship between admission and drug, which is more likely to be reliable in larger, well-designed studies, as well as publication bias, as smaller studies which demonstrate a stronger link between medication safety and hospital admissions were more likely to be published. Restricting to studies with a sample size of over 5000 patients, the estimate of the proportion of hospitalisations caused by ADEs ranged between 1% and 8% (Beijer and de Blaey 2002).

2.3.3 Summary of individual studies

Several studies were identified in this review which were not included in the systematic reviews above or were published at a later date. There was a wide range in sample size, varying from 186 participants in chart review studies in a single facility to reviews of electronic records of over 50 million patients using an

administrative database. The proportion of hospital admissions which were attributed to an ADE varied a lot across studies, with individual estimates likely affected by study design. There was a clear link between sample size and the estimate: smaller studies which had access to patient charts reported a much larger proportion of ADE-related hospitalisations (11%-25%), whilst larger studies which employed data mining of electronic health records produced much lower estimates (0.5%-6.5%). In addition, there was a large variation in the estimated proportion of preventable ADEs (11%-72%), although not all studies evaluated preventability. Proportion of severe ADEs ranged from 8% to 38%, although there was no consistency in the definition of severe ADE. Two studies defined severe ADEs leading to death (Pirmohamed, James et al. 2004) (Wu, Jen et al. 2010). One study defined severe ADEs as those leading to disability or death (Leendertse, Egberts et al. 2008). The summary of study results is listed in Table 2.2. Details of the design of studies in the review are listed in Table 2.3.

2.3.4 ADEs linked to cardiovascular medications.

Cardiovascular medications, such as angiotensin converting enzyme inhibitors, angiotensin-II receptor blockers, diuretics and amiodarone are commonly associated with adverse drug events (Gandhi, Weingart et al. 2003, Gurwitz, Field et al. 2003, Howard, Avery et al. 2003, Samoy, Zed et al. 2006). This section offers a description of the safety profile of each of these medications.

ACE inhibitors and ARBs

Medications acting on the renin-angiotensin-aldosterone system (RAAS) such as ACE inhibitors are used in routine clinical practice for the treatment of hypertension, heart failure and for the improvement of outcomes following myocardial infarction. The use of RAAS agents and loop diuretics has seen a significant rise in the past decade with the implementation of a pay-for-performance scheme in UK primary care in 2004 (Quality and Outcomes Framework (QOF) (Department of Health 2004), which included RAAS agents and loop diuretics as first-line treatment for hypertension and other cardiovascular conditions, such heart failure and coronary heart disease (National Health Service 2007). Long-term management of cardiovascular conditions is generally carried out in the community; as a result these medications are prescribed on a very large scale in primary care. As well as providing a clear clinical benefit to patients

Table 2.1: Summary of the results of studies of hospital-acquired ADE

Author	Year	N	Proportion of hospital episodes with ADE	Preventable %	Severe/life-threatening %
Lindley[1]	1992	416	27%	Not assessed	Not assessed
Bates[2]	1995	4031	6.5%	28%	30% severe, 13% life-threatening
Smith[3]	1996	20695	6.9%	N/A	Not assessed
Bates[4]	1997	4108	4.6%	28%	30% severe, 13% life-threatening
Classen[5]	1997	91574	2.4%	N/A	6%
Cox[6]	2001	21635	0.06%	N/A	N/A
Davies[7]	2009	3322	15.8%	53%	N/A

with hypertension and cardiovascular problems, ACE inhibitors present a risk of unintended ADEs. In particular, these drugs cause electrolyte imbalances (abnormal levels of serum potassium, sodium and other substances) through different mechanisms: the action of ACE inhibitors on the RAAS causes a rise in the net retention of potassium, which causes hyperkalaemia. This risk is elevated in some patient groups, such as patients with renal impairment and diabetes, which may impair elimination of sodium or potassium, exacerbating the risk of hyperkalaemia (Ahuja, Freeman et al. 2000, Maddirala, Khan et al. 2008). A recent US-based study used a cohort of 60464 patients prescribed RAAS agents in a primary care clinic to estimate that 0.59% (SE 0.0003) of patients newly prescribed a RAAS agent who did not receive monitoring were hospitalised for hyperkalaemia (Smith, Raebel et al. 2011).

[1] Lindley, Tully et al. 1992
[2] Bates, Cullen et al. 1995
[3] Smith, Raebel et al. 2011
[4] Bates, Spell et al. 1997
[5] Classen, Pestotnik et al. 1997
[6] Cox, Anton et al. 2001
[7] Davies, Green et al. 2009

Table 2.2: Summary of the results of studies of community-acquired ADE

Author	Year	N	Proportion of admissions linked to ADE	Preventable %	Severe/ life-threatening %
Gandhi[8]	2000	2248	2.8%	13%	13%
Howard[9]	2006	Systematic review of 13 studies	Median 3.7% (range 1%-15%)	Only assessed preventable ADEs	N/A
Gandhi[10]	2003	661	25%	11%	13%
Gurwitz[11]	2003	27617	50.1 per 1000 patient-years	13.8 per 1000 patient-years	38%
Pirmohamed[12]	2004	18820	6.5%	72%	2.3% led to death
Morris[13]	2004	50,000	1%	Only assessed preventable ADEs	N/A
Passarelli[14]	2005	186	11.3%	23%	36.7%
Samoy[15]	2006	565	24.1%	72.1%	8.1%
Leendertse[16]	2008	12793	5.6%	46.5%	15.6% leading to disability or death
Wu[17]	2010	≈60000000	0.5%	N/A	4.7% led to death

[8] Gandhi, Burstin et al. 2000
[9] Howard, Avery et al. 2007
[10] Gandhi, Weingart et al. 2003
[11] Gurwitz, Field et al. 2003
[12] Pirmohamed, James et al. 2004
[13] Morris, Rodgers et al. 2004
[14] Passarelli, Jacob-Filho et al. 2005
[15] Samoy, Zed et al. 2006
[16] Leendertse at al. 2008
[17] Wu, Jen et al. 2010

This is an important concern during therapy because RAAS agents are frequently used in patients with chronic kidney disease (CKD) and diabetes to treat renal disease. Regular monitoring of ACEI is suggested in a number of national guidelines, particularly in the presence of pre-existing renal dysfunction and heart failure (National Institute for Health and Care Excellence 2011, British National Formulary (online) 2015), which are discussed in further detail in Section 3.4.

Diuretics

Diuretics reduce blood pressure by lowering fluid levels and inhibiting chloride and sodium reabsorption (Padilla, Armas-Hernandez et al. 2007). There are several types of diuretic, including loop, thiazide and potassium-sparing varieties. One of the side effects of loop and thiazide diuretics is the depletion of serum potassium and sodium which may result in hypokalaemia and hyponatraemia (Fichman, Vorherr et al. 1971). Similarly to ACE inhibitors and ARBs, this effect is usually benign if appropriate and timely action is taken in response to abnormal test results. Abnormally low levels of serum potassium or sodium pose a risk of severe complications including potentially fatal cardiac arrhythmias. Elderly patients are at particular risk due to a higher prevalence of chronic kidney disease and diabetes, predisposing this patient group to electrolyte abnormalities. Hyponatraemia can also cause impaired gait which, in combination with frailty places elderly patients at particularly high risk of falls (Renneboog, Musch et al. 2006) and fractures, even when controlling for the presence of osteoporosis (Kinsella, Moran et al. 2010). The largest study of ADEs linked to diuretics during established therapy (n=76264) reported 50 cases of drug-related hospitalisation caused by diuretics (0.066% of exposed patients) (Schneeweiss, Hasford et al. 2002).

Amiodarone

Amiodarone, used to treat cardiac arrhythmias, may produce adverse effects to patient health due to the toxic effects arising from its chemical composition. The high iodine content (37.3%), with 10% of content released as free iodine can cause disorders of thyroid function (Loh 2000) by inhibiting triiodothyronine (T3) and thyroxine (T4) (Toft and Boon 2000). This adverse effect is common, with previously conducted cohort studies reporting between 16-37% of patients with thyroid dysfunction during amiodarone therapy (Martino, Safran et al. 1984, Raeder, Podrid et al. 1985, Albert, Alves et al. 1987, Shukla, Jowett et al. 1994,

Tsadok, Jackevicius et al. 2011). Toxic effects of amiodarone have been shown to additionally produce hepatic, respiratory and cardiac adverse effects (Stelfox, Ahmed et al. 2004). Cohort studies reported 52% (Raeder, Podrid et al. 1985) and 88% (Shukla, Jowett et al. 1994) of patients on regular dose therapy with at least one ADE. However, most cohort studies investigating the risks of amiodarone are limited by small sample size, with a study typically involving up to 200 participants.

An additional potential ADE linked to amiodarone is major cardiovascular adverse event (MACE). A hospital-based retrospective analysis of 817 patients prescribed amiodarone over a 6 year period found 56 episodes of MACE using chart review. Using a Cox proportional hazards model controlling for patient characteristics and risk factors for MACE (hypertension, diabetes, ischaemic heart disease and left ventricular dysfunction) (Schuhlen, Kastrati et al. 1998, Boersma, Poldermans et al. 2001) found the presence of AIT to be a predictor for MACE (HR 2.68, 95% C.I. 1.53,4.68) (Yiu, Jim et al. 2009).

2.3.5 Critique of methods used in studies of ADE in healthcare

The literature review used broad inclusion criteria and selected studies featuring a wide range of study designs and methodology. In order to gain a better understanding of the effect of different methodological decisions on study results, several aspects of the included studies were investigated further which were of particular importance to their quality and internal validity. These included the chosen definitions of drug-related harm used across studies, methods used to gather data on ADEs and criteria used to establish a causal link to a drug or error.

Differences in definition of ADE and which drug-related events should be included

Studies selected in the review were inconsistent in their definition of ADE. Certain studies used the classic WHO definition which considered an ADE to be an unintended effect of a medication which occurs at normal dose used for therapy, which includes both non-preventable ADEs and those caused by errors (Classen, Pestotnik et al. 1997). Other authors, such as Pirmohamed et al and Davies et al (Pirmohamed, James et al. 2004, Davies, Green et al. 2009) applied the definition of ADR developed by Edwards and Aronson (Edwards and Aronson 2000), which characterised an ADR as an appreciably harmful reaction from a medicine which requires action or prevention on behalf of the healthcare professional.

Table 2.3: Description of study design in studies of incidence of ADE identified in the review						
Author	**Year**	**Design**	**ADE identification method**	**Severity**	**Preventability**	**Outcomes**
Hospital-acquired ADEs						
Lindley	1992	Prospective cohort	Chart review	Not assessed	Investigation of prescription records	ADR
Bates	1995	Prospective cohort	Chart review, ADE reports	Manual review by two physicians of all potential ADE cases	Manual review by two physicians of all potential ADE cases	ADE
Smith	1996	Retrospective cohort	Spontaneous reports	Assessment by pharmacist	Assessment by pharmacist	ADR
Bates	1997	Case-control	Chart review, ADE reports	Manual review by two physicians of all potential ADE cases	Manual review by two physicians of all potential ADE cases	ADE
Classen	1997	Case-control	Automated signals from electronic records	Assessment by physician	Naranjo algorithm	ADE
Cox	2001	Retrospective cohort	Spontaneous reports	Spontaneous reports	Spontaneous reports	ADR
Davies	2009	Prospective cohort	Chart review	Manual assessment by investigators	Naranjo algorithm	ADR

Table 2.3 continued

Author	Year	Design	ADE identification method	Severity	Preventability	Outcomes
Community-acquired ADEs						
Gandhi	2000	Retrospective cohort	Chart review, patient surveys	Chart review, patient surveys	Naranjo	Drug complications
Howard	2003	Prospective cohort	Chart review, interviews	Investigator review	Investigator review using Hepler criteria for preventability, Hallas criteria for causality	ADE
Gandhi	2003	Retrospective cohort	Chart review	Assessed by investigators - severe, life-threatening	Assessed by investigators	ADE
Gurwitz	2003	Retrospective cohort	Multiple: discharge summaries, chart review, automated signals, electronic records, incident reports	Not assessed	Not assessed	ADE
Pirmohamed	2004	Prospective cohort	Chart review, interviews	Not assessed	Naranjo, Hallas criteria	ADE
Morris	2004	Retrospective cohort	Large database	Diagnosis codes	Not assessed	drug-related harm
Passarelli	2005	Prospective cohort	Chart review	Chart review	Naranjo	ADE
Samoy	2006	Prospective cohort	Chart review	Chart review	Independetly assessed by 3 reviewers	ADE
Wu	2010	Retrospective cohort	Large database	Not assessed	Not assessed	ADE

Aronson et al placed more explicit stress on the presence of harm, whilst the original WHO definition did not exclude adverse effects which may not cause immediate harm, such as abnormal electrolyte levels. Therefore, the use of the WHO definition may potentially identify a larger set of events in a longitudinal analysis. Other studies opted to include all drug complications, which include any problem related to prescription drugs, including ADEs, non-adherence, undertreatment or overtreatment (Johnson and Bootman 1995, Gandhi, Burstin et al. 2000, Howard, Avery et al. 2003).

Methods for identification of ADEs and establishing causal relationship with drug therapy used in previous studies

The reported incidence of ADEs varied greatly across studies included in this review. Some of this variation may be explained by the multitude of methods used to identify and classify ADEs across studies. The most common method of identification of ADEs amongst selected studies was chart review. Studies used this method on its own (Lindley, Tully et al. 1992, Gandhi, Weingart et al. 2003, Passarelli, Jacob-Filho et al. 2005, Samoy, Zed et al. 2006, Davies, Green et al. 2009), as well as in combination with other methods, such as ADE reports from clinicians (Bates, Cullen et al. 1995, Bates, Spell et al. 1997), patient interviews or surveys (Gandhi, Burstin et al. 2000, Howard, Avery et al. 2003, Pirmohamed, James et al. 2004). Direct access to patient charts enabled investigators to obtain detailed information on drug treatments, laboratory results and symptoms experienced, allowing them to examine temporal relationships between medication use and an adverse event. This may be used to accurately attribute causality and determine the extent to which the event may have been avoided with better care. A disadvantage of this method is its labour intensive nature, which limits sample size as the number of charts which can be potentially reviewed is limited by the size of the investigating team. Another potential limitation is the need for ethics approval to gain access to patient records, which may delay the investigation.

Another method used in the studies selected for review was spontaneous reports of ADE by clinicians (Smith, Bennett et al. 1996, Cox, Anton et al. 2001). This is arguably the simplest method of collection of data on ADEs as it does not require intensive data collection on site. In the UK, the nationwide ADR reporting system is the Yellow Card scheme. It was set up in 1964 with the aim of recording ADRs in routine clinical care and collects reports from healthcare professionals, including doctors, pharmacists, and nurses, as well as patients. Data on ADRs are

collated and presented by the Medicines and Healthcare products Regulatory Agency (MHRA) on a yearly basis. In 2012, the MHRA received 26,037 reports of ADRs (Medicines and Healthcare Products Regulatory Agency 2013). However, high levels of underreporting of ADEs potentially undermine the accuracy of this method of estimating the incidence of medication-related harm in the UK. A previous study has found that only 1 in 20 ADRs are reported in primary care and hospital setting (Hazell and Shakir 2006). Studies comparing different methods of collection have shown that spontaneous reporting has consistently produced lower estimates of incidence of ADE across studies, compared to chart review (Madsen 1993, Chan and Critchley 1994), patient monitoring (Orsini, Orsini et al. 1995) and electronic record search (Cox, Anton et al. 2001).

The adaptation of computerised physician entry systems (CPOE) and decision-support systems which incorporate automated signals for required action introduced a new source of data on ADEs both in hospital and community settings. One study in the review used automated signals which identified ADEs from electronic records (Classen, Pestotnik et al. 1997). A study of ADEs in primary care went further and used automated signals in combination with other methods of data collection, including chart review, incidence reports and electronic patient records to maximise reliability (Gurwitz, Field et al. 2003). Using alternative sources of data alongside automated signals may identify clinically important events which are missed by the electronic algorithm and thus provide a more reliable identification method.

Several studies have used previously validated criteria for establishing causal link to a drug, such as the Naranjo algorithm (Naranjo, Busto et al. 1981) or Hallas algorithm for ADE-related hospital admissions (Hallas, Harvald et al. 1990). The use of such tools improves the reliability of the link between ADE and medication use, reducing the rate of false positives which are related to another cause. However, these methods are most applicable in combination with chart review and may be less practical in studies which employ large databases as the level of information listed in EHRs may be insufficient to reliably attribute causailty.

The development of large databases of electronic health records which combine prescription, laboratory test orders and results and systematically coded diagnoses offer an opportunity for large-scale analysis of ADEs in the community setting. Amongst the studies found in this literature review, two assessed ADEs by identifying episodes on a hospital database using standard diagnostic codes ICD-10 codes (Wu, Jen et al. 2010). This allowed a large sample size (over

500,000 suspected ADEs in nearly 60 million admissions) over a long study horizon (10 years). Another study identified 2000 patients with biochemical ADR in a sample of 74,000 patients on CPRD (McDowell, Coleman et al. 2010). The reliability of coding was ensured through quality control mechanisms (Jick, Kaye et al. 2003). This method is has the potential to be more reliable than spontaneous reporting of ADEs which relies on incentives of healthcare professionals and patients to report ADEs, which are not always clear. The limitation of large database research for ADE identification is its inherent reliance on the quality the underlying coding of diagnoses which are used to attribute causal link to a drug. In most cases it is not possible to independently verify the causal link, unless access to original patient charts is possible in order to apply algorithms to establish causality. Even with access to patient charts, however, systematically double checking though hundreds of thousands of records could prove to be an enormous and unnecessarily burdensome task.

Retrospective analysis of ADEs using electronic records extracted from healthcare providers or administrative databases may offer a compromise between accuracy and sample size; aspects which are both important in conducting a high-quality study of medication safety.

Methods used to assess preventability of ADEs

The Harvard Medical Practice Study (described in Section 2.2) and the large follow-up study published by Thomas et al revealed the extent of the problem of adverse events in hospitals and highlighted the problem of hospital-acquired adverse drug events in healthcare (Brennan, Leape et al. 1991, Thomas and Brennan 2000). One of the key findings of the above studies is that a large proportion of AEs and ADEs in healthcare may be prevented through better quality care: Brennan et al demonstrated that around a third of all hospital-acquired adverse events were caused by error, whilst around a half of the subset of serious adverse events resulting in death were preventable; Thomas et al that 35% of ADEs occurring in hospital were caused by a medication error. Since the publication of these large-scale epidemiological studies of AEs and ADEs in healthcare, research in this area has mainly focused on estimating the extent of preventable harm in the secondary care setting. A preventable ADE was generally defined as drug-related harm caused by a medication error which could have been avoided with better quality care. In order to judge whether an event may have been prevented through better care, most studies enlisted the help of a

clinical pharmacist or physician with experience in prescribing. The proportion of ADEs considered to be preventable ranged from 20-28%, with similar estimates from hospital-based and community-based studies (Bates, Cullen et al. 1995, Leape, Bates et al. 1995, Bates, Spell et al. 1997, Gandhi, Weingart et al. 2003, Gurwitz, Field et al. 2003). ADEs leading to hospitalisation tended to be more often preventable than not (range 67-72%) (Howard, Avery et al. 2003, Pirmohamed, James et al. 2004). The high rate of preventability amongst these ADEs is explained by the fact that the more severe events are more likely to be caused by error, and it is not unreasonable to assume that ADEs causing hospitalisation are more severe than other community-acquired ADEs. The association between severity and preventability of ADE was also reported in other studies. Gurwitz et al reported that 27.6% of ADEs in primary care were avoidable, compared with 42% amongst life-threatening ADEs (Gurwitz, Field et al. 2003).

2.4 Costs of ADE

Economic cost is defined as the opportunity cost incurred as a result of a specific event or decision. In other words, under conditions of scarcity, the use of resources to mitigate the effects of an ADE will result in the loss of opportunity of achieving a health gain elsewhere in the healthcare system. The costs of an ADE may include the cost of treatment, specifically the cost of the services which the patient may need to use, such as excess hospital admissions, medications or primary care clinic visits. Wider societal costs may include additional informal care or loss of income due to inability to work by the patient or their carer. The scope of a costing analysis depends on the perspective taken by the investigators. On one end of the spectrum is the societal perspective, which includes all costs incurred by society, which involve both direct medical cost and indirect costs, such as lost productivity. The other end of the spectrum is the perspective of a particular organisation or service, which is much narrower and only considers direct cost of delivering a particular type of care. This review aims to cover cost estimates from studies taking all perspectives, including societal and provider-specific, with a particular focus on community-acquired ADEs.

Preventable harm from medical and medication errors in healthcare are widely considered to impose a considerable economic cost on health systems. Given that cost containment in healthcare has become a priority in many developed countries, the cost of preventable adverse events has attracted the attention of

policy makers. In the UK, the Department of Health estimated a direct cost burden to the NHS of £2bn per annum from additional hospital stays due to adverse events in 2000 and £400 million in legal settlements, with a potential liability of £2.4bn (Department of Health 2000). Kohn et al (Kohn 1999) reported an estimated an annual cost of adverse events of $17 to $20 billion in the US, which included both health and non-health costs such as lost productivity in 1999. These estimates were derived using extrapolations from smaller studies whose generalisability to the entire population was not guaranteed, which may have resulted in unreliable estimates. Similarly, there was little indication of how much of this cost may have been avoided as only a subset of adverse events in healthcare are preventable. In order to report the cost of preventable harm in healthcare, this literature review focused on individual-level costing studies.

The review employed several criteria to evaluate the information contained in studies identified in the literature search. Studies found in the literature search were assessed on quality of the costing analysis. This was based on the following criteria:

i. Clearly stated perspective of the costing study e.g. third payer, societal, single provider
ii. Included a description of the methods used to identify ADEs
iii. Inclusion of all relevant direct medical costs of treatment, including primary care, hospital care, tertiary/social care, medications, tests
iv. If a third payer provider perspective is not clearly stated, inclusion of societal and indirect costs, such as loss of productivity and costs of informal care
v. The time horizon of the study and data collection, which should reflect the nature of ADEs studied and the time scale of impact on healthcare resource-use
vi. Methods for attributing cost of treatment to the ADE e.g. employing pharmacists or clinicians with experience in prescribing to ascertain the link between drug and adverse event

Most studies examining the cost of ADEs were set in the hospital setting, which is consistent with the disproportionate focus of the wider medication safety literature on secondary care.

2.4.1 Hospital-based studies

A nested case-control study in two US hospitals which examined 4108 admissions found that ADEs occurring in hospital were associated with an excess length of stay of 2.2 days and statistically significant excess cost of $3244. The excess cost for preventable ADEs was $5857, although it did not reach statistical significance at 5% level (Bates, Spell et al. 1997). The presence and preventability of an ADE were examined by two independent reviewers. Costs were estimated by matching ADE patients with non-ADE controls, adjusting for patient covariates using linear regression and attributing additional cost per ADE by multiplying the intensity of resource use (represented by DRG code) by the length-of-stay of each case.

Another large-scale matched case-control study of hospital-acquired ADEs in the US (1580 cases with ADE matched to 20197 controls). The authors reported that cases had an excess length of stay of 1.74 days and cost of $2013 attributable to the ADE, compared to controls, after adjusting for patient characteristics (Classen, Pestotnik et al. 1997). The authors used a similar method of identifying and costing ADEs as the previously described study by Bates et al to find the cost of excess hospital stay attributed to ADE. The excess cost varied considerably across the different types of ADE, ranging from excess cost of $677 for itching to $9022 for drug-induced fever.

A case-control study in Spain of 20014 inpatient episodes found an excess cost of €1641 attributed to the ADE (Pinilla, Murillo et al. 2006). The authors focused on ADEs caused by error (i.e. those which were preventable) and estimated the excess cost by matching patients with ADE to controls. The cost of stay was calculated by microcosting each service and item used by patients, such as bed cost, procedures, drugs and other consumables, rather than through DRG or another grouping method, thus potentially providing a more accurate estimate than other studies which used costing based on length-of-stay. However, the authors only included costs in hospital, choosing to ignore direct medical costs incurred in other settings following discharge, thus limiting the accuracy of the cost estimates.

2.4.2 Community care-based studies

Patients suffering an ADE in the community may incur direct medical cost of treatment across multiple healthcare sectors. Substantial costs may be incurred as a result of emergency hospital admission. A meta-analysis of studies of the cost of community-acquired ADEs leading to hospitalisation published between

1982 and 1996 found a median length of stay of 8.7 days and a cost of DM (Deutschmarks) 4046 (at 1995 price levels), which corresponds to around £1386 per admission in 1995 GBP (Goettler, Schneeweiss et al. 1997). Based on evidence from studies, the authors concluded that 1/3 of the cost is preventable. The authors synthesised data on excess length of stay of ADEs as a proxy for the excess resource use associated with ADEs, and did not include costs outside the hospital. In addition, the average cost of a hospital bed-day in Germany was applied to estimate the cost, which does not inform the reader of the cost burden of difference types of ADE and policy-makers which ADEs or drug classes should be targeted with interventions to reduce cost.

A literature review of 45 economic studies of ADE published between 1990 and 2003 reported a mean cost of hospital-acquired ADEs between 2284 and 3093 USD (5585 USD for preventable ADEs, at 2000 price level) (Rodriguez-Monguio, Otero et al. 2003). For community-acquired ADEs resulting in hospital admission, the excess cost attributable to the event was 3066-7062 USD. The methodological strength of this review is the inclusion of studies which clearly reported the cost components in their costing analyses and synthesised all costs in terms of USD, which allowed comparisons. However, it focused on costs arising from hospital care only, either in case of ADE-related hospitalisations or ADEs acquired in hospital. It did not evaluate costs incurred primary care, prescribed medications, social care or indirect or societal costs of drug-related harm.

A matched case-control study measured the resource use and healthcare cost of 1210 elderly patients with ADE matched to ADE-free controls in a US primary care facility. ADEs were identified and costed using electronic health records. The authors concluded that a community-acquired ADE results in an excess cost of USD 1310 (in 1999 prices) over the 6 weeks following an ADE (Field, Gilman et al. 2005). The study employed robust methods of identifying ADEs from multiple sources, which included ADE reports by providers, search of free-text fields, hospital discharge and emergency department notes and automated signals of inappropriate drugs or drug-drug interactions. The costing method was more inclusive than other similar studies; data on resource use in primary care, hospitals, home care, ambulance transport and referrals to other services was collected. Unit costs were derived for individual cost items such as the time of healthcare professionals, bed costs, medications and procedures from Medicare fee schedule, which reflected more closely the true cost from the perspective of the third-party payer. The strength of the study was the inclusion of resource use from services outside hospital, including primary care, medications and

ambulance transport. The results of the study were limited by a short follow-up period of only 6 weeks, meaning that the estimate only took into account the short-term impact of ADEs on resource utilisation.

A large prospective cohort study of ADE-related hospitalisations in the UK found that 6.5% of 18820 admission were caused by ADRs, with 72% of these considered avoidable. The average length of stay for an ADR-related admission was 8 days costing an average of GBP 1824 per patient (in 2002 prices), which when extrapolated to national level corresponds to GBP 466 million per year, which translated to GBP 656 million in 2014 GBP (Pirmohamed, James et al. 2004). However, the authors used an average cost per bed day to estimate the excess cost attributable to ADEs, which was insensitive to the type of ADE and did not reflect differences in complexity of treating different types of ADE. In addition, costs incurred outside the hospital setting were not mentioned in the study.

Johnson and Bootman (1995) developed a cost-of-illness model to estimate the effect of drug-related morbidity in the primary care setting. The model consisted of a decision-tree structure reflecting the treatment pathway of community-acquired ADEs, various services used such as GP visits, medications, emergency department, hospital admissions and long-term care. Probabilities and costs were populated from literature and expert opinion. Cost per patient was estimated by folding back the tree and a population-level estimate was produced by extrapolating the model results. The authors estimated the total healthcare cost of treating drug-related problems to be USD 76.6 bn (1994 cost year) , with 62% of the cost occurring in hospital, at a mean cost of USD 5411 per admission (Johnson and Bootman 1995). However, the authors did not make a distinction between preventable and non-preventable ADEs and thus did not estimate how much of this cost could have been avoided. In addition, the authors used only literature and expert opinion instead of observational data, meaning the results were prone to inaccuracies arising from the inconsistency in reported incidence of ADE in the literature described in the previous section.

Gyllensten et al analysed electronic records of 596 patients with ADE in primary care in Sweden over a 3 month follow-up period. The authors estimated the direct attributable cost of ADE of USD 2076 (at 2008 price level) and noted that 46% of cost was incurred outside the inpatient hospital setting. Outpatient treatment was the 2nd highest direct cost item representing USD 547 per patient. The authors also estimated indirect costs, such as loss of productivity from sick leave and

disability, which was estimated at USD 1719 per patient. ADEs were identified by a pharmacist, with causality established from patient records and history. The strength of the study was the inclusion of a wide range of services in cost estimation and inclusion of indirect costs. The potential limitation is the short observation period, which missed out costs from services used beyond the 3 month follow-up period. Another limitation is the lack of adjustment for confounding factors, meaning that the attributable cost may have included cost of care due to more frail and multimorbid condition of patients with ADE.

2.4.3 Other costs: productivity losses

Whilst ADEs and medication errors may incur substantial healthcare cost, the evidence of the indirect costs of ADEs, such as lost productivity and income, is sparse. Evidence suggests that indirect costs of medical errors may be around USD 878 million annually (in 1989 USD), including 276 million in lost wages and 441 million in lost productivity (Johnson, Brennan et al. 1992). Estimates were derived by assessing the potential litigation costs of AE using interviews of 794 individuals with adverse events and simulating population-level cost which included both direct medical expenses and lost productivity. However, specific studies of individual-level indirect cost attributable to ADEs have not been conducted.

Cost estimates based on excess hospital stay may be prone to variations in hospital costs and charges across sites. Most studies investigating ADEs were carried out in one or two sites, meaning that the cost estimate may be explained by inter-site variation, rather than the true healthcare cost. Leendertse et al conducted a multi-centre study involving 21 hospitals in the Netherlands reported a mean excess length of stay attributable to ADE of 8 days (Leendertse, Egberts et al. 2008), which is similar to the estimates found in Pirmohamed et al and Goettler et al. The similar length of stay estimated from a large sample size and multiple sites supports the cost estimates found in costing studies of the impact of ADE in terms of hospital cost.

Most studies examining the healthcare cost of ADEs used excess length of stay as a proxy measure for the additional cost of ADE. This is a convenient approach as it is relatively easy to measure from accessible electronic or paper-based hospital records. Although hospital cost represents a large proportion of the total healthcare cost and is often the largest cost item in budget impact studies, it fails to take into account the impact of ADEs on the direct medical cost incurred in

other sectors of healthcare – in particular follow-up primary care (GP consultations and medications) and other services, such as intermediate, social care and ambulance service. This is an important methodological limitation as follow-up care in primary, social care and additional medications following discharge may be substantial. Most of the studies estimating costs have a short follow-up, probably owing to their design which is geared towards measuring clinical, rather than economic effects of ADEs. This limits the ability to extrapolate resource use and cost over a longer time horizon, such as one year or beyond, which may be of interest when building an economic model. In addition, it fails to take into account long-term impact in terms of resource use and cost, particularly in community-based services.

Table 2.4: Summary of studies of cost of community-acquired ADE						
Study	**Year of publication**	**Cost year**	**Currency**	**Description**	**Original cost**	**Cost in 2014 GBP**
ADE-related admissions						
Goettler et al	1997	1995	DM	Estimated as the average LOS due to an ADR observed between 1975 and 1996, multiplied by cost per day in hospital	4046	3043.153
Rodriguez-Monguio	2003	2000	USD	Estimates of ADE-related admission	1286-4441	1223-4225
Field et al	2005	1999	USD	Primary and secondary care resource use in the 6 months following an ADR	1310	1278.489
Pirmohamed et al	2004	2002	GBP	Median LOS multiplied by cost per bed day	1824	2565.966
Johnson and Bootman	1995	1994	USD	Model based on costof hospital admission, GP visit, medication	5411	6830.171
Gyllensten	2014	2008	USD	Estimated hospital inpatient and outpatient cost	2076	1167.466
ADEs not leading to hospitalisation						
Rodriguez-Monguio	2003	2000	USD	Estimates from studies reporting cost of ED visits caused by ADEs	329-422	313-401
National level estimates						
Pirmohamed et al	2004	2002	GBP	Based on incidence rate of ADE, no. of hospital admissions per year and cost per ADE-related admission	466 million	656 million

2.5 Discussion

This chapter described the issues of adverse events, adverse drug events and medication errors through a summary of key policy papers and a review of literature on the incidence, nature, preventability and cost of ADEs in healthcare. Healthcare safety has become a priority for healthcare policy-makers in the past two decades following the emergence of evidence from observational studies on the harm caused by the very treatments and processes which are delivered to improve health. This included evidence from the Harvard Medical Practice Study (Leape, Brennan et al. 1991) and other large-scale studies which followed, which reported that up to 3.7% of hospitalisations resulted in an adverse event and up to 98,000 patients died yearly as a consequence of AEs in the US. The Department of Health found similarly worrying results in the UK, suggesting that up to 850,000 adverse events occur in British hospitals yearly. Reports by influential policy-making organisations, such as "To err is human" by the US institute of Medicine and "Organisation with a Memory" by the UK Department of Health highlighted the failures of the respective healthcare systems at identifying, preventing and learning from adverse events. This has precipitated a large body of research on the nature and causes of adverse events in healthcare. Medications, which are the most common form of medical intervention, have predictably been identified as a leading cause of iatrogenic harm.

The issue of medication-related adverse events first became the focus of epidemiological research in the 1960s, when the first small and medium-scale studies estimated the incidence of ADEs in hospitals. Medication harm constitutes a large proportion of iatrogenic harm. According to a systematic review of hospital-based studies published between 1990 and 2005, 6% of hospital episodes result in an ADE, although the range was 0.2% to 60.7% due to differences in data collection methods and definitions of ADE/ADR used. Medication harm may result in unnecessary hospital admission (3.7% of admissions in UK are caused by ADE, range 1.4% to 15.4%) (Howard, Avery et al. 2007) or excess treatment in primary care (2.8% of patients on chronic drug therapy in the community setting in the UK suffer an ADE) (Gandhi, Weingart et al. 2003). Another study reported an incidence of preventable community-acquired ADEs of 13.8 per 1000 patient-years of therapy (Gurwitz, Field et al. 2003).

ADEs can occur in any setting of healthcare. Most studies focused on hospital-acquired ADEs due to the practical ease of data collection: access to charts and existing measures in place, such as patient rounds and proximity of healthcare professionals to collect data. Estimates of the proportion of hospital episodes which included an ADE across individual studies selected in this review ranged between 0.06% and 15.8%. The estimated proportion of hospital admissions caused by ADEs ranged between 0.5% and 25%. Data regarding the overall rate of community acquired ADEs (not necessarily resulting in hospital admission) was more sparse. The wide range in estimates may be explained by the heterogeneity in methods of identifying ADEs (self-reporting, electronic record search, chart review), medications included in the study and lack of agreement of what constitutes an ADE and publication bias.

The review has identified several studies which have estimated the direct medical costs of ADEs. Estimates from studies conducted outside the UK were converted to GBP and all costs inflated to 2014 to allow cross-country comparisons. The cost of hospital admissions caused by ADEs varied from £1167 to £6830. This places a substantial burden on healthcare systems, with yearly excess cost resulting from ADE-related hospitalisations of £466 million in 2002, which is £656 million in 2014 when medical cost inflation is taken into account. The actual direct medical cost of ADEs would be higher as the above estimate did not include the medical cost of treating ADEs which do not result in hospital admission. Indirect cost impact of ADEs through lost productivity and the cost of informal care is poorly understood due to difficulties in estimating indirect costs. Most studies in the review used excess hospital stay as a proxy for the cost impact of ADE due to the ease of applying unit costs for additional days spent in hospital. Inpatient admissions are the largest cost item in healthcare and methods based on estimation of total cost based on this service alone may yield a good approximation of the true cost. However, it is unclear which parts of the healthcare system are mostly used in the treatment and whether there may be cost spill overs, for example into social or intermediate care services.

Some of the studies in the review measured the extent to which ADEs observed were avoidable. Studies were consistently reporting that a large proportion of ADEs were preventable (20%-28%) and that the majority of severe ADEs requiring hospitalisation were preventable (67%-72%). This may be explained by the fact that most ADEs are minor (such as electrolyte disorders, minor illness) and are an unavoidable risk of drug therapy. Most drugs prescribed in primary care have been tested for the risk of severe idiosyncratic ADEs before approved

for patient use; therefore the risk of such severe or life-threatening events under conditions of normal therapy is very unlikely. Therefore, it is intuitive to assume that a larger proportion of severe ADEs occurring in healthcare are a consequence of a medication error.

Monitoring error is a major cause of ADEs in primary care. Inadequate monitoring has been shown to be a cause of 36% of preventable community-acquired ADEs (Gurwitz, Field et al. 2003) and 22% of ADE-related hospitalisations (Howard, Avery et al. 2007). Despite the fact that monitoring errors may cause almost as many ADEs as prescribing errors, most literature on medication safety has focused on problems in prescribing. The evidence on the role of monitoring and development of monitoring indicators to improve quality of care is sparse in comparison to the literature on inappropriate prescribing.

Assessment of evidence of the incidence of ADEs in healthcare is complicated by a lack of agreement on the nomenclature used to define preventable harm caused by medication use. Many pharmacoepidemiological studies identified adverse drug reactions (ADR) under the WHO definition and distinguished between dose dependent (Type A) and idiosyncratic (Type B) reactions to a drug. However, this limits medication safety research as this definition does not allow for a distinction between events caused by errors (which are preventable and can be targeted by interventions) and those which are non-preventable. This gave rise to a new set of definitions, which divide medication-related harm into ADRs, adverse drug events (ADEs) and preventable adverse drug events (pADE), further complicating cross-study comparisons.

There was no "gold standard" method for identifying ADEs across studies in the review, although several methods were commonly used. The most commonly used methods included chart review, data from spontaneous reports of ADE, automated electronic signals and systematic searches of EHRs. Chart review was used by several studies in the review and arguably provided the best method of observing ADEs in practice, including prospective or retrospective observation. Access to all patient information, including charts, discharge forms and information provided by healthcare professionals in charge of care may be used to accurately pinpoint the causes of an ADE and identify the active failures which have led to it. However, this method is resource-intensive and may be impractical on a large scale. It may be inadequately powered in a study of rare ADEs, and the need for physical data collection may introduce new sources of bias, such as confirmation bias or measurement error. In order to improve identification of

ADEs from chart review, some studies employed multiple physicians to independently verify records.

Spontaneous reporting is a convenient method as it employs readily available data from routinely collected ADE reports. The disadvantage of this method is the high level of underreporting of ADEs (Hazell and Shakir 2006). This was confirmed from the findings of this review, as spontaneous reporting consistently provided lower estimates compared to other methods of collecting data on ADEs. Several studies have made head-to-head comparisons of spontaneous reporting and alternative methods, showing that this method consistently missed ADEs (Madsen 1993).

The number of studies which have used electronic health records (EHRs) to study ADEs has risen in recent years due to the widespread adoption of electronic data recording in primary care. This has provided opportunities for large-scale studies of rare ADEs and identification of medication errors across multiple facilities and regions. Decision support systems for physicians which are built into the practice electronic system often incorporate automated signals for medication errors and ADEs, which are an additional source of data for studies of ADEs. The use of algorithms based on coding systems for diagnoses and clinical events, such the Read hierarchy or International Classification of Disease (ICD-10) codes, is a more reliable and systematic method of identifying ADEs and minimising the risk of missing clinically relevant events. The conclusions drawn from EHR data may be limited by the inability to access the full patient file, including patient charts and physician's notes, as well as no opportunity for direct observation by an investigator. This may miss out key information regarding the nature and causes of an ADE. In addition, the quality of recording of diagnoses and clinical events may not always be guaranteed, and the investigator is unable to observe that which has not been recorded electronically. To mitigate these issues, some studies chose to cross-check with other sources of data, such as patient charts or automated ADE signals, in to minimise the number of missed events in the data (Gurwitz, Field et al. 2003).

The issues related to analysis of incidence of ADEs are directly applicable to analyses of the cost impact of ADEs. In addition, the inconsistency of costing methods complicated comparison across studies. Most studies only included a limited perspective of cost of hospital, using excess length of stay as the primary measure of additional resource use attributable to ADEs. A number of studies have included the cost of other services and more detailed microcosting methods,

such as costing based on resource-use grouping (DRG) and costing of individual elements used in treatment of ADEs. Most studies elected to take a third-payer cost perspective, meaning that the impact of ADEs in terms of societal costs, such as loss of wages and productivity, informal care was unclear.

Several studies used validated methods such as the Naranjo or Hallas algorithms to establish causality and determine the severity of the ADE. This provided a systematic method of establishing the causal link to a drug and determining the severity of ADE.

Several studies included in the literature review described in this chapter have identified cardiovascular drugs as a medication area frequently associated with community-acquired preventable ADEs and ADE-related hospital admissions. Certain medications in this therapeutic area, including ACE inhibitors and diuretics in the treatment of hypertension, ACE inhibitors indicated for heart failure and amiodarone used in the treatment of cardiac arrhythmia in primary care pose a risk of ADEs. The safety profile and therapeutic outcomes from these medications may be improved through regular monitoring by identifying ADEs at an early stage of development before they have caused harm to the patient. The next chapter will explore the topic of medication monitoring by examining the current guidelines for both initiation and follow-up monitoring of these medications in UK primary care. Chapter 3 also includes a review of studies which have investigated the compliance to necessary monitoring of this subset of medications and lays out a rationale for conducting an observational study of monitoring to improve knowledge in this area of medications safety.

Chapter 3:

Literature review: medication monitoring for the prevention of ADEs

3.1 Introduction to the chapter

This chapter includes a review of guidelines for monitoring of study medications and the existing literature on the use of monitoring for the prevention of adverse drug events in the primary care setting. Compliance to recommended monitoring in ACEI, loop diuretics and amiodarone was assessed. This is followed by a review of studies which examined the effectiveness and cost-effectiveness of interventions to improve monitoring.

3.2 Aims and objectives of chapter

1. Outline recommendations for follow-up monitoring of medications included in the studies and appraise the quality of UK guidelines for follow-up monitoring of cardiovascular medications.
2. Review previous studies investigating the completion of monitoring both in the UK and abroad.
3. Examine studies evaluating the effectiveness and cost-effectiveness of interventions to improve monitoring in primary care.

3.3 Background of medication monitoring

In Chapter 1 of this thesis, monitoring was defined as a tool that can be used to guide treatment in order to achieve therapeutic goals whilst simultaneously protecting the patient from ADEs (Pirmohamed and Ferner 2003). Regular monitoring during the course of therapy can improve the safety profile of certain medications by identifying early signs of onset of ADEs. This is the case with a number of cardiovascular medications, such as ACE inhibitors and diuretics, which can cause mild ADEs that have the potential to progress to severe conditions, and medications which may cause organ dysfunction, such as amiodarone-induced hypothyroidism or thyrotoxicity. The safety profiles of medications included in this study were described in more detail in Chapter 2.

In certain cases, monitoring is necessary in order to avoid potentially harmful under-treatment of severe cardiovascular conditions. For instance, in the treatment of cardiac arrhythmia, amiodarone may be the best therapeutic option for some patients and choosing not to undergo drug therapy due to an elevated risk of thyroid dysfunction may expose the patient to the risk of cardiac arrest and sudden death (Jafari-Fesharaki and Scheinman 1998). Carrying out treatment, even in the case of a potential risk of thyroid dysfunction may, in combination with regular monitoring, is likely to be the optimal therapeutic strategy. In the case of ACE inhibitors and loop diuretics, regular monitoring of serum potassium, sodium and creatinine levels may help to identify the onset of serum potassium abnormalities or acute kidney injury and allow the patient to control hypertension or the progression of heart failure. Post-treatment monitoring may be carried out following discontinuation of amiodarone therapy as thyroid dysfunction linked to amiodrone may develop several months after the last administered dose of the drug (British Thyroid Association 2006).

Inappropriate decisions regarding the monitoring of medications, such as failure to carry out recommended tests at the recommended frequency listed in guidelines or failure to respond to abnormal test results in a timely manner constitute monitoring errors. Monitoring error has been identified as the leading cause of preventable ADEs across multiple healthcare settings. A systematic review of observational studies of community-acquired ADEs identified five studies which attributed events to different causes, with the proportion of ADE-related hospitalisations caused by a monitoring error ranging from 21% to 42% of cases (Howard, Avery et al. 2007). Another systematic review of studies of preventable ADEs in primary care reported that inadequate monitoring was the most common cause of preventable ADEs leading to hospitalisation (45.4%), ahead of prescribing errors and other medication-related issues, such as patient non-adherence (Thomsen, Winterstein et al. 2007).

Observational analyses of the causes of ADEs in a range of healthcare settings have identified inadequate monitoring as a leading cause of drug-related harm. An observational study of community-acquired ADEs in the US ambulatory care setting in a cohort of 27617 elderly (65 years and older) patients found 421 preventable ADEs. The authors reported that 36.6% of these pADEs were caused by monitoring errors, including inadequate frequency and failure to respond to abnormal test results (Gurwitz, Field et al. 2003).

The same authors analysed cases of ADE over a 9 month follow-up period in the long-term care sector in the US. The authors identified 338 ADEs linked to one or more errors in the medication management process through assessment of electronic health records, electronic alerts and incident reports by pharmacists. They concluded that failure either to conduct monitoring or to respond to abnormal test results was linked to preventable ADEs in 80% of cases. The causal link was established by physicians who reviewed each individual case to identify a relationship between error and event. An ADE was deemed to be caused by a monitoring error if it might have been prevented had tests been carried out appropriately (Gurwitz, Field et al. 2005).

Budnitz et al reviewed 21000 cases of ADE-related emergency department visits reported to a US-based electronic ADE surveillance system. The authors found that amongst patients aged 65 and over, 54.4% of visits which resulting in hospital admission due to an ADE acquired in outpatient care were linked to a drug which commonly requires regular monitoring, although the methods for establishing the causal link between a monitoring error and the ADE were not described (Budnitz, Pollock et al. 2006).

Excessive use of monitoring in situations where it is unwarranted is an unnecessary use of resources and can even result in harm to the patient. A study of 162 patients monitored during inpatient or outpatient treatment with digoxin applied criteria for appropriateness of monitoring developed by an expert panel. The authors reported that only 16% of the tests carried out during the study period had an appropriate indication, such as presence of known risk factors, change in dosage or evidence of previous digoxin toxicity (Canas, Tanasijevic et al. 1999). The authors found that many tests were carried out too frequently in patients on stable therapy, constituting a waste of healthcare resources. The study estimated that savings of over USD 460k (1999 USD) may have been achieved if the number of referrals for digoxin level tests was reduced by 80%. Moreover, conducting TDM shortly after a dose change may lead to erroneous treatment decisions as the short time interval may be insufficient to reach a steady state in terms of pharmacokinetics. A study of TDM of antiepileptics in hospital inpatients found a similar result: only 27% of drug level tests had an appropriate indication (Schoenenberger, Tanasijevic et al. 1995).

Literature on medication safety has mainly focused on errors at the prescribing stage of medication management. This may be explained by the fact that inappropriate medication use is associated with lower patient quality-of-life

(Classen, MStat et al. 2006) and increased risk of hospitalisation (Fick, Waller et al. 2001, Klarin, Wimo et al. 2005). Previous studies, however, have shown that inappropriate prescribing accounts for only a small proportion of community-acquired ADEs. A systematic review of studies into prescribing error found that 6.6% of ADE-related hospitalisations were linked to inappropriate medications (Budnitz, Lovegrove et al. 2011). This may be explained by the fact that ADEs in primary care most frequently arise from common medications used in the treatment of chronic conditions, the risks of which are known to physicians. A systematic review of studies of preventable drug-related hospitalisations found that the three drugs most commonly involved in causing ADE-related hospitalisations were antiplatelets, diuretics and NSAIDs (Howard, Avery et al. 2007). Failures in monitoring have received considerably less attention than errors in prescribing despite the fact that inadequate monitoring is at the root of a substantial proportion of ADEs in primary care. A systematic review of studies of preventable ADEs in ambulatory care in the US reported that the proportion of ADEs caused by monitoring errors (36%) was comparable to the proportion of ADEs caused by inappropriate prescribing (43%), whilst failure to monitor drugs was the most common cause of drug-related hospital admissions (involved in 45% of ADE-related hospitalisations) (Thomsen, Winterstein et al. 2007). Therefore, monitoring errors merit as much attention from policy-makers and medication safety researchers as inappropriate prescribing in primary care.

3.4 Review of guidelines for follow-up monitoring of ACEI, loop diuretics and amiodarone.

One area in which monitoring may improve the safety of medications is cardiovascular therapy (Raebel 2012). In UK primary care, monitoring of cardiovascular medications, such as ACE inhibitors, loop diuretics and amiodarone is recommended in a number of national guidelines including those issued by the National Institute for Health and Care Excellence (NICE), health service organisations and organisations specialising in specific disease areas such as British Thyroid Association and British Hypertension Society (Sever, Beevers et al. 1993, British Thyroid Association 2006, Primary Care Quality & Information Service 2008, National Institute for Health and Care Excellence 2011). This section examines each of the study medications in the context of guidelines published in the UK and their impact on recommendations for appropriate practice for regular monitoring in UK general practice.

The risks associated with the use of ACEI and diuretics are well recognised in clinical practice and precautionary measurement of urea, electrolytes and creatinine has been considered an essential element of antihypertensive therapy for many years (Sever, Beevers et al. 1993). Therapy with ACEI agents is associated with hyperkalaemia and a risk of impaired renal function in some patients whilst loop diuretics are associated with hypokalaemia and hyponatraemia, as described in Chapter 2 of this thesis. Complications of these events are some of the most common severe ADEs: one in seven severe ADEs leading to intensive care involved electrolyte disorders or acute kidney injury (AKI) linked to ACEI and loop diuretics (Darchy, Le Mière et al. 1999). Periodic monitoring of serum potassium and creatinine in ACEI and, additionally, serum sodium in loop diuretics is used to identify abnormal electrolyte levels and deteriorating renal function before complications occur.

Chapter 2 examined the risk profile of amiodarone and found an association with hypothyroidism and thyrotoxicity based on evidence from previous studies. Thyroid dysfunction develops over a period of several weeks and can thus be managed through early detection using thyroid function tests (Podrid 1995). A review of studies evaluating guidelines for thyroid function test (TFT) monitoring of amiodarone therapy noted a large variation in the criteria for monitoring owing to a lack of supporting evidence of the benefits of monitoring beyond expert opinion (Stelfox, Ahmed et al. 2004). Monitoring of thyroid function during amiodarone therapy in UK primary care is recommended in a number of published guidelines.

ACEI, loop diuretics and amiodarone are mentioned in a large number of national and local guidelines for monitoring. This section summarised the information listed in those which were deemed relevant for this programme of research. The full summary of guidelines is included in Table 3.1 below.

Some of the guidelines, including those published by NICE and the BNF, issued specific instructions for subgroup of patients. For instance, increased frequency of monitoring in ACEI and loop diuretics is recommended in the presence of heart failure, co-prescription with certain medications and pre-existing renal dysfunction (National Institute for Health and Care Excellence 2011) (British National Formulary (online) 2015). Therefore, recommendations are reported separately for all patients on the drug and specific sub-groups.

Other national guidelines which mentioned monitoring of ACEI, diuretics and amiodarone were the Scottish Intercollegiate Guidelines Network and the Primary

Care Quality & Information Service in Wales (Scottish Intercollegiate Guidelines Network 2001) (Primary Care Quality & Information Service 2008). The most reliable source of recommendations for monitoring of amiodarone therapy was the British Thyroid Association, which listed detailed information on the required test and frequency (British Thyroid Association 2006).

The recommendations by national bodies served as the basis for local published guidelines on appropriate monitoring practice at the primary care trust (PCT) level. Whilst the guidelines published by NICE or BNF listed general recommendations for which tests should be carried out and produced warnings about the risks of therapy, the local guidelines represented a specific checklist which included information on the required test, frequency and action in response to an abnormal test result. They combined evidence from NICE and BNF with other published sources to provide a rationale for their recommendations. The guidelines examined in this chapter mostly relied on other previously published recommendations to support the included information and did not refer to original studies of monitoring.

One of the sources cited in local guidelines was a set of guidelines published in the British Medical Journal (BMJ) by a guideline developing group comprised of GPs and pharmacists, which advised monitoring renal function and serum potassium at least once in 12 months (Eccles, Freemantle et al. 1998). However, the authors stated that they had no basis for recommending any particular frequency of monitoring, although they noted that testing should be carried out once in 12 months without stating any further evidence to support this claim. The only recommendation listed was for physicians to regularly ask about the adverse effects of ACEI (it did not mention who or when they should ask), referring to evidence from expert opinion.

Another cited source is a study by Smellie et al (Smellie, Forth et al. 2007), which is itself based on recommendations previously published by the Royal College of Physicians of London and the Renal Association (Joint Specialty Committee on Renal Medicine of the Royal College of Physicians of London and the Renal Association 2006). This publication recommended regular monitoring of renal function and electrolytes, justifying this by referencing previously published recommendations by NICE (which have since been updated, NICE guidelines 108 and 127 described earlier), the British Hypertension Association (Sever, Beevers et al. 1993) and a guideline produced by the US Department of Health (Chobanian, Bakris et al. 2003). The latter recommended serum potassium and

creatinine once or twice per year, with more frequent monitoring in the presence of heart failure or diabetes, but made no reference to evidence from studies. Other guidelines published in the US include the ACC/AHA Guidelines for the management of heart failure, which listed the need for regular monitoring of serum electrolytes and renal function (Yancy, Jessup et al. 2013), but noted that the main source of the recommendation came from expert opinion.

Similarly to ACEI and loop diuretics, guidelines for TFT monitoring in amiodarone relied on a weak evidence base to support their recommendations. The British Thyroid Association, for instance, referred to evidence from non-experimental descriptive studies, but failed to cite any specific published work to support the need for regular monitoring of thyroid function. Supporting evidence for local guidelines did not include any studies of monitoring of amiodarone and relied on references from the BNF, British Thyroid Association and manufacturer's guidelines (Cordarone X 2004).

In addition to providing information on the required test and frequency in follow-up monitoring, many of the listed guidelines included recommendations for action in case of an abnormal test result. In the context of ACEI monitoring, there was little agreement in terms of required action in case of abnormal test results: in case of a rise of creatinine by over 50% compared to baseline levels, therapy should be modified (London and South East Medicines Information Service 2008), reduced or stopped (North West 2002, Primary Care Quality & Information Service 2008) or only stopped (Erewash 2006); the GP should consider modifying or stopping treatment if serum potassium rises above 5mmol/L (North West 2002), 5.5mmol/L (Erewash 2006, London and South East Medicines Information Service 2008) or 5.9mmol/L (Primary Care Quality & Information Service 2008). Similarly, there was little agreement on the required response to abnormal test results in loop diuretic monitoring: stop treatment if serum potassium is above >5.5mmol/L or 6mmol/L (depending on source); review therapy (London and South East Medicines Information Service 2008) or add a potassium-sparing diuretic if serum potassium falls below 3mmol/L (North West 2002). One guideline failed to list any recommended action (Erewash 2006). In the case of amiodarone, three of the guidelines recommended stopping therapy in case of development of AIT, and thyroid replacement therapy in case of AIH, without interruption of amiodarone therapy (North West 2002, London and South East Medicines Information Service 2008, Primary Care Quality & Information Service 2008). However, only two guidelines listed any definition of AIT (raised T3 and T4, low TSH) (London and South East Medicines Information Service 2008,

Primary Care Quality & Information Service 2008) and only the Thyroid Association guidelines listed the specific ranges of test results at which abnormal thyroid function should be diagnosed (British Thyroid Association 2006). One guideline made no reference at all to required action, mentioning that specialist advice should be sought (Erewash 2006).

Table 3.1 Summary of information included in guidelines for follow-up monitoring.

Drug	Indication	Test and frequency	Reference
ACEI	All patients	Renal function & electrolytes at regular intervals	(British National Formulary (online) 2015)
		Serum potassium and creatinine 1-2 weeks after initiation, then following dose change	(Scottish Intercollegiate Guidelines Network 2001)
		Serum potassium once in 12 months	(Primary Care Quality & Information Service 2008) (North West 2002, Erewash 2006, London and South East Medicines Information Service 2008)
	Co-prescribed medications which increase risk of renal/electrolyte dysfunction (CCB, diuretics)	Regular monitoring of renal function, serum potassium and sodium during therapy	(National Institute for Health and Care Excellence 2011) (British National Formulary (online) 2015)
	Chronic heart failure	serum potassium, sodium, creatinine and estimated glomerular filtration rate (eGFR) once in 6 months	(National Institute for Health and Care Excellence 2010) (National Institute for Health and Care Excellence 2011) (British National Formulary (online) 2015)
	Chronic heart failure and risk factors (e.g. pre-existing renal dysfunction	Same as above, but more frequent	(National Institute for Health and Care Excellence 2010) (National Institute for Health and Care Excellence 2011) (British National Formulary (online) 2015)

Loop diuretics	**All patients**	**Regular monitoring of electrolytes**	**(British National Formulary (online) 2015)**
		Serum potassium 4-6 weeks after initiation, then following dose change	(Scottish Intercollegiate Guidelines Network 2001)
		Serum potassium once in 12 months	(Primary Care Quality & Information Service 2008)
		U&E once in 6 months	(Erewash 2006)
		Urinalysis once in 12 months, U&E after dose change	(North West 2002)
	Heart failure	U&E once in 6 months	(National Institute for Health and Care Excellence 2011) (London and South East Medicines Information Service 2008)
Amiodarone	All patients	TFT once in 6 months	(British National Formulary (online) 2015) (Primary Care Quality & Information Service 2008) (Erewash 2006) (North West 2002)
		TSH every 6 months up to 12 months after discontinuation	(British Thyroid Association 2006)

3.5 Review of studies of completion of monitoring

3.5.1 Aims and objectives

This section of the review examines previous research on the practice of monitoring for the prevention of ADEs in primary care for the medications included in this study. It examines the reported estimates of compliance to guidelines for monitoring, with a particular focus on follow-up monitoring during ongoing therapy. The objectives of the review are as follows:

- Summarise the reported estimates of the proportion of completed recommended monitoring of ACEI, loop diuretic and amiodarone therapy
- Identify any previous studies investigating the effect of monitoring on the risk of ADEs
- Comment on the quality of identified studies and differences in the methods used to identify and classify monitoring errors

- Examine the factors associated with the conduct of monitoring reported in studies

3.5.2 Review

Section 3.4 reported that published guidelines for monitoring in the UK agreed on the need to monitor all patients during ongoing therapy with ACEI, loop diuretics and amiodarone. It also demonstrated a number of deficiencies with regard to the content of these guidelines and highlighted a lack of good evidence to support recommendations.

In order to examine the evidence relating to the conduct of monitoring and effect of monitoring on the risk of ADEs in the context of the medications in the study, a literature review was performed in order to identify relevant studies on the subject published between 1990 and 2014. The review distinguishes between studies of monitoring conducted at initiation of therapy and those on monitoring during the course of therapy. Studies of initiation monitoring were included in order to better assess the predictive factors for non-compliance to guidelines in monitoring overall. The search was conducted using a MeSH search builder by applying the following terms: 'medication'; 'monitoring'; 'laboratory test'; this was added to search terms for specific medications: 'angiotensin converting enzyme inhibitor'; 'loop diuretic'; 'thiazide diuretic'; 'amiodarone'; 'cardiovascular'. The terms were used either alone or in combination in order to find studies from relevant areas. A manual search of abstracts was carried out and, for studies which were deemed relevant, the full text was obtained in order to extract information on study results and methods used identify monitoring tests..

The immediate information which was searched for in the abstract was the type of monitoring which the study focussed on: initiation or follow-up. The next important piece of information was whether ADEs were an outcome, as the link between monitoring and ADE outcomes was of interest in this study.

The design and methods employed in each selected study were assessed for the purpose of a) assessing the quality of the study and thus the reliability of its conclusions; and b) to identify the various methods for the identification of monitoring error and heterogeneity in study design across studies, which may contribute to differences in estimates of monitoring completion across studies.

3.5.3 Summary of studies

The review identified 13 studies published between 2005 and 2012 which satisfied the inclusion criteria. Three studies were UK-based and 10 were conducted in the USA. Retrospective cohort design was used in 7 studies, 3 studies involved a pre-post intervention, 2 used an RCT and one study employed a cross-sectional design. The main design characteristics and the findings from the studies are summarised in Tables 3.2 and 3.3.

Type of monitoring:

The studies identified in the review made a distinction between monitoring carried out prior to or shortly after the start of therapy (initiation monitoring), and testing carried out at regular intervals during therapy (follow-up monitoring). Two studies focused specifically on monitoring at initiation of therapy (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005), 6 studies looked into follow-up monitoring only (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, Matheny, Sequist et al. 2008, Avery, Rodgers et al. 2012, Bundy, Marsteller et al. 2012, Akbarov, Kontopantelis et al. 2015), whilst the 3 remaining studies investigated both types of monitoring (Johnson, Canty et al. 2010, McDowell, Coleman et al. 2010, Tjia, Fischer et al. 2011).

Studies applied specific criteria to identify patients who required either initiation or follow-up monitoring and assessed the patterns of testing at the corresponding stage of monitoring. For example, Simon et al identified a cohort of patients who were newly initiated onto a course of therapy with a drug which is indicated for initiation monitoring in guidelines. This study assessed the presence of laboratory tests in the period 180 days prior to start of therapy up to 14 days after initiation (Simon, Andrade et al. 2005). Raebel et al focused specifically on follow-up monitoring and applied rigorous methods to ensure that patients included in the analysis are long-term users of the study drug in order to avoid mistakenly including patients receiving initiation monitoring for a new course of therapy (Raebel, McClure et al. 2007). The authors set the inclusion criteria of a minimum of 12 months of continuous prescribing of the drug with tolerated gaps of recorded prescriptions of maximum 60 days.

An alternative approach by Hurley et al involved a "look back" period of 6 months prior to the observed period during which prescription records were assessed for exposure to the study drug, but not included in the assessment of outcomes (Hurley, Roberts et al. 2005). This was done to ensure that those patients whose

exposure to regular monitoring was observed during the study period were on prevalent therapy. Some studies classified initiation and follow-up monitoring as any testing which is undertaken before and after initiation, respectively (McDowell, Coleman et al. 2010). However, monitoring is likely to be more intensive in the first few weeks of therapy as the response to the drug and the risks of ADE in individual patients are assessed. This may lead to an overestimation of monitoring if testing shortly after initiation is not carried on in the form of regular follow-up monitoring during long-term therapy.

Summary of findings from individual studies:

The proportion of patients with completed initial monitoring varied between 47-87% for ACE inhibitors (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005, McDowell, Coleman et al. 2010, Tjia, Fischer et al. 2011), 42-64% for amiodarone (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005, Tjia, Fischer et al. 2011) and 47-61% for loop diuretics (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005, McDowell, Coleman et al. 2010, Tjia, Fischer et al. 2011). Follow-up monitoring was completed in 28-95% in ACEI therapy (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, Matheny, Sequist et al. 2008, McDowell, Coleman et al. 2010, Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015), 36-89% in loop diuretics (Matheny, Sequist et al. 2008, McDowell, Coleman et al. 2010, Avery, Rodgers et al. 2012, Bundy, Marsteller et al. 2012) and 20-70% in amiodarone therapy (Raebel, Chester et al. 2006, Johnson, Canty et al. 2010, Tjia, Fischer et al. 2011, Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015).

Comment on study design:

There was a significant heterogeneity in terms of study design and methods used in published literature on follow-up monitoring of the medications included in this study. It included retrospective cohort studies (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, McDowell, Coleman et al. 2010, Tjia, Fischer et al. 2011, Akbarov, Kontopantelis et al. 2015), pre-post intervention studies (Johnson, Canty et al. 2010, Bundy, Marsteller et al. 2012) and randomised controlled trials (RCTs) (Matheny, Sequist et al. 2008, Avery, Rodgers et al. 2012). The review included studies which both aimed to describe monitoring compliance and to estimate the effect of an intervention to improve monitoring. In case of intervention studies, this review looked at the control arm in order to gauge monitoring error rates under standard care conditions (Matheny, Sequist et al. 2008, Avery, Rodgers et al. 2012). Included studies ranged from relatively

small, single-site investigations to large analyses of multiple facilities or electronic patient records; the range of sample size of the studies selected in this review was 416 to 52906 patients.

Sources of data: electronic health records:

The inclusion criteria specified that all studies in the review had to be based in the primary care setting. All studies included in the review opted to use electronic health records to analyse monitoring. This may be explained by several factors: analysis of monitoring requires long-term follow-up in order to observe the conduct of monitoring during established therapy, EHRs include systematically coded data which allows the use of code lists to identify prescriptions of high-risk medications and laboratory tests conducted during therapy and EHR systems are a common platform for decision support software used for interventions to improve monitoring, which are described further in Section 3.6. There were differences in how studies obtained information on the compliance to recommended monitoring. Johnson et al observed monitoring tests recorded in electronic health records to test the effectiveness of an intervention to improve amiodarone TFT monitoring in 905 patients using an observational study design (Johnson, Canty et al. 2010). Bundy et al used the same method to observe monitoring in a cohort of 2013 patients on regular therapy with medications indicated for follow-up monitoring including ACE inhibitors and loop diuretics (Bundy, Marsteller et al. 2012). Two studies involved an RCT to test the effectiveness of interventions to improve medication safety and used EHRs to identify monitoring compliance. Avery et al collected electronic health records from 72 general practices in three English counties and identified 9071 patients on ACEI/ARB/diuretic and 493 patients on amiodarone therapy at risk of a monitoring error according to the study criteria applied at baseline (Avery, Rodgers et al. 2012). Matheny et al tested an electronic decision support system in 2790 patients on ACEI and 2508 on loop diuretics by observing monitoring tests through the electronic health record (Matheny, Sequist et al. 2008).

Other studies selected in this review were descriptive analyses of monitoring compliance which used electronic health records from primary care. A UK-based study used electronic data from 50 general practices to assess the feasibility of an active monitoring system using routine electronic records in primary care in Salford, England (Akbarov, Kontopantelis et al. 2015). The authors included previously developed prescribing and monitoring indicators, which included monitoring of ACE inhibitors (presence of U&E test in the previous 15 months in

patients ≥75 years old on regular therapy, N=8123) and amiodarone (presence of thyroid function test in previous 6 months in patients on regular therapy, N=271). Another study of monitoring within English general practice used EHRs from the Clinical Practice Research Datalink (CPRD) database to investigate compliance of GPs to both baseline and follow-up monitoring of a set of cardiovascular medications (McDowell, Coleman et al. 2010). Even with restrictive inclusion criteria, the study achieved large sample sizes for each medication cohort: ACEI (N=11245), ARBs (N=1630) and loop diuretics (N=416). Similar studies were conducted in the US: Simon et al employed electronic administrative data linked to a laboratory to estimate monitoring compliance in ACE inhibitors (N=20445), loop diuretics (N=35707) and amiodarone (N=1161), amongst other medications (Simon, Andrade et al. 2005). Hurley et al used claims data from a US health maintenance organisation (HMO) to identify monitoring in a large population of patients on regular ACE inhibitor therapy (N=12382 to 14843 across a 3 year period) (Hurley, Roberts et al. 2005). Raebel et al used a similar approach to access claims data in a database containing EHRs of over 2 million patients between 1999 and 2001 (Raebel, Lyons et al. 2005). The study obtained large sample sizes for each medication cohort (ACEI, N=48682; diuretics, N=78903; amiodarone, N=1551). The authors used the same approach to investigate follow-up monitoring of ACE inhibitors in 52906 patients on established therapy with the drug. Tjia et al used a retrospective design to identify compliance to both baseline and follow-up monitoring in 4703 patients prescribed medications indicated for monitoring, which included ACEI, loop diuretics and amiodarone (Tjia, Fischer et al. 2011).

It is not always possible to guarantee the consistency and reliability of information recorded in EHRs, which is one of the potential limitations of using this method for research in pharmacoepidemiology. It is rarely possible or feasible to verify the accuracy of recorded information due to several reasons, including lack of access to original patient files, requirements for ethics approval and the size of the dataset. If access to the original patient file is available, it is possible to cross-check the information recorded on the EHR against patient medical records. This was the case in the study by Raebel et al included in this review, which assessed the quality of electronic data against medical records in a sample of patients in a subset of drug-monitoring pairs included in the study, which included amiodarone and thyroid function tests (Raebel, Lyons et al. 2005).

Drugs assessed

There was no single standard method for the selection of medications for analysis across studies in the review, although many studies included all three study medications despite the different methods used in the selection process. All studies measured monitoring compliance separately for each drug, recognising that recommendations for monitoring and required tests vary by medication. Some studies chose to focus on medications from a single drug class, such as biochemical monitoring in cardiovascular medications or specifically compliance to recommended monitoring in amiodarone (Johnson, Canty et al. 2010, McDowell, Coleman et al. 2010). In other studies, medication selection was based on their inclusion in practice guidelines for monitoring. For example, one study included medications which were listed as requiring regular monitoring by the American Heart Association and American College of Cardiology (Simon, Andrade et al. 2005). Another study additionally considered guidelines published by the American Diabetes Association and the National Kidney Foundation and included medications based on the level of harm which may arise from inadequate follow-up monitoring (Matheny, Sequist et al. 2008). Hurley et al and Raebel et al used a more sophisticated medication selection method by looking into Black Box warnings listed in the Physicians Desk Reference as well as published guidelines, and used feedback from physicians and pharmacists to validate the medication list (Hurley, Roberts et al. 2005, Raebel, Lyons et al. 2005). Other factors taken into account in medication selection was predominant use of the drug in the primary care setting, previous evidence of poor compliance to monitoring guidelines and association with drug-related hospitalisations (Tjia, Fischer et al. 2011).

A different approach taken by Avery et al considered drug-monitoring pairs which represented difficulties in medication management in primary care with the potential to lead to severe harm to patient health, and issues which are detectable on a general practice electronic system. This method was used with the view of creating specific guidelines which would be adaptable within existing general practice systems and enable the intervention to be targeted where it is most likely to improve medication safety (Avery, Rodgers et al. 2012). The same method was used by Akbarov et al to assess the rate of monitoring error in ACEI and amiodarone therapy in 50 English general practices (Akbarov, Kontopantelis et al. 2015).

Criteria for identifying monitoring error

All studies included a description of how they determined inappropriate monitoring. Methods used to identify the presence of error varied widely across studies. Specifically, in the case of ACE inhibitors and loop diuretics it included absence of U&E (Avery, Rodgers et al. 2012, Bundy, Marsteller et al. 2012), serum potassium, sodium and creatinine (McDowell, Coleman et al. 2010), serum creatinine and potassium (Raebel, McClure et al. 2007), either serum creatinine or potassium (Tjia, Fischer et al. 2011) or serum potassium alone (Matheny, Sequist et al. 2008). In the case of amiodarone, all studies which listed this medication assumed that a TFT monitoring test is required, but there was no consistency on the definition of TFT. Across studies in this review, it was defined as the presence of TSH, triiodothyronine (t3), or free thyroxine (t4) (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005) or TSH only (Johnson, Canty et al. 2010, Tjia, Fischer et al. 2011).

Studies differed in their approach to identifying missed tests: some looked through records over a certain time period following initiation (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, McDowell, Coleman et al. 2010, Tjia, Fischer et al. 2011), others at overdue monitoring over a time period preceding a fixed date (Matheny, Sequist et al. 2008, Avery, Rodgers et al. 2012, Bundy, Marsteller et al. 2012) and one study looked at regular points during therapy (around 0, 6 and 12 month point after initiation) (Johnson, Canty et al. 2010).

There was a broad consensus across studies regarding the required frequency of testing of ACEI, loop diuretics and amiodarone. However, some authors chose to add additional time to this interval in order to take into account missed tests or delays in the reporting of tests by linked laboratory electronic systems. As a result, the time interval over which studies assessed the presence of monitoring was 12 months (Hurley, Roberts et al. 2005, Matheny, Sequist et al. 2008, Bundy, Marsteller et al. 2012), 13 months (Raebel, McClure et al. 2007) or 15 months (Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015). One study only looked at follow-up monitoring between months 3 and 6 after initiation (Tjia, Fischer et al. 2011). The four studies which looked at amiodarone monitoring also used different time intervals: tests recorded over a 6 month period (Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015), strict 6-monthly intervals (Johnson, Canty et al. 2010), and tests between months 2 and 7 of therapy (Tjia, Fischer et al. 2011).

Other design aspects: identifying prevalent therapy

Some studies applied specific conditions in order to identify patients at a specific point during the course of therapy. Monitoring requirements vary according to the stage of therapy: patients newly introduced to a high-risk drug often require frequent monitoring to ensure tolerability to a drug whilst those on longer term therapy may not need to be observed as often, as the risks of therapy have been adequately assessed and the outcomes of therapy are less uncertain. A common strategy was to identify patients who have been newly prescribed, defined by the absence of a prescription from the same drug class in the preceding 6 months (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005, Tjia, Fischer et al. 2011) or 12 months (McDowell, Coleman et al. 2010), in order to identify patients who require initiation monitoring according to guidelines.

Within studies of follow-up monitoring, an important aspect of study design was to identify prevalent therapy and ensure that those patients who were identified as being on the drug have not been discontinued and thus no longer require regular monitoring. This was ensured by setting inclusion criteria to include patients with a record of continuous prescribing of a study medication in the EHR, such as in Avery et al (Avery, Rodgers et al. 2012). Hurley et al included patients with a continuous record of prescriptions during the 12 month follow-up period and 6 months prior to the index date in order to exclude patients who had recently begun therapy and thus required initiation, rather than follow-up monitoring (Hurley, Roberts et al. 2005). The authors defined continuous use as a minimum of 6 months' supply of medication over the duration of the observed period, with a prescription of the drug at the end of the period to prevent misclassification of patients who had been discontinued early. Another study identified continuous use as prescription of the study medication for 6 out of 12 previous months (Bundy, Marsteller et al. 2012).

In their study of ACEI follow-up monitoring, Raebel et al restricted the study cohort to patients with a record of continuous prescribing of the drug for a minimum of 12 months (Raebel, McClure et al. 2007). The authors excluded any patient with a gap between prescriptions lasting more than 1.5 times the dispensed days' supply. McDowell et al identified discontinuation of therapy as gaps in prescribing of at least 30 days (McDowell, Coleman et al. 2010).

Some studies did not mention specific methods used to identify early discontinuation (Tjia, Fischer et al. 2011), thus it cannot be guaranteed that all patients who were identified as having an error were still on the drug and

required regular laboratory testing. In addition, few studies which looked specifically at follow-up monitoring ensured that included patients are not required to undergo initiation monitoring, which is more frequent than monitoring during established therapy. This may have overestimated monitoring completion rates reported by these studies.

Identifying factors associated with monitoring:

Several studies of monitoring reported estimates of the effect of different patient factors on completion of recommended monitoring tests. Akbarov et al used a mixed-effects logistic regression model and reported that monitoring errors were less likely to occur in older patients (OR 0.38-0.46 for different age groups compared to a baseline group aged 18-50), patients co-prescribed 8-10 medications (OR 0.53; 95% C.I. 0.3,0.93) and >10 medications (OR 0.36; 0.21,0.64) compared to a baseline of 0 or 1 co-prescriptions, male compared to female patients (OR 0.75; 0.64,0.88) (Akbarov, Kontopantelis et al. 2015). Simon et al fitted a logistic regression model to find that monitoring errors at initiation of therapy were more frequent in women than men in diuretics (OR 1.25; 95% C.I. 1.2,1.31) and more frequent in men than women in amiodarone thyroid function monitoring (OR 1.31; 1.02,1.67). The authors also found that older patients were more likely to be monitored appropriately compared to younger age groups in the cohort, but did not report actual odd ratios obtained (Simon, Andrade et al. 2005). They also reported that patients on diuretics with 6 or more comorbidities compared to those with 0 or 1 comorbidities were more likely to have an error (OR 3.63; 3.33,3.96). In patients prescribed amiodarone, no association was found between multimorbidity and monitoring. Raebel et al fitted a logistic regression model in their study of follow-up urea & electrolyte monitoring of ACE inhibitor therapy (Raebel, McClure et al. 2007). They reported a statistically significant effect of several patient factors on conducted monitoring at the 5% significance level: female patients compared to male patients (OR 1.06; 1.02,1.11), and older patients (different patient groups had an OR ranging 1.15-2.1 compared to a baseline of <50 years old). There were other factors associated with appropriate monitoring: number of health service contacts during the study period (outpatient visits: OR 1.46, hospitalisations: OR 1.15); patients with certain diagnoses (heart failure, CKD or diabetes) and those co-prescribed certain drugs (digoxin, potassium supplements or diuretics) were also more likely to receive U&E tests during therapy. Bundy et al reported that the number of co-prescribed medications was independently associated with regular monitoring but failed to report adjusted OR estimates (Bundy, Marsteller et al. 2012).

Outcomes used in the studies

The primary outcome used in most of the studies identified in the review was the proportion of patients with completed monitoring tests. Although the main purpose of medication safety monitoring is to prevent the occurrence of ADEs, the evidence of the effect of monitoring on the risk of ADE remains sparse. A systematic review of studies of amiodarone monitoring published between 1966 and 2000 found no studies which compared patients with and without recommended monitoring of amiodarone in terms of the incidence of ADEs (Stelfox, Ahmed et al. 2004). The review conducted in this study found few studies which have investigated the effectiveness of regular amiodarone monitoring at preventing ADEs.

A large retrospective cohort study of over 27,000 patients on ACE inhibitor or spironolactone investigated the effect of regular monitoring of serum potassium during the first year of therapy on the incidence of serious hyperkalaemia. Presence of hyperkalaemia was defined as serum potassium≥6mmol/L or ICD-9 code for hyperkalaemia in combination with an emergency department visit or hospital admission. A monitoring error was identified as the absence of a record of serum potassium during the first 12 months of therapy. The authors controlled for confounding arising from differences in characteristics of patients with or without regular monitoring by applying inverse probability weights based on the propensity score of receiving appropriate monitoring. Amongst patients prescribed ACEI either alone or in combination (n=26794), the unadjusted RR for severe hyperkalaemia was 1.38 (95% C.I. 0.9,2.14) (131/18937 patients with ≥1 tests vs. 24/7857 patients with no tests), but after adjusting for confounding the RR was 0.5 (0.37,0.66). In patients with chronic kidney disease stage 3 or 4 the adjusted RR was 0.19 (0.11,0.36), whilst in patients with no CKD, no statistically significant difference was found (Raebel, Ross et al. 2010). The study concluded that regular monitoring of potassium levels in ACEI therapy is only justified in patients with CKD, whilst extending monitoring guidelines to all patients may be ineffective.

A UK-based retrospective study of monitoring of antihypertensive medications using records of 74096 patients investigated the compliance of GPs to both baseline and follow-up monitoring guidelines using CPRD. The authors divided monitoring into two stages: baseline monitoring, which was defined as the presence of serum creatinine, urea, potassium or sodium tests in patient

electronic records in the 365 days prior to initiation of therapy and 2 days after; follow-up monitoring, which was defined as the presence of the above tests in the 6 months following index date. The authors reported that 47.2% of all patients received required baseline monitoring, which varied by medication: 53.8% for ACEi (n=11245), 35.3% for ARBs (n=1630) and 35.8% for loop diuretics (n=416) (McDowell, Coleman et al. 2010). The authors identified ADEs using Read diagnosis codes and abnormal laboratory test results using the following definitions: hyperkalaemia was defined as serum potassium ≥5.5mmol/L, hypokalaemia as serum potassium <3.5mmol/L and hyponatraemia as serum sodium ≤130mmol/L. The authors additionally assessed the effect of monitoring on hospitalisation and death. The vast majority of ADEs (88%) were found through abnormal monitoring tests; therefore it was difficult to measure the effect of monitoring on the rate of ADE due to the inability to observe ADE outcomes in patients with no record of laboratory tests during the study period. A Cox proportional hazards model was used to infer a link between monitoring and hospitalisation (HR 1.37, 95% C.I. 1.21,1.55). The authors noted that the apparent increased risk of hospitalisation in monitored patients may be explained by hospital admissions resulting from abnormal laboratory test results observed in the analysis. The authors noted that electrolyte ADEs were likely to be identified following laboratory tests during an unrelated investigation as most cases of hyperkalaemia identified in the cohort were asymptomatic. It was difficult to determine the indication for conducted tests, meaning that some tests may have been undertaken for reasons unrelated to drug therapy.

This study had a number of limitations. Although the sample size in the ACEI group was large, the number of patients included for other medications was comparatively small, given the potential sample size which could have been obtained from a large database. The small sample size was a result of restrictive inclusion criteria, which specified the inclusion of patients for whom drug therapy was indicated for hypertension. The exclusion of other indications, such as heart failure, may have missed out frail and elderly patients which were at greater risk of hyperkalaemia, which may explain the comparatively low estimate of compliance to follow-up monitoring guidelines.

Table 3.2: Summary of studies of initiation monitoring								
Author	**Year of publication**	**Design**	**Setting**	**Country**	**Medication**	**N**	**Proportion of patients with monitoring**	**Error identification method**
McDowell et al	2010	Retrospective cohort	Primary care	UK	ACEI	9071	0.472	Absence of test in 365 before index date and 2 days after in electronic record
					loop diuretics	493	0.472	
Simon et al	2005	Retrospective cohort	Primary care	US	ACEI	20445	0.667	Absence of creatinine or potassium test in 6 months prior or 14 days after in electronic data
					loop diuretics	35707	0.667	
					amiodarone	1161	0.423	Absence of TFT or LFT in 6 months before or 14 days after initiation
Raebel et al	2005	Cross-sectional	Primary care	US	ACEI	48682	0.608	Absence of serum creatinine or potassium in 180 days before or 14 days after initiation
					diuretics	78903	0.605	
					amiodarone	1551	0.428	Absence of TFT or LFT in 180 days before or 14 days after initiation
Johnson et al	2010	Pre-post intervention	Primary care	US	amiodarone	905	0.490	Absence of TFT (TSH) at initiation of therapy using electronic record
Tjia et al	2011	Retrospective cohort	Primary care	US	ACEI	1757	0.873	Absence of serum creatinine or potassium in 180 days prior or 14 days after initiation
					amiodarone	55	0.636	Absence of TSH test in 180 days prior or 14 days after initiation

Table 3.3: Summary of studies of follow-up monitoring								
Author	**Year of publication**	**Design**	**Setting**	**Country**	**Medication**	**N**	**Proportion of monitored**	**Error identification method**
Avery et al	2012	RCT	Primary care	UK	ACEI/loop diuretics	9071	0.886	Absence of U&E in previous 15 months for ACEI/diuretics using electronic data
					amiodarone	493	0.511	Absence of TFT in previous 6 months for amiodarone using electronic data
McDowell et al	2010	Retrospective cohort	Primary care	UK	ACEI	11245	0.538	Absence of serum potassium, sodium and creatinine in first 6 months of therapy using CPRD
					loop diuretics	416	0.358	
Hurley et al	2005	Retrospective cohort	Primary care	US	ACEi (potassium)	12382 to 14843	0.58-0.62	Absence of serum potassium in 12 month period using clinic and pharmacy claims data
					ACEI (creatinine)	12382 to 14843	0.62-0.68	Absence of serum creatinine in 12 month period using clinic and pharmacy claims data
Raebel et al	2007	Retrospective cohort	Primary care	US	ACEI	52906	0.675	Absence of serum creatinine and potassium in 13 month period using electronic data
Matheny et al	2008	RCT	Primary care	US	ACEI	2790	0.5	Absence of serum potassium in previous 12 months using electronic data
					loop diuretics	2508	0.517	
Johnson et al	2010	Pre-post intervention	Primary care	US	amiodarone	905	0.7	Absence of TFT (TSH) test at 6 months or 12 months after initiation using electronic record
Bundy et al	2012	Pre-post intervention	Primary care	US	ACEI/loop diuretics	2013	0.58	Overdue monitoring of U&E in previous 12 months using electronic record
Tjia et al	2011	Retrospective cohort	Primary care	US	ACEI	1757	0.281	Absence of serum creatinine or potassium in weeks 3-6 after initiation using electronic record
					amiodarone	55	0.2	Absence of TSH test in months 2-7 using electronic record

3.5.4 Comment on study design and methods

Findings from studies selected in this review found a positive association between a number patient factors and the conduct of monitoring. These include increasing age (Simon, Andrade et al. 2005, Raebel, McClure et al. 2007, McDowell, Coleman et al. 2010), number of co-morbidities (Raebel, Lyons et al. 2005, Simon, Andrade et al. 2005, McDowell, Coleman et al. 2010), sex (male) (Simon, Andrade et al. 2005, McDowell, Coleman et al. 2010) or female (Raebel, McClure et al. 2007), presence of CKD, diabetes or heart failure (Raebel, McClure et al. 2007), smoking status (McDowell, Coleman et al. 2010), economic deprivation (McDowell, Coleman et al. 2010) and recent hospital admission (Raebel, McClure et al. 2007). This suggests that the GPs take into account the health of the patient and presence of potential risk factors for ADE in their decision to conduct regular monitoring. Older patients with multiple comorbidities and specific conditions were more likely to receive appropriate monitoring.

The ordering and completion of recommended monitoring is just one component of a correct monitoring programme. The act of ordering a test, in itself, would not prevent an ADE. Failing to adequately respond to abnormal results can expose a patient to an ADE. Despite evidence suggesting that over a third of preventable ADEs are linked to inadequate response to laboratory test results (Gurwitz, Field et al. 2003), no comprehensive studies of monitoring response errors and their impact of the risk of ADE in primary care were found as part of this review.

A number of explanations for the poor uptake of monitoring in primary care have been presented in the literature. Shared responsibility for therapy by more than one physician or the inability to track patient adherence to ordered tests may lead to poor monitoring. Some GPs may choose not to order tests due to a perceived low risk of ADE (Goldman, Soran et al. 2010). GPs may fail to respond to abnormal laboratory tests if they lack a good understanding of the reasons for monitoring and conduct tests in order to comply with protocol (Clague, Twum-Barima et al. 1983). Other reasons relate to the quality of guidelines: specifically, the lack of specific information in guidelines, such as risk factors which merit extra attention and necessary action in response to an abnormal test result, and lack of supporting evidence for the effectiveness and cost-effectiveness of the recommended regimen (McDowell and Ferner 2011).

One issue which has hindered research in monitoring failures is lack of a working definition of monitoring error in the medication safety literature. One study involving a survey of 32 primary care professionals defined a monitoring error as

failure to carry out monitoring of a prescribed medicine in a way which would be considered acceptable in routine general practice, with 70% average approval from the respondents across the medications included in the survey (Alldred, Standage et al. 2008). They defined acceptable practice as monitoring according to recommendations set out in guidelines. However, this definition assumes consistency in routine general practice and broad agreement on the parameters of monitoring, such as required laboratory tests and frequency of monitoring, across prescribing physicians. A lack of understanding and agreement of what may be considered appropriate monitoring during routine therapy in published guidelines outlined in Section 3.4 resulted in discordance of monitoring practices: an audit and questionnaire study of monitoring in UK general practices reported that only 35% of GPs conducted regular monitoring of renal function in ACEI therapy (Kalra, Kumwenda et al. 1999).

All studies estimating the rate of compliance to recommended monitoring included in this review identified the presence or absence of a potential monitoring error based on the presence of a laboratory test within a fixed time interval. This analytical design may be restrictive in certain cases. The choice of length of the observed follow-up is arbitrary in most cases, which may have resulted in missed tests which were carried out beyond the observed period. In addition, this design does not allow assessment of the frequency of monitoring during ongoing therapy, meaning the authors could not ascertain which monitoring regimen was applied to the study patients, and what potential impact it had on the safety of therapy. This design is not appropriate for a study examining the effect of monitoring on patient outcomes as the typical length of follow-up (6 months – 1 year) is too short to capture the effect of monitoring on the risk of ADE.

3.6 Review of studies evaluating interventions to improve medication monitoring

Recognising the problem of inadequate monitoring of high-risk medications in primary care, a number of interventions to improve monitoring have been proposed and evaluated. Most of the interventions involved information technology or additional support for GPs in the medication management process.

3.6.1 Objectives of review

i. Identify studies which have evaluated interventions to improve compliance to monitoring guidelines in primary care.
ii. Examine the effectiveness and cost-effectiveness of interventions reported in studies.
iii. Assess the quality of the evidence from these studies.

3.6.2 Search strategy

A literature review was performed in order to identify relevant studies which evaluated interventions to improve medication monitoring published from 1960 to 2014. The search was conducted using a MeSH search builder by applying the following terms: 'medication'; 'monitoring'; 'laboratory test'; this was combined with search terms relating to an intervention: 'trial'; 'intervention'; 'control'; 'evaluation'. The terms were used alone or in combination in order to find studies from relevant areas. Additionally, those studies identified in the review outlined in Section 3.5 which involved an intervention were included. A manual search of abstracts was carried out and for studies which were deemed relevant the full text was obtained in order to extract information on the type of intervention, study medications and the effect of the intervention on the proportion of completed monitoring.

The review selected 12 eligible studies which have evaluated one or more interventions to improve monitoring. This included 8 randomised controlled trials and 4 evaluations with a pre-post intervention design. Nine studies evaluated improvements in follow-up monitoring, whilst three focus on monitoring at initiation of therapy. Ten of the twelve studies evaluated the effect on monitoring of several medications or several classes in a single healthcare facility. A summary of included studies can be found in Table 3.4.

3.6.3 Summary of studies of information technology interventions

Investment in electronic technology in the primary care setting has focused on the implementation of electronic health records (EHR) systems, which are longitudinal collections of patient-centric health care information available across providers, care settings and time, and computerised provider order entry (CPOE) systems, which facilitate the ordering of diagnostic and laboratory tests and the retrieval of results via an electronic platform (Black, Car et al. 2011). The systematisation of medication management via introduction of CPOE systems has reduced the rate of prescribing errors in primary care (Bates, Teich et al. 1999) and has created new opportunities for technological interventions in the area of medication monitoring.

Systems for storage and retrieval of routine electronic records can be used to create algorithms to support practitioner decision-making, known as computerised decision support systems (CDSS). In the context of monitoring, EHR, CPOE and prescribing records can be linked to combine data on dosage, test results and patient risk factors to provide safety prompts, issue reminders for overdue monitoring and suggest possible actions in response to abnormal test results.

Several studies have assessed the effectiveness of CDSS interventions to improve monitoring of high-risk medications in primary care. Pharmacist-led CDSS was found to be effective at improving monitoring in several RCTs (Raebel, Lyons et al. 2005, Raebel, Chester et al. 2006, Avery, Rodgers et al. 2012). Pharmacist input was found to increase monitoring compliance compared to GP electronic reminders alone (Feldstein, Smith et al. 2006). On the other hand, several RCTs found no evidence of improvement in monitoring compliance after application of CDSS (Palen, Raebel et al. 2006, Matheny, Sequist et al. 2008, Lo, Matheny et al. 2009). A positive effect of electronic aids has been additionally observed in several pre-post studies (Hoch, Heymann et al. 2003, Stewart, Loftus et al. 2003, Steele, Eisert et al. 2005, Koplan, Brush et al. 2012). However, this improvement in monitoring does not necessarily translate into reduction in morbidity and mortality and improvement in quality of life, as demonstrated by a number of studies which failed to observe an effect of pharmacist-led interventions on final patient outcomes (Krska, Cromarty et al. 2001, Roberts, Stokes et al. 2001).

Ensuring the safe use of medicines is more challenging in primary care than in the inpatient setting due to a typically longer length of follow-up, more frequent drug-drug interactions and infrequent patient-provider contact (Budnitz and

Layde 2007). In addition, the process of monitoring is often data intensive, with the need to evaluate a large number of test results to inform a decision. Interventions to support medication monitoring have the potential to be effective as they represent a systematic assessment of a patient's medication safety needs using routinely recorded information and professional support from pharmacists, which adds an extra layer of safety to ensure that monitoring is carried out in a timely fashion and appropriate action is taken in case of an abnormal test result.

However, the potential of such interventions to improve the safety of care is often constrained by design issues. For instance, evidence suggests that "active" reminders which generate a prompt and require an action are more effective at improving monitoring than "passive" alerts which simply provide information or suggest an action without interrupting the work flow of the physician. However, "active" reminders are considered by some physicians as too invasive, which may cause them to be ignored or dismissed as disruptive, a phenomenon known as "alert fatigue" (Shah, Seger et al. 2006). This issue is particularly relevant when the signal threshold is low, which may cause as many as 88% of drug interaction and 69% of drug allergy alerts to be ignored (Payne, Nichol et al. 2002). The acceptance of active drug safety alerts can be improved by categorising alerts according to importance, with less important issues relegated to passive alerts (Shah, Seger et al. 2006).

Studies evaluating interventions to improve monitoring seldom address the fact that the effectiveness of a preventive strategy is often dependent on the extent to which it has been applied to the target population. Laboratory monitoring of medications, as with any prevention strategy, involves the commitment of physician time and the costs of carrying out the laboratory test. Additional costs of an intervention to improve monitoring may be incurred, such as the cost of implementing an electronic reminder system based on the CPOE system or pharmacist review. The benefits of monitoring are measured in terms of the number of prevented ADEs and associated loss of patient utility and additional treatment cost. These costs and benefits do not follow a linear relationship with the scale of the monitoring programme. Patients at the higher end of the risk scale for ADE are monitored first, followed by patients at lower risk for whom the benefits of regular monitoring may be lower. Extending monitoring to hard-to-reach patients will raise its average cost per patient. When these effects are combined, the expansion of the preventive strategy will exhibit a progressively lower marginal return to the resources used (Warburton 2005). As a result, improving the proportion of completed recommended tests from 70% to 80% is

not likely to be as effective or cost-effective as, say, a move from 50 to 60%, despite the fact that the absolute change in monitoring completion is the same.

In summary, IT-based interventions and pharmacist-led interventions have been shown to improve compliance to monitoring guidelines in primary care. Pharmacist input in combination with an IT intervention was shown to produce a larger effect than electronic reminders alone, suggesting the need for professional support in addition to CDSS. However, implementation of these interventions in practice may not be justified based on improvement in rates of monitoring alone, which represent a process outcome. Studies linking monitoring to incidence of ADEs and patient health and utility outcomes are necessary to demonstrate the superiority of these interventions relative to standard practice.

3.6.4 Studies which evaluated the cost-effectiveness of interventions to improve monitoring

The review identified two studies which assessed the cost-effectiveness of interventions to improve monitoring of one or more of the study medications. Elliott et al conducted a cost-effectiveness analysis alongside an RCT to evaluate a pharmacist-led IT intervention to reduce prescribing and monitoring errors (details of the clinical study found in Table 3.4, Avery et al study) (Elliott, Putman et al. 2014). The intervention was dominated in the case of ACEI and loop diuretic monitoring errors (higher cost and lower effectiveness); in contrast, the intervention was highly likely to be cost-effective at a ceiling willingness-to-pay of GBP 20,000 per QALY gained in the case of amiodarone (GP 515 per QALY, 100% of being cost-effective at ceiling WTP of GBP 20,000 per QALY).

The other economic evaluation found in the review was a cost-effectiveness of an electronic reminder system to improve monitoring of ACE inhibitor and ARB therapy (Smith, Raebel et al. 2011). At the base-case scenario, intervening in all patients on ACEI or ARBs wasn't found to be cost-effective. A subgroup analysis of patients with CKD found the intervention to be dominant, with a 95% probability of cost-effectiveness at the same WTP threshold.

Table 3.4: Summary of studies evaluating interventions to improve monitoring						
Author	**Year of publication**	**Design**	**Type of intervention**	**Type of monitoring**	**Drug**	**Effect of intervention**
Krska et al	2001	Randomized trial	Pharmacist medication review	Follow-up	12 drug classes, including ACEI and diuretics	reduction of monitoring issues by 70% compared to 14% in control group
Hoch et al	2003	Pre-post intervention study	Electronic physician reminder	Follow-up	Diuretics	absolute increase of 3%
Stewart et al	2003	Pre-post intervention study	Electronic reminder	Follow-up	Amiodarone	absolute improvement of ordering of follow-up thyroid function tests by 29%
Raebel et al	2005	Randomized trial	Electronic reminder and pharmacist outreach	Initiation monitoring	14 drugs, including amiodarone	9% absolute increase
Steele et al	2005	Pre-post intervention study	Electronic reminder	Follow-up within recommended interval for each drug	All medications due for monitoring, including ACEI	Improved test ordering by a third (12% absolute improvement) in the case of displayed messages; background messages did not have an effect
Feldstein et al	2006	Cluster randomized trial	Electronic physician reminder; patient telephone reminder; pharmacist patient outreach	Follow-up monitoring within 1 month	10 medications, including ACEi and diuretics	HR: EMR 2.5 (1.8-3.5); AVM 4.1 (3-5.6); pharmacist 6.7 (4.9-9)
Raebel et al	2006	Randomized trial	Electronic reminder and pharmacist outreach	Follow-up within recommended interval for each drug	9 drugs, including amiodarone	absolute increase of 6%
Palen et al	2006	Randomized trial	Electronic reminder	Follow-up within recommended interval for each drug	10 medications, including ACEi and diuretics	No statistically significant effect
Matheny et al	2008	Randomized trial	Electronic reminder	Follow-up within recommended interval for each drug	15 drugs, including ACEI and diuretics	No statistically significant effect
Lo et al	2009	Randomized trial	Electronic reminder	baseline	11 drug classes, including ACEI and diuretics	No statistically significant effect
Koplan et al	2012	Pre-post intervention study	Electronic reminder	initiation monitoring	Warfarin	absolute improvement of ordering of initiation monitoring by 5%
Avery et al	2012	Randomized trial	Pharmacist-led IT support	Follow-up	ACEI, diuretics and amiodarone	Reduction in monitoring errors in ACEI (OR 0.63, C.I. 0.41,0.95)
Elliott et al	2014	Cost-effectiveness analysis	Pharmacist-led IT support	Follow-up	ACEI, loop diuretics, amiodarone,	Dominated for ACEI monitoring (higher cost, lower QALYs generated), cost-effective in amiodarone TFT monitoring: £515/QALY, 100% probability of cost-effective at £20k/QALY
Smith et al	2011	Cost-effectiveness analysis	Electronic reminder for follow-up mponitoring	Follow-up	ACEI	Base case analysis: $522,000 per QALY, <1% prob. Of cost-effectiveness; subgroup analysis of CKD patients found dominance, 95% prob. Of cost-effectiveness at WTP $30,000/QALY

3.7 Discussion

3.7.1 What we know

This chapter described the appropriate practice of monitoring of ACE inhibitors, loop diuretics and amiodarone in the context of UK-specific published guidelines on the conduct of monitoring of these medications in primary care. This was followed by a review of published studies examining compliance to recommended monitoring of the study medications and a review of studies assessing the efficacy and cost-effectiveness of interventions to improve monitoring. National guidelines for monitoring have been published by a range of organisations, such as NICE (National Institute for Health and Care Excellence 2010, National Institute for Health and Care Excellence 2011), BNF, disease-specific organisations (British Thyroid Association 2006, Joint Specialty Committee on Renal Medicine of the Royal College of Physicians of London and the Renal Association 2006), health-service specific (Scottish Intercollegiate Guidelines Network 2001, Primary Care Quality & Information Service 2008) and other influential organisations in healthcare. The information in national guidelines was cited in local guidelines which defined specific parameters of monitoring, such as required test, frequency and action in case of abnormal test result designed to inform GPs of the best practice of monitoring patients prescribed these medications (North West 2002, Erewash 2006, London and South East Medicines Information Service 2008).

NICE lists ACEI as first-line treatment of hypertension in patients <55 years old and in chronic heart failure. ACEI and diuretics are listed as second-line treatment of resistant hypertension in patients of all ages. Whilst NICE is clear about the need to conduct renal function and electrolyte tests at initiation and following each dose change, the recommendations regarding regular monitoring during long-term therapy with ACEI and loop diuretics are less detailed. Monitoring tests should be repeated "as appropriate", with more careful monitoring in case of worsening renal function, elevated serum potassium levels or co-prescription with another drug which is likely to cause hyperkalaemia or adversely affect renal function. Regular monitoring is left at the discretion of the prescribing GP, although patients with heart failure should get a renal function test once in 6 months as part of chronic disease management, meaning that patients on ACEI treatment for heart failure should be offered regular monitoring by default.

There was little agreement on the appropriate frequency of regular monitoring of U&E and renal function during ACEI and diuretic therapy: BNF recommended

repeated monitoring during therapy (British National Formulary (online) 2015), with no details of the required frequency; Welsh guidelines recommended monitoring once every 12 months (Primary Care Quality & Information Service 2008), whilst Scottish guidelines (SIGN) did not specify the need for regular monitoring except following a dose change (Scottish Intercollegiate Guidelines Network 2001).

Several guidelines developed by guideline developing groups and specialist organisations such as the Royal College of Physicians and Renal Association listed the need for regular monitoring once in 6 or 12 months (Joint Specialty Committee on Renal Medicine of the Royal College of Physicians of London and the Renal Association 2006). All guidelines follow a similar pattern of cross-referencing to other guidelines and expert opinion, with no reference to published randomised or observational studies of monitoring. These recommendations have been cited in local guidelines developed by PCTs to guide the practice of monitoring in their local areas, which uniformly recommend monitoring once in 12 months, despite absence of credible evidence to support the need for regular monitoring in all patients on regular ACEI or diuretic therapy. Guidelines disagreed on the thresholds for hyperkalaemia and AKI and whether therapy should be discontinued or modified in the eventuality of an ADE.

There was more agreement on the required frequency of TFT monitoring during amiodarone therapy. National guidelines, including ones published by the BNF and British Thyroid Association, recommend the measurement of TFT once in 6 months to prevent AIH and AIT (British Thyroid Association 2006, British National Formulary (online) 2015). However, there were inconsistencies in how TFT was defined as it consists of a combination of three separate measurements. Local guidelines combined evidence from national guidelines with the drug manufacturer's own recommendations for thyroid monitoring and were in agreement regarding the need to conduct TFT monitoring every 6 months during therapy and in the 12 months following cessation of therapy due to the persistent risk of thyroid dysfunction. Specification of thyroid dysfunction, which tests should be carried out (TSH, t3 or t4) and what actions are required in case of an abnormal test result was vague. Similarly to ACEI and loop diuretics, guidelines made no reference to RCTs or observational studies to support their recommendations.

The monitoring of ACE inhibitors, loop diuretics and amiodarone in primary care has been described in a number of large, representative studies in the US and

UK. The proportion of included patients with missed monitoring varied substantially across studies, although all studies agreed that compliance to monitoring recommendations in ACEI, diuretics and amiodarone was poor. GPs frequently failed to undertake necessary tests at the beginning of therapy or during long-term treatment, even in patients with known risk factors for ADE. The review identified studies with a variety of study designs, including RCTs, retrospective and prospective cohort and cross-sectional studies. All studies used EHRs to observe monitoring, although there was substantial variation in how monitoring errors were defined and identified in electronic data. There was little consistency in key design aspects, such as length of each cycle used to observe monitoring tests, inclusion criteria for patients on newly initiated or established therapy and criteria used to identify discontinuation of therapy. These discrepancies in methodology made it difficult to compare findings across studies and attribute differences in monitoring compliance to a particular aspect of study design.

Several patient factors independently associated with the conduct of monitoring were identified in previous studies. These include age, sex, number of co-morbid conditions, number of co-prescribed medications, presence of specific conditions which increase the risk of electrolyte ADEs, such as CKD, diabetes and heart failure; specific medications which warrant close monitoring of renal function and urea & electrolytes, such as diuretics. This suggests that the approach to regular monitoring of patients who are older, have more concomitant conditions and drug therapies and have pre-existing conditions which may predispose them to a higher risk of ADE differ from those who do not present with these risk factors. Monitoring practice may have been influenced by published guidelines into practice for monitoring which explicitly indicate more careful monitoring in the presence of these risk factors. These observations may also be explained by risk profiling of patients by GPs in order to manage workload and avoid conducting tests when perceived risk of ADE is low. In the context of this study, it represents a risk of selection bias, as patients assigned to different monitoring regimens may not be directly comparable. This is the rationale for deriving information on patient characteristics, diagnoses and prescribed medications in order to control for effects which may confound the relationship between monitoring and the risk of ADE drawn from observational data. Methods for identifying covariates for the analysis are outlined in Chapters 4 and methods for controlling confounding in the statistical analysis are described in Chapter 5 of this thesis.

A number of explanations for the poor uptake of monitoring in primary care have been presented in the literature. Shared responsibility for therapy by more than one physician or the inability to track patient adherence to ordered tests may have led to missed tests. Some GPs may choose not to order tests due to a perceived low risk of ADE (Goldman, Soran et al. 2010). GPs may fail to respond to abnormal laboratory tests if they lack a good understanding of the reasons for monitoring and order tests in order to comply with protocol (Clague, Twum-Barima et al. 1983). Other reasons relate to the quality of guidelines: specifically, the lack of specific information in guidelines, such as risk factors which merit extra attention and necessary action in response to an abnormal test result, and lack of supporting evidence for the effectiveness and cost-effectiveness of the recommended regimen (McDowell and Ferner 2011).

Errors at the monitoring stage of medication therapy are poorly understood. In Chapter 1 of this thesis it was demonstrated that a medication safety monitoring strategy has the potential to be effective under a set of conditions, namely the presence of a measurable proxy marker which accurately tracks the patient's condition in a timely fashion, with a sufficiently high signal-to-noise ratio to allow the physician to identify patients with ADE and take appropriate corrective action. Even when these conditions have been met, heterogeneity of patients in terms of stage and complexity of disease, co-prescribed medications and risk factors represents a further challenge to prescribers. A monitoring strategy is a complex intervention involving decisions at multiple stages over an extended period of time, the performance of which is influenced by a large number of interacting factors which may not necessarily be within the prescribing physician's control. As such, it is a process which may entail significant scope for failure. A GP may fail to order a required monitoring test due to time constraints or insufficient understanding of the drug regimen and the associated risks. If a test is ordered, inadequate monitoring may occur due to suboptimal care, such as failure to respond to an abnormal laboratory test result. Alternatively, the reason for the failure may be due to patient non-adherence to an ordered test.

This approach to dealing with errors in the monitoring process fits conveniently into the human error model proposed by Reason, which hypothesises that errors in complex systems such as healthcare are rooted in systemic causes and views human error as a function of the underlying organisational process (Reason 1990). This theory has seen support within both the patient safety literature and policy. Within the context of this framework, interventions targeting safety systems (e.g. electronic reminders) and the workforce (e.g. greater involvement

of pharmacists or nurses in the monitoring process) can be seen as efforts to reduce the impact of human error in medication management by addressing systemic shortcomings.

However, it is important to recognise that the absence of a laboratory test may not necessarily constitute the presence of an actual error. Not all GPs agree with monitoring guidelines on which patients should be monitored and how frequently. For this reason, lapses in monitoring are often referred to as potential, rather than actual monitoring errors (Hurley, Roberts et al. 2005). This is in contrast to the current approach of studies of monitoring, which conclude that missed laboratory tests constitute an error or oversight (Simon, Andrade et al. 2005, Coleman, McDowell et al. 2010). It is clear, however, that in many cases monitoring was not carried out on purpose, due to a low perceived risk of harm. As a result, classifying such events as "errors" presumes that the action was guided by incorrect judgement. In many cases monitoring guidelines are supported by little or no evidence of the preventive effect of the monitoring strategy, the presumption of incorrect judgment is therefore subjective.

This pragmatic approach to monitoring may be illustrated by an example in the context of cardiovascular medication therapy: appropriate clinical practice at the initiation of a RAAS agent requires the measurement of renal function prior to beginning treatment in order to establish the risk of hyperkalaemia. In many cases, the elevated risk of hyperkalaemia due to poor renal function or diabetic nephropathy is outweighed by the cardiovascular benefit of drug therapy; treatment accompanied by close monitoring is the optimal strategy (Palmer 2004). In this case the risk of ADE persists throughout the length of therapy, justifying the use of follow-up monitoring in accordance to guidelines. However, in the absence of clear risk factors the probability of hyperkalaemia is significantly lower, and the logic behind the need to carry out monitoring is less clear. Physicians, working under clear time and resource constraints, are thus less likely to strictly adhere to monitoring guidelines in patients that they perceive to be low-risk.

This "human error"-based approach to dealing with monitoring errors is demonstrated in the design of interventions to improve monitoring of high-risk medications. It fails to recognise the possibility that the issue stems from a tendency by prescribing physicians to disagree with the recommendations in guidelines and scepticism towards the perceived benefits of the regimen. Section 3.6 demonstrated that pharmacist-supported IT-based interventions are more

effective than IT-based decision support alone at improving compliance to monitoring guidelines. The possible explanation is that a pharmacist provides a second opinion on the risks of medication therapy and the correct action to mitigate this risk, providing both professional support to their prescribing colleagues and educational intervention for physicians as they gain additional knowledge of the risks of the drug from the pharmacist. However, definitive judgment of whether a pharmacist-led intervention is an effective tool to manage drug safety during ongoing therapy cannot be made without evidence of reduction in the risk of ADE and benefit in terms of final patient outcomes.

Pharmacist input at the dispensing stage is a vital component of the medication safety system in the UK NHS which is designed to prevent prescribing and administration errors. The effectiveness of pharmacist input to reduce monitoring errors in primary care may be constrained by limited access to information about which patients require follow-up monitoring during therapy. Recognising this issue, NHS England and the Health and Social Care Intervention Centre have piloted a system which gives community pharmacists access to patient prescription records from general practices (Phramaceutical Journal 2014). The knowledge and expertise of a pharmacist may be similarly applied to improve the safety of medications during ongoing care. As part of the reconfiguration of primary care in England, additional numbers of clinical pharmacists will be employed in general practices in order to assist GPs in the medication management process (GP Online 2015). The knowledge of GPs and pharmacists in medication safety can be combined in order to improve the safety of ongoing medication management. The role of pharmacists in the medication management process extends beyond medication review; a recent study of errors in the hospital setting found that over half of the medication errors intercepted by pharmacists were as a result of consultative activities, whilst the proportion intercepted through medication review constituted 37% (Patanwala, Sanders et al. 2012), suggesting that communication between healthcare professionals is no less important in the medication safety process than the technical skill of the pharmacist. However, striving to reduce the risk of undesirable effects of medical care to zero may entail excessive use of resources which are better used elsewhere (Warburton 2005).

3.7.2 What we don't know

The review did not find a single study which looked at the effect of conducting follow-up monitoring at different frequency. All studies treated monitoring as a

binary exposure by comparing patients with and without monitoring during the follow-up period. The observation period ranged from 12 to 15 months in studies of ACEI and loop diuretics and 6 to 12 months in amiodarone, which was insufficient to observe frequency of monitoring during long-term therapy. An observational study with considerably longer follow-up period is required in order to observe whether GPs adhere to long-term monitoring goals. This gap in evidence makes it difficult to assess the effect of non-compliance to monitoring guidelines on the safety of medications, patient outcomes and cost in terms of scarce healthcare resources. This lack of evidence may have contributed to the lack of consensus on how frequently monitoring should be carried out during long-term therapy with ACEI, loop diuretics, amiodarone and other medications prescribed in primary care. GPs are less likely to adhere to a strict monitoring regimen over the full period of therapy which may last many years if they are sceptical about its benefits. Moreover, the lack of evidence regarding the costs and benefits of regular monitoring makes it difficult to justify the introduction of new interventions aimed at improving compliance to monitoring guidelines. This programme of research aimed to fill this evidence gap by assessing the effectiveness and cost-effectiveness of monitoring of these medications using routinely collected data from GP practices and hospitals across the UK.

The vast majority of studies of monitoring were limited to an assessment of inadequate monitoring and its determinants. Only two studies identified in the review have investigated the effect of monitoring of ACEI or amiodarone on ADE outcomes (McDowell, Coleman et al. 2010, Raebel, Ross et al. 2010). The review identified several studies which estimated the effect of interventions to improve monitoring, although only two economic analyses of monitoring involving the study medications were found (Smith, Raebel et al. 2011, Elliott, Putman et al. 2014). The evidence of the effect of monitoring on the risk of ADEs and the economic rationale for conducting regular monitoring remains sparse at best.

The ordering and completion of recommended monitoring is just one component of a correct monitoring programme. The act of ordering a test, in itself, would not prevent an ADE. Failing to adequately respond to abnormal results can expose a patient to an ADE. Despite evidence suggesting that over a third of preventable ADEs are linked to inadequate response to laboratory test results (Gurwitz, Field et al. 2003), no comprehensive studies describing the incidence of monitoring response errors and their impact of the risk of ADE in primary care were found as part of this review.

3.7.3 Design aspects

The literature review in this Chapter noted a lack of a working definition of monitoring error in the medication safety literature. One study involving a survey of 32 primary care professionals defined a monitoring error as failure to carry out monitoring of a prescribed medicine in a way which would be considered acceptable in routine general practice, with 70% average approval from the respondents across the medications included in the survey (Alldred, Standage et al. 2008). They defined acceptable practice as monitoring according to recommendations set out in guidelines. However, this definition assumes consistency in routine general practice and broad agreement on the parameters of monitoring, such as required laboratory tests and frequency of monitoring, across prescribing physicians. A lack of understanding and agreement of what may be considered appropriate monitoring during routine therapy in published guidelines outlined in Section 3.4 resulted in discordance of monitoring practices: an audit and questionnaire study of monitoring in UK general practices reported that only 35% of GPs conducted regular monitoring of renal function in ACEI therapy (Kalra, Kumwenda et al. 1999).

All studies estimating the rate of compliance to recommended monitoring included in this review identified the presence or absence of a potential monitoring error based on the presence of a laboratory test within a fixed time interval. This analytical design may be restrictive in certain cases. The choice of length of the observed follow-up is arbitrary in most cases, which may have resulted in missed tests which were carried out beyond the observed period. In addition, this design does not allow assessment of the frequency of monitoring during ongoing therapy, meaning the authors could not ascertain which monitoring regimen was applied to the study patients, and what potential impact it had on the safety of therapy. This design is not appropriate for a study examining the effect of monitoring on patient outcomes as the typical length of follow-up (6 months – 1 year) is too short to fully capture the effect of monitoring on the risk of ADE.

The reviews of literature described in this chapter drew attention to several methodological aspects which are important in an investigation of the effect of regular monitoring on the risk of ADE. The use of observational data to estimate the causal effect of monitoring needs to deal with confounding which arises from patient factors influencing both the chosen monitoring strategy and the potential

risk of ADE, meaning that comparisons of the risk of ADE across groups of patients assigned to different monitoring regimens may be confounded by other factors. An investigation of follow-up monitoring needs to separate patients newly prescribed the drug from prevalent users in order to make a distinction between initiation and follow-up monitoring during established therapy, which represent two distinct stages of monitoring conducted for different purposes. Previous studies have identified either newly initiated or prevalent therapy by looking up prescription records of the study drug in the 6 or 12 months prior to the index prescription (Hurley, Roberts et al. 2005, Tjia, Fischer et al. 2011). Follow-up monitoring was observed during long-term therapy, rather than in the first weeks after initiation. In addition, there should be a way to systematically identify patients who have been discontinued in order to avoid including patients who no longer require regular monitoring. Studies dealt with this issue by setting criteria for discontinuation, such identifying gaps in dispensed supply of over 30 days (McDowell, Coleman et al. 2010, Raebel, Ross et al. 2010).

3.7.4 Concluding remarks

The findings from the literature review outlined in this chapter suggest that a large proportion of long-term users of cardiovascular medications, which include ACE inhibitors, diuretics and amiodarone, do not receive appropriate follow-up monitoring in primary care, despite published recommendations which indicate the need for regular monitoring of these medications to prevent ADEs. In addition, although a number of studies in the UK and other countries have investigated the compliance to recommendations for monitoring of these medications in primary care, the review in this chapter found only two previously published studies which explicitly aimed to estimate the effect of regular monitoring on the risk of ADE during antihypertensive therapy. The review found no studies which investigated the effect of thyroid function monitoring on the risk of amiodarone-induced ADEs. The review identified two studies of the cost-effectiveness of regular monitoring of ACE inhibitors and amiodarone, although both studies relied on modelling methods to estimate the treatment effect of monitoring on the reduction of the risk of ADE, rather than observing the effect of monitoring from longitudinal data. Until a clear link between monitoring and the risk of specific ADEs using well-designed, population-level observational or randomised studies, GPs will remain sceptical about the need for regular monitoring and until there is clear evidence regarding the cost-effectiveness of strategies for regular monitoring, the case for rolling out new interventions to

improve monitoring will be weak. New evidence on the effectiveness and cost-effectiveness of monitoring based on data observed in real practice will contribute towards evidence-based guidelines, which would both ensure that GPs conduct monitoring in areas where it would effectively improve the safety of pharmacotherapy and reduce the wasteful use of scarce healthcare resources on unnecessary tests and investigations. The link between follow-up monitoring and final patient outcomes needs to be established before the evaluation and implementation of interventions designed to improve monitoring, which may prove to be unnecessary and wasteful if such a link is absent.

The analysis of previous studies of monitoring has identified several design aspects which are important in a study of monitoring. A reliable method for identifying monitoring regimens and potential cases of inappropriate monitoring is needed, as well as methods to ensure that prevalent therapy is correctly ascertained by identifying periods of discontinuation. This is particularly important in an observational study where patients may be frequently discontinued or moved between different treatment regimens. A study needs a systematic method for identification of ADE cases. In the case of an observational design with non-random assignment into comparison groups, steps must be taken to avoid the confounding effect of patient factors which may influence the decision to conduct regular monitoring. These aspects were incorporated into the study design of the analysis of monitoring in this programme of research. The manner in which exposures and outcomes were ascertained from CPRD and HES data is outlined in Chapter 4 of this thesis. The methods for controlling confounding arising from non-random assignment are handled in Chapter 5.

Chapter 4:

Methods for data extraction, manipulation and identification of exposures, outcomes and covariates for the analysis

4.1 Introduction to chapter

This chapter describes the aims of the time-to-event analysis of monitoring and outlines the methods of extraction and selection of a study cohort which was used to answer the research questions set out in Chapter 1 of this thesis. It sets out a justification for the use of observational data from large databases to carry out this research, offers a description of the capabilities of the Clinical Practice Research Datalink and Hospital Episode Statistics for pharmacoepidemiological research. It proceeds to outline the methods of extraction of electronic patient records and identification of the exposures, outcomes and covariates required for analysis.

4.2 Aims of analysis of the effect of monitoring frequency on risk of ADEs

The aim of this analysis was to investigate the effect of different frequencies of conducting routine monitoring of medications during long-term therapy with ACE inhibitors, loop diuretics and amiodarone in terms of the risk of specific adverse drug events (ADEs) linked to these medications.

The following monitoring drug-test pairs were selected for analysis:

i. Urea & electrolytes, serum potassium, sodium and creatinine and ACEI
ii. Urea & electrolytes, serum potassium, sodium and creatinine and loop diuretics
iii. Thyroid function tests and amiodarone

The following objectives were set for the analysis of monitoring and ADEs:

1. Conduct a descriptive analysis of monitoring frequency for each individual medication cohort. Assess differences in the practice of monitoring across patient groups with different characteristics.
2. Assess the effect of patient covariates on the choice of monitoring regimen within each medication cohort in order to design appropriate statistical tools to control for selection bias arising from non-random assignment into comparator groups with different frequencies of monitoring.
3. Specify a time-to-event regression model and estimate the effect of monitoring frequency on the following outcomes:
    - Hyperkalaemia, AKI during ACEI therapy
    - Hypokalaemia, hyponatraemia and fall-related injuries during loop diuretic therapy
    - AIH, AIT and major adverse cardiovascular event (MACE) during amiodarone therapy

4.3 Rationale for the study.

The research described in this thesis focused on a subset of cardiovascular medications which are associated with ADEs in patients and included ACE inhibitors (ACEI), loop diuretics and amiodarone. The decision to concentrate on these three medications was motivated by the following factors:

i. ACEI, loop diuretics and amiodarone were found to be amongst the top 10 medications associated with preventable drug-related hospital admissions (Howard, Avery et al. 2007).

ii. The benefits of regular monitoring have been previously discussed in the case of ACEI and diuretics (Coleman, McDowell et al. 2010, Raebel 2012) and amiodarone (Podrid 1995, Stelfox, Ahmed et al. 2004). In UK primary care, regular monitoring of these medications for prevention of ADEs is recommended in several published guidelines in primary care, which were described in detail in Chapter 2 of this thesis.

iii. There is limited evidence of the efficacy of regular monitoring programmes involving ACEI, loop diuretics or amiodarone and almost no evidence of their cost-effectiveness. There is a need for

further research linking monitoring of these medications to ADE outcomes using observed data from actual practice in order to inform the implementation of future guidelines concerning regular monitoring of these drugs

4.4 Comparators

The analysis aimed to estimate the effect of different frequencies of monitoring on the risk of occurrence of specific ADE outcomes linked to each study medication. In order to achieve this, the analysis compared three alternative strategies: monitoring in line with the frequency recommended in guidelines, compared against strategies involving conducting monitoring below or above this recommended level.

A major methodological issue in this analysis was identification of ADEs in patients who have not had a record of laboratory tests during the follow-up period. Certain ADE outcomes used in the analysis are typically identified through blood tests (such as hyperkalaemia using serum potassium, AKI using creatinine measurements). In patients who have had no laboratory tests during the time interval it was not possible to establish whether absence of recorded ADEs is attributed to actual absence of ADE, or lack of an opportunity to diagnose an abnormal test result. An additional analysis was conducted in this cohort in order to examine the risk of bias arising from its inclusion or exclusion in the final time-to-event analysis. The results of this investigation are described in Section 6.10.

The rationale for using three comparison groups to classify alternative ways of conducting regular monitoring was that this design offered the simplest method of comparison across strategies involving different frequency of monitoring. The inclusion of a larger number of comparison groups representing specific monitoring frequencies would complicate the analysis and increase the opportunity for misclassification errors due to the difficulty to pinpoint the exact frequency of routine monitoring within individual patients across a lengthy observation period based on retrospective analysis of EHRs. Monitoring at recommended frequency was chosen as the baseline comparator in order to test the hypothesis that the current acceptable practice in monitoring of ACEI, loop diuretics and amiodarone is superior to alternatives involving monitoring at a less or more frequent rate. In a case where sufficient evidence suggests that an alternative may yield a better outcome in terms of ADE prevention, a more in-

depth analysis may be conducted in order to identify what frequency may be considered optimal in order to inform clinical guidelines in this specific case.

4.5 Background of large databases of EHRs

The source of primary care data used to populate this analysis was the Clinical Practice Research Datalink (CPRD). Separate cohorts of patients prescribed each of the study medications were created. Separate analyses were carried out to estimate the effect of monitoring within each medication cohort due to differences in monitoring guidelines and the outcomes across study medications.

4.5.1 Medication safety research using large databases of electronic patient records

The use of routinely collected patient data for observational research on the effect of medications has substantially increased in recent years due to advances in technology and policy in the UK and abroad. A major factor which has increased the capabilities of secondary healthcare data is the widespread adoption of EHRs as a primary method of recording and storing routine patient data across healthcare. As well as benefiting practice through improvements in efficiency, safety of care and transparency, electronic records can be anonymised and used in research applications. The increased use of routinely collected EHRs to conduct research in the use of medications has led to the development of large databases which are designed to systematically collect, collate and clean electronic records from healthcare providers in order to facilitate their use by researchers. Examples of such databases in operation in England are the Clinical Practice Research Datalink (formerly the General Practice Research Database (GPRD/CPRD) (Jick, Kaye et al. 2003), QResearch (Hippisley-Cox, Stables et al. 2004), the Health Improvement Network (THIN) (Lewis, Schinnar et al. 2007), and Hospital Episode Statistics (HES) (Health & Social Care Information Centre 2015). The detailed information on episodes of care and diagnoses can complement existing sources of systematically collected data, such as the British Household Panel Survey (BHPS) (Taylor 2010), Office of National Statistics (ONS) mortality register and others in order to broaden the options for retrospective observational research in UK healthcare.

4.5.2 Advantages and limitations of the use of large databases of EHRs in the context of this study.

One of the main advantages of the use of large databases of EHRs in studies of monitoring is the use of systematic methods of recording and maintaining routine clinical data. Patient records are collected systematically and the completeness and reliability of data is checked using quality assurance mechanisms. This was important for this study as exposure and outcomes needed to be accurately ascertained from patient records to minimise differential misclassification. Data quality assurance mechanisms used in CPRD and HES are described in Section 4.6.

The continual improvements in data quality and quality assurance may act as a double edged sword. Increases in observed event rates, such as laboratory tests or diagnoses, may be attributed to improved recording rather than improvements in care quality. It is often difficult to separate this effect from the data. Trends in monitoring over calendar years were assessed in order to examine the potential effect of improved recording on the study results in Section 6.8. The completeness of EHRs could not be guaranteed prior to the introduction of Payment by Results in 2004 (Audit Commission 2010). For this reason, a decision was made to restrict the data used for this study to the period after 01.01.2003 in order to provide a balance between completeness of EHRs and length of follow-up in the study.

Large databases of longitudinal electronic patient data are useful in studies of monitoring due to the large, generally representative samples and a typically long period of follow-up. This improves internal validity and gives the investigator the opportunity to observe outcomes over a period of several years, which is seldom possible with physical collection of data from patient charts. In the UK, 98% of the population is registered with a GP and the ubiquitous use of computerised recording of patient data in general practices means that the patient populations in databases of GP EHRs have a high likelihood of being nationally representative.

The data collected within large databases reflects day-to-day practice in a real-life environment. This is a useful feature which allows this study to observe GP decision-making in follow-up monitoring and its impact on medication safety. The downside of this design is the potential for changes in behaviour by both patient and provider in response to events occurring during study follow-up, which may produce biased estimates of treatment effect. Such retrospective changes are outside the control of the investigator as they are observed retrospectively. In the

context of this study, certain design decisions were made in order to limit the effect of these changes, such as only considering the first case of ADE to avoid confounding the analysis with reactive monitoring in response to the ADE. The issue of changing exposure is discussed in more detail in Section 6.12.

Electronically recorded patient records tend to employ standard coding to denote clinically important events, such as prescriptions, diagnoses, procedures, consultations or episodes of care. Commonly used coding systems include the World Health Organisation (WHO) International Classification of Diseases (ICD) and Read codes to label diagnoses, Office of Population Censuses and Surveys (OPCS) classification for procedures, Healthcare Resource Group (HRG) for case mix analysis, BNF codes for medications and others. This aspect was useful for this study as it allowed for data mining of a large number of records of prescriptions, episodes and diagnoses in order to systematically determine exposure status, presence of ADEs and covariates. Search algorithms developed for this study may be re-used in subsequent studies using databases which employ the same coding systems (or other systems which have previously been cross-mapped, such as Read and ICD-10 codes), allowing the investigator to build upon previously developed methods and continually improve search queries. Consistent use of coding allows for observation of an outcome over multiple years or differences across geographical locations or facilities, reducing the risk of observation bias.

The format of data in most large databases of EHRs is suitable for a retrospective cohort study design. Non-random assignment of patients into comparison treatment groups can be adjusted using regression or propensity score methods, allowing for a quasi-experimental design to measure the causal effect of monitoring. Large databases have been previously used as a source of data for observational studies of medication safety. Applications have included the use of QResearch to analyse the safety profile of statins in over 2 million patients (Hippisley-Cox and Coupland 2010) and drug-related acute kidney injury in a cohort of over 2 million patients using CPRD (Dormuth, Hemmelgarn et al. 2013).

Large databases of EHRs are often linked via a common anonymised patient ID number, which is the case for CPRD and HES. This was one of the practical reasons for choosing these two databases to conduct the analysis in this study, as it allowed the direct linkage of GP data on monitoring to diagnoses and episodes of care occurring in hospital or other types of data, such as Office of National Statistics (ONS) mortality data. This allowed the current study to observe

treatment of ADEs both in general practice and hospitals, albeit only for a proportion of the cohort registered in GP practices which have given consent for their patient IDs to be linked to HES (currently stands at 59%). The extent of linked data to this day is limited, and considerable effort is often required to obtain records from other services. This study examined primary and secondary care data, but did not have the resources necessary to access other services that the study patients may have used, such as social or intermediate care. In addition, the CPRD database is limited to practices using the Vision electronic system. Healthcare in the UK is known for relatively low inter-operability of electronic systems across providers, creating problems for studies looking at outcomes across multiple services (Franklin, Berdunov et al. 2014). Until inter-operability improves across services, the potential of observational research using routinely recorded electronic records in the UK will be heavily constrained by access issues.

The biggest limitation of using retrospectively collected observational data to study monitoring is confounding arising from non-random assignment into different treatments. Decisions relating to the frequency of monitoring are likely to be influenced by patient factors, such as the presence of pre-existing conditions, treatments or patient characteristics which may increase the perceived risk of ADE. Therefore, a retrospective cohort study comparing groups of patients assigned to different monitoring regimens needs to handle systematic differences in characteristics across comparison groups. The issue of confounding arising from non-random treatment assignment, as well as methods used to limit the potential effect of this type of confounding on the results of this study are described in more detail in Section 5.5.

4.6 Description of data sources used in this thesis

4.6.1 General information

The Clinical Practice Research Datalink (formerly known as the General Practice Research Database) is a database of routinely collected primary healthcare records from 643 participating general practices in England. It contains anonymised records of over 3 million patients with a length of follow-up of up to 14 years. It is managed by the Medicines & Healthcare products Regulatory Agency (MHRA) in collaboration with the National Institute for Health Research (NIHR).

General practices in the UK act as the "gatekeeper" of the healthcare system and are generally the first port of call for individuals who seek medical treatment on a non-emergency basis. As such, recording data at the general practice level allows to form a comprehensive picture of population health. Each contact with primary care is recorded electronically, with the resulting EHRs managed by one of a number of electronic clinical systems. These include systems such as Vision, EMIS, TPP and SystmOne. CPRD employs EHRs managed by Vision, a system provided by In Practice Systems Ltd (INPS) to around 2300 general practices in the UK (Phoenix 2014). Participating practices input electronic records using a standardised protocol and the database is updated continuously on a monthly basis, with the application of systematic methods to ensure data quality. The methods of data input and consistency of coding of prescriptions, diagnoses and medical encounters is assessed on an ongoing basis, with practices which fail to comply with standards of data quality subject to removal from the database over time (Jick, Kaye et al. 2003). Consistency of data input ensures reliability of the information recorded on the database and high level of internal validity through minimisation of biased coding and measurement.

The Hospital Episode Statistics (HES) database is a collection of electronic records of admitted care at NHS hospitals in England. It was developed following the publication of a report by a steering group chaired by Dame Edith Korner in 1987 which highlighted the fact that only 10% of admitted care records were collected by the DoH (Health & Social Care Information Centre 2015). The HES database is split into separate sections covering episodes of care, diagnoses and procedures.

4.6.2 Representativeness and validation

Two quality aspects of an automated healthcare database which affect the external validity of studies which use it as a source of data are representativeness and validation. The degree to which a database is representative of the population from which practices and patients were recruited may be assessed using certain characteristics of patients, such as age and female/male ratio. The patient population on CPRD was shown to be representative of the UK population using census data (Gelfand 2005). Validity refers to the extent to which the population in the database is consistent in terms of clinical characteristics and outcomes to the patient population which it represents. This aspect is clearly important as validity of the database will impact directly on the validity of the results and conclusions drawn from studies employing the database as a source of patient records.

The practices included on the CPRD database are assessed periodically for consistency of recording of prescriptions and diagnostic data. Those practices which are deemed to have inconsistencies of records, such as significant gaps in recorded data which was unlikely to occur naturally, are removed from the database. Additional measures to ensure the high standard of recorded data include continuous comparison of the incidence of certain conditions found in observational studies, including breast cancer, influenza and colorectal cancer, which have been shown to have a similar incidence rates in CPRD records (Jick, Kaye et al. 2003).

The accuracy of diagnostic records on the GPRD/CPRD database has been extensively assessed in the literature. A systematic review of studies validating diagnoses in individual disease areas using alternative sources of data such as GP questionnaires or chart reviews has reported a positive predictive value (PPV) of over 90% for a number of common chronic conditions such as inflammatory bowel disease, diabetes and psychosis. A number of mental health conditions, such as dementia had PPVs of between 80% and 90%. Amongst the conditions which were relatively poorly recorded (PPV<80%) were cardiovascular events, atrial fibrillation and cardiac arrhythmia (Khan, Harrison et al. 2010).

There are certain limitations in quality of recorded data on CPRD. The first refers to uncertainty regarding the completeness of hospital referrals data. For this reason hospital referral records were not used to identify hospitalisations. These records were derived straight from the linked Hospital Episode Statistics database. The second limitation refers to missing data within lifestyle information category, which has been noted in a previous investigation of GPRD data (Gelfand 2005). For these reasons, smoking status and body mass index were not used to populate covariates in the analysis. This does not constitute a significant limitation of the study as lifestyle factors of patients have not been found to influence the conduct of monitoring.

A study evaluating the validity of HES in 2002 reported that data completeness was poor for episodes, diagnoses and procedures based on comparisons with alternative data sources for records between 1990 and 2000 (Williams and Mann 2002). This is due to low quality of recording of hospital data during this time period and has been observed by a number of authors (Dixon, Sanderson et al. 1998). The quality of recording in the years after 2000 has been on a positive trend, in particular following the introduction of Payment by Results in 2004 (Audit Commission 2010). A systematic review of studies comparing routinely

recorded EHRs in hospitals in the UK to case note review demonstrated a statistically significant improvement in completeness of diagnosis records (96% vs. 74%) and an improvement (albeit not sufficiently large to reach statistical significance) in the accuracy of recorded episodes (77% vs. 86%) (Burns, Bottle et al. 2011). The quality of recorded data on HES has been investigated using records from clinical registers for a number of outcomes, including all-cause mortality (Aylin, Bottle et al. 2007), cancer (Garout, Tilney et al. 2008) and outcomes of vascular surgery (Aylin, Lees et al. 2007). The HES database has previously been used as a source of data for analyses of the incidence of ADE-related hospitalisations (Waller, Shaw et al. 2005, Patel, Bell et al. 2007).

4.6.3 Included data

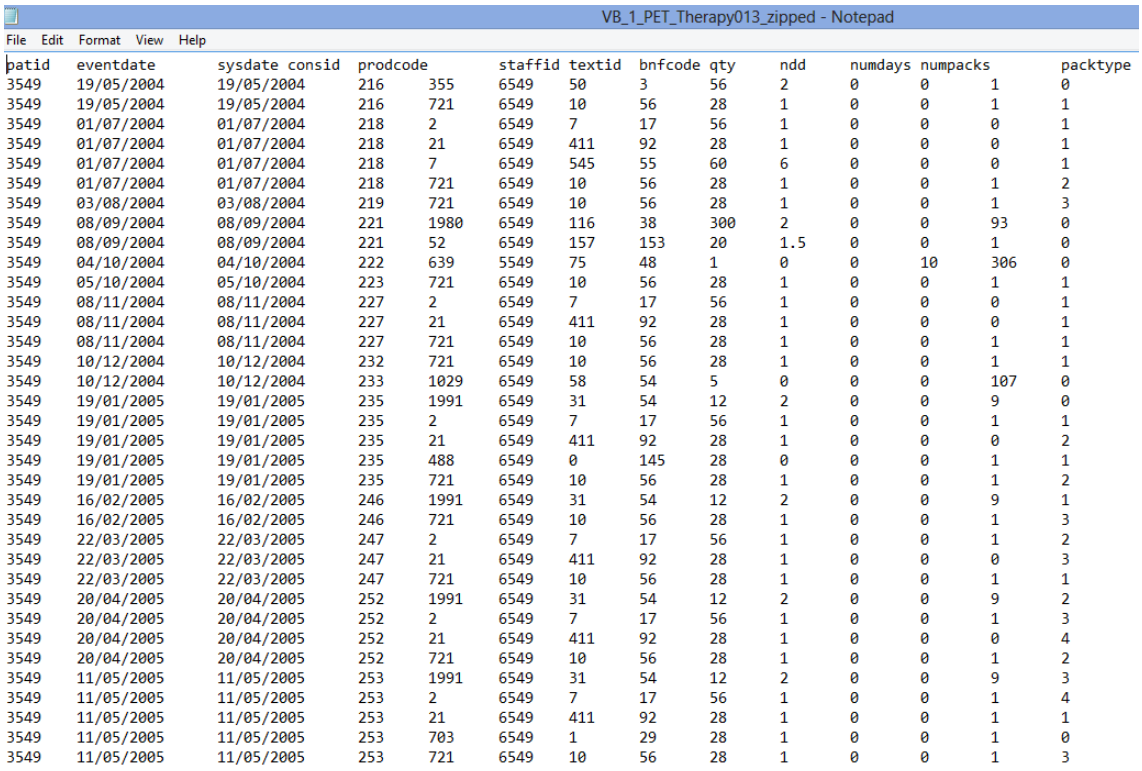
The type of data included on CPRD include patient demographics, details of practice, prescriptions, consultations, clinical events (diagnoses, clinical characteristics, lifestyle information) and laboratory test orders and results. Data is extracted in a series of files in plain text format with delimited values. Files contain one record per line and can be linked together using the unique patient identifier.

Prescription records include a product code linked to a CPRD dictionary which allows the user to identify the drug name, formulation and dosage. Information on the BNF chapter is included, which allows additional options for the searching and categorisation of prescription data.

The clinical sheet includes information on all clinical events, which include diagnoses, immunisation, referrals and lifestyle information (smoking, alcohol consumption and weight). All events are assigned a Read code to aid search and categorisation of diagnoses into groups. Read codes have been partially mapped to the International Classification of Disease, 10th revision, which is the standard coding system for diagnoses in secondary care in the UK. This allows the researcher to identify hospital admissions linked to conditions which were initially identified in primary care, and vice versa, with a high level of accuracy.

Screen shots of data extracted from CPRD are included in Figs 4.1 and 4.2.

Fig. 4.1. Raw plain text format of prescription data from CPRD

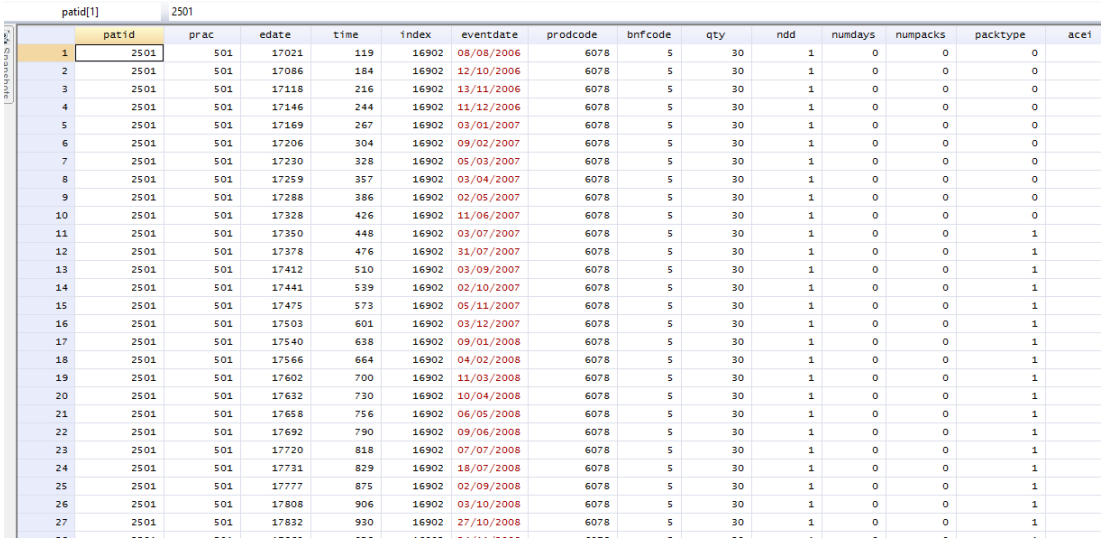


VB_1_PET_Therapy013_zipped - Notepad

File Edit Format View Help

patid	eventdate	sysdate	consid	prodcode	staffid	textid	bnfcode	qty	ndd	numdays	numpacks	packtype
3549	19/05/2004	19/05/2004	216	355	6549	50	3	56	2	0	0	1	0
3549	19/05/2004	19/05/2004	216	721	6549	10	56	28	1	0	0	1	1
3549	01/07/2004	01/07/2004	218	2	6549	7	17	56	1	0	0	0	1
3549	01/07/2004	01/07/2004	218	21	6549	411	92	28	1	0	0	0	1
3549	01/07/2004	01/07/2004	218	7	6549	545	55	60	6	0	0	0	1
3549	01/07/2004	01/07/2004	218	721	6549	10	56	28	1	0	0	1	2
3549	03/08/2004	03/08/2004	219	721	6549	10	56	28	1	0	0	1	3
3549	08/09/2004	08/09/2004	221	1980	6549	116	38	300	2	0	0	93	0
3549	08/09/2004	08/09/2004	221	52	6549	157	153	20	1.5	0	0	1	0
3549	04/10/2004	04/10/2004	222	639	5549	75	48	1	0	0	10	306	0
3549	05/10/2004	05/10/2004	223	721	6549	10	56	28	1	0	0	1	1
3549	08/11/2004	08/11/2004	227	2	6549	7	17	56	1	0	0	0	1
3549	08/11/2004	08/11/2004	227	21	6549	411	92	28	1	0	0	0	1
3549	08/11/2004	08/11/2004	227	721	6549	10	56	28	1	0	0	1	1
3549	10/12/2004	10/12/2004	232	721	6549	10	56	28	1	0	0	1	1
3549	10/12/2004	10/12/2004	233	1029	6549	58	54	5	0	0	0	107	0
3549	19/01/2005	19/01/2005	235	1991	6549	31	54	12	2	0	0	9	0
3549	19/01/2005	19/01/2005	235	2	6549	7	17	56	1	0	0	1	1
3549	19/01/2005	19/01/2005	235	21	6549	411	92	28	1	0	0	0	2
3549	19/01/2005	19/01/2005	235	488	6549	0	145	28	0	0	0	1	1
3549	19/01/2005	19/01/2005	235	721	6549	10	56	28	1	0	0	1	2
3549	16/02/2005	16/02/2005	246	1991	6549	31	54	12	2	0	0	9	1
3549	16/02/2005	16/02/2005	246	721	6549	10	56	28	1	0	0	1	3
3549	22/03/2005	22/03/2005	247	2	6549	7	17	56	1	0	0	1	2
3549	22/03/2005	22/03/2005	247	21	6549	411	92	28	1	0	0	0	3
3549	22/03/2005	22/03/2005	247	721	6549	10	56	28	1	0	0	1	1
3549	20/04/2005	20/04/2005	252	1991	6549	31	54	12	2	0	0	9	2
3549	20/04/2005	20/04/2005	252	2	6549	7	17	56	1	0	0	1	3
3549	20/04/2005	20/04/2005	252	21	6549	411	92	28	1	0	0	0	4
3549	20/04/2005	20/04/2005	252	721	6549	10	56	28	1	0	0	1	2
3549	11/05/2005	11/05/2005	253	1991	6549	31	54	12	2	0	0	9	3
3549	11/05/2005	11/05/2005	253	2	6549	7	17	56	1	0	0	1	4
3549	11/05/2005	11/05/2005	253	21	6549	411	92	28	1	0	0	1	1
3549	11/05/2005	11/05/2005	253	703	6549	1	29	28	1	0	0	1	0
3549	11/05/2005	11/05/2005	253	721	6549	10	56	28	1	0	0	1	3

Fig. 4.2: Cleaned and formatted prescription data following input into Stata 11



patid[1] 2501

	patid	prac	edate	time	index	eventdate	prodcode	bnfcode	qty	ndd	numdays	numpacks	packtype	acei
1	2501	501	17021	119	16902	08/08/2006	6078	5	30	1	0	0	0	
2	2501	501	17086	184	16902	12/10/2006	6078	5	30	1	0	0	0	
3	2501	501	17118	216	16902	13/11/2006	6078	5	30	1	0	0	0	
4	2501	501	17146	244	16902	11/12/2006	6078	5	30	1	0	0	0	
5	2501	501	17169	267	16902	03/01/2007	6078	5	30	1	0	0	0	
6	2501	501	17206	304	16902	09/02/2007	6078	5	30	1	0	0	0	
7	2501	501	17230	328	16902	05/03/2007	6078	5	30	1	0	0	0	
8	2501	501	17259	357	16902	03/04/2007	6078	5	30	1	0	0	0	
9	2501	501	17288	386	16902	02/05/2007	6078	5	30	1	0	0	0	
10	2501	501	17328	426	16902	11/06/2007	6078	5	30	1	0	0	0	
11	2501	501	17350	448	16902	03/07/2007	6078	5	30	1	0	0	1	
12	2501	501	17378	476	16902	31/07/2007	6078	5	30	1	0	0	1	
13	2501	501	17412	510	16902	03/09/2007	6078	5	30	1	0	0	1	
14	2501	501	17441	539	16902	02/10/2007	6078	5	30	1	0	0	1	
15	2501	501	17475	573	16902	05/11/2007	6078	5	30	1	0	0	1	
16	2501	501	17503	601	16902	03/12/2007	6078	5	30	1	0	0	1	
17	2501	501	17540	638	16902	09/01/2008	6078	5	30	1	0	0	1	
18	2501	501	17566	664	16902	04/02/2008	6078	5	30	1	0	0	1	
19	2501	501	17602	700	16902	11/03/2008	6078	5	30	1	0	0	1	
20	2501	501	17632	730	16902	10/04/2008	6078	5	30	1	0	0	1	
21	2501	501	17658	756	16902	06/05/2008	6078	5	30	1	0	0	1	
22	2501	501	17692	790	16902	09/06/2008	6078	5	30	1	0	0	1	
23	2501	501	17720	818	16902	07/07/2008	6078	5	30	1	0	0	1	
24	2501	501	17731	829	16902	18/07/2008	6078	5	30	1	0	0	1	
25	2501	501	17777	875	16902	02/09/2008	6078	5	30	1	0	0	1	
26	2501	501	17808	906	16902	03/10/2008	6078	5	30	1	0	0	1	
27	2501	501	17832	930	16902	27/10/2008	6078	5	30	1	0	0	1	

4.6.4 Description of data contained in HES

The HES database contains episode-level records of hospital admissions, diagnoses and procedures carried out during the hospital stay. Similarly to CPRD, records are split into a number of plain text files with delimited values for each column of data. Data sheets include patient demographics, admissions, hospital episodes, diagnoses, procedures and additional hospital care including critical and maternity care.

Summary of variables included in the raw data extracted from HES is included in Table A2 in the Appendix.

4.6.5 Coding systems

Much of the clinical information on CPRD is coded in order to optimise data volume and facilitate the search of specific items of interest for researchers. All clinical information, including diagnoses, disease history and tests are coded using the Oxford Medical Information System (OXMIS) and Read dictionaries.

The Read classification system is a collection of over 100,000 alphanumeric codes used to denote diseases, diagnoses, symptoms, clinical events, as well as personal and family history of disease in patients. It was developed by Dr James Read during the 1980s and has been used by the NHS as the primary classification system in primary care since 1990 (Dave and Petersen 2009). It facilitates systematic input of clinical information by GPs based on key terms which are cross-referenced to codes in the database. The codes are formed of five alphanumeric characters, with each character representing a level of detail of a particular condition. For example, the root G includes all circulatory system diseases, code G2 represents hypertensive diseases, G21 is hypertensive heart disease, within this category can be found G210z which stands for malignant hypertensive heart disease. This facilitates the application of search algorithms, which may be designed to capture specific diagnoses or entire categories of disease.

Prescriptions are classified according to a code based on the Prescription Pricing Authority dictionary. Additionally, prescriptions on CPRD care classified according to BNF code, which allows the development of search queries based on either the generic name or BNF chapter.

The HES database uses a different set of coding systems to denote diagnoses and procedures, which reflects the differences in the recording of clinical data between primary and secondary care providers in England. Diagnoses are recorded using the 10th revision of the International Classification of Diseases (ICD-10), which is the international standard diagnostic classification system created and maintained by the World Health Organisation (World Health Organisation 1992). Procedures are recorded using the 4th revision of the Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures (OPCS-4) system (Health & Social Care Information Centre), which was developed by the Clinical Classifications Service.

4.7 Study design, inclusion criteria and extraction methods

4.7.1 Study approval

The study protocol was reviewed by the Independent Scientific Advisory Committee (ISAC) for Medicines and Healthcare products Regulatory Agency in line with requirements for the extraction and dissemination of CPRD data. Linked data from the HES database was additionally requested. The protocol was approved in May 2014 (ISAC protocol no. 14_067). A copy of the approved protocol is available upon request from the author.

4.7.2 Study design

This study employed a retrospective cohort design using CPRD primary care records with linkage to secondary care records from HES spanning from 2003 to 2012. In the cohort study, monitoring within CPRD-participating practices was examined using descriptive statistics. The next stage of analysis involved an investigation of the effect of frequency of monitoring on the risk of ADE separately for each medication cohort using time-to-event analysis. This involved identifying monitoring frequency from CPRD test records and diagnoses of ADEs in primary care using CPRD clinical records and in secondary care through a search of HES episode and diagnosis data. In the amiodarone cohort, the results of the time-to-event analysis were used to estimate transition probabilities for a decision-analytic model simulating the outcomes of alternative strategies involving different frequency of regular monitoring. This was combined with unit costs and utilities sourced from the literature to estimate the cost-effectiveness of alternative frequencies in amiodarone monitoring. This chapter focuses specifically on the methods of cohort selection, data extraction, cleaning and preparation for analysis. Statistical methods specific to each analysis in the study are described in detail in Chapter 5. The methods of the economic analysis are described in Chapter 7 of this thesis.

4.7.3 Study medications

The final list of medications and corresponding CPRD product codes are included in Table A3 in the Appendix.

ACE inhibitors

The initial list of medications used to compile the inclusion criteria for the ACE inhibitor data extraction was derived from Section 2.5.5.1 of the BNF. It included captopril, cilazapril, enalapril maleate, fosinopril sodium, imidapril, lisinopril, moexipril, perindopril, quinapril, ramipril and trandolapril. Combined formulations, such as ACEI-diuretic combinations, were included in the list.

Loop diuretics

The initial list of medications used to compile the inclusion criteria for the loop diuretic data extraction were derived from Section 2.2.2 of the BNF (Joint Formulary Committee 2014). It included furosemide, bumetanide and torasemide.

Amiodarone

Amiodarone hydrochloride is found in Section 2.3.2 of the BNF. Product codes corresponding to all available formulations of amiodarone were searched in the CPRD product dictionary, which included both tablet and injection varieties.

4.7.4 Data sources

Individual-level records of patient demographics, prescriptions, referrals, laboratory tests, clinical event and consultation events were extracted from CPRD. Data extracted from HES included episodes, diagnoses and procedures.

Exposure and outcomes were observed from the start of the observation period (90 days following initiation) to the end of observed follow up (which is listed in Section 4.8 and differed according to the medication studied) or the final recorded data (death or transfer out date). Baseline characteristics of patients were collected by observing CPRD diagnosis and prescription records over a 12 month period prior to initiation date.

4.7.5 Selection of study cohort

The front-end online CPRD system was queried using the following inclusion criteria:

i. Patients ≥18 years old
ii. Registered at an up-to-standard practice

iii. First prescription of one of the study medications between 01.01.2003 and 31.12.2012, with at least 12 months of registration at the practice without a record of prescription of the included medication.

The study population was comprised of adults (aged ≥18 old) who were newly initiated with a course of therapy involving one of the study medications. Exclusion of patients registered in practices which were not up-to-date according to CPRD was done in order to ensure maximum data quality and minimise the effect of missing data. Only patients with newly prescribed medications were included in order to separate initiation monitoring from follow-up monitoring during ongoing therapy, as initiation therapy has a different purpose and was outside the scope of this study. The requirement of a minimum 12 month prescription-free period prior to initiation ensured that patients were on a new course of therapy.

The starting point for data management was the identification of the index date, which corresponds to the initiation of therapy with the drug of interest. The "therapy" file was used to identify the first recorded prescription which signals the beginning of therapy, as the inclusion criteria specified that a minimum of 12 prescription-free months prior to the first prescription. The "patient" data sheet was used to extract demographic information for each included patient.

4.7.6 Initial data cleaning

Certain variables did not add useful information and were removed in order to optimise data processing. A data check was performed to ensure that all included events fell within the study dates (01.01.2003-31.12.2012).

Laboratory test data from CPRD was used for two purposes: identifying evidence of monitoring during therapy and abnormal test results which indicated the presence of ADE. All relevant test records were included to identify exposure as they indicated that monitoring was conducted. Specific test records were selected in order to identify values which were above a certain range, suggesting the presence of an ADE. For example, in order to identify the presence of thyroid function monitoring, all records of TSH tests were identified, with dates extracted in order to identify whether monitoring was conducted during drug exposure. TSH test records with values below 0.1 μmol/L or above 10 μmol/L were identified in order to identify cases of thyrotoxicity and hypothyroidism, respectively. Test results which were not informative or unlikely values of test results were removed

following consultation with a GP with extensive experience in prescribing and drug monitoring in primary care (Prof. Anthony Avery). Specifically, serum potassium levels of over 20 mmol/L, serum creatinine values of over 1000 µmol/L were flagged up as potentially spurious based on the fact that such a high value was unlikely to be observed in practice. A manual search of these records was conducted to check the coherence of information included in all fields. Most tests with missing or implausible values in the test result field also had missing information for type of test, qualifier, lower and upper thresholds, in which case a decision was made to remove them from the dataset.

Dates are recorded on CPRD as strings in DD/MM/YYYY format. In order to facilitate data cleaning and identification of clinical events, all dates were converted using Stata code into a numeric format consisting of the number of days elapsed since 1960.

An algorithm was developed and applied in Stata to flag up duplicate values. A manual check of flagged values was performed to remove those values which were confirmed as duplicates

4.8 Identifying exposure

Two types of exposure were considered in this study: exposure to the study medication and exposure to the monitoring regimen assigned by the GP during the course of therapy. Drug therapy rarely remains unchanged over the course of observation, with changes in dosage, frequency of administration or discontinuation of therapy routinely occurring in response to the progression of disease, non-response or adverse reactions to the drug. In order to determine whether a patient was on active therapy during a particular time interval, the dates of prescription and the type of medication prescribed were extracted from the "therapy" file. Exposure to monitoring was identified by searching the CPRD laboratory test file for required monitoring tests. The search algorithm incorporated a Read code list developed using information from a previous investigation of monitoring in UK general practice and advice from a GP with extensive experience in prescribing and monitoring of medications in UK primary care (Prof. Anthony Avery). The list of Read codes used to extract laboratory tests is found in Table A4 in the Appendix.

4.8.1 Identifying exposure to the study medication

Exposure to drug therapy was established by identifying BNF codes corresponding to the study medications in the CPRD therapy file. CPRD product codes were matched to BNF codes for ACE inhibitors found in BNF Section 2.5.5.1, loop diuretics in Section 2.2.2 and amiodarone in Section 2.3.2. The date of the first prescription of a study medication was identified as the index date, which was the starting point for the analysis period for each patient. This value was applied to each patient record in order to standardise all dates to identify the number of days elapsed since initiation. This simplified coding as each patient had a unique initiation date, whilst length of observation and the cut-off points for observed exposure were common across patients in the same medication group.

Each medication cohort was analysed individually due to the differences in monitoring requirements and the ADE outcomes assessed in each study medication. Once a patient begins a new course of therapy with one of the study medications, they were placed in the appropriate cohort, with all exposures and outcomes assessed irrespective of other medication therapies which may have occurred in the same time period. This means that the same patient may have been placed in more than one medication-specific cohort for analysis. As each of the study medications presented unique risks of ADE and thus required a separate monitoring regimen, there were no major methodological limitations to this approach. The co-prescription of ACE inhibitors and loop diuretics may increase the risk of electrolyte ADEs and thus influence frequency of monitoring, which was controlled by adding a dummy variable for co-prescription of another antihypertensive medication (ACE inhibitor, diuretic or ARB) to control for the potential confounding effect of combined therapy in the analysis.

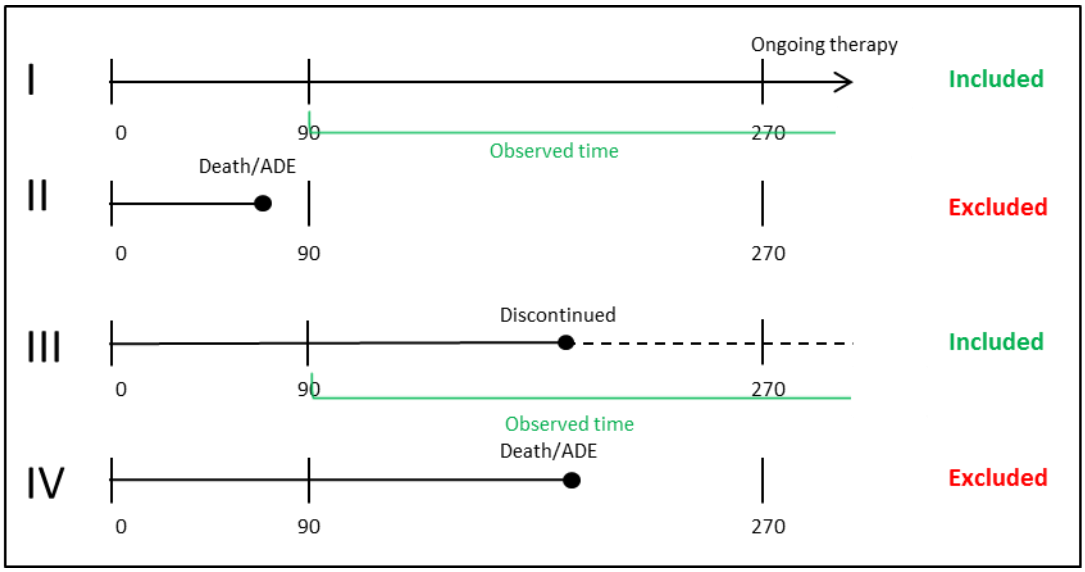
The length of observation of monitoring exposure and ADE outcomes was 3.75 years for ACE inhibitors and loop diuretics and 2.5 years for amiodarone. The choice of interval length was based on the recommended interval for monitoring of these medications listed in guidelines, which was 12 months for ACE inhibitors and loop diuretics and 6 months for amiodarone. The analysis involved assessment of frequency of monitoring by observing recorded tests across multiple monitoring "cycles" which corresponded to the recommended interval + 25% leeway for missed tests (guidelines for monitoring of study medications were described in Chapter 3 of this thesis). This method had a dual purpose: to allow the observation of the frequency of monitoring tests conducted in each monitoring interval and to identify the continuity of exposure to the drug by

identifying records of prescription of the study medication within each monitoring interval.

During a lengthy observation period (3.75 years for ACEi, diuretics; 2.5 years for amiodarone), an assumption of continuous therapy during the entire period of follow-up could not be made. For a number of patients in the cohort drug therapy may have been discontinued, either permanently or for a short period before resuming. A measure of continuity of therapy beyond the period of initial monitoring had to be included in order to ensure that all patients in the final study cohort were required to undergo follow-up monitoring.

Continuity of therapy in the initial stages of the observation period was ensured by requiring each patient to have a computerised prescription history of the target drug between days 90 and 270 from the date of the initial prescription. In other words, patients who were discontinued within the first 3 months of therapy and had no repeat prescription of a drug within the same class within the subsequent 6 months were assumed to have had their therapy withdrawn. These patients were removed from the dataset as the aim of the study was to assess the effect of follow-up monitoring during ongoing therapy. Figure 4.3 illustrates this approach in the context of three hypothetical scenarios.

Fig. 4.3. Description of cohort selection based on continuation of therapy in the first 9 months for all medication cohorts



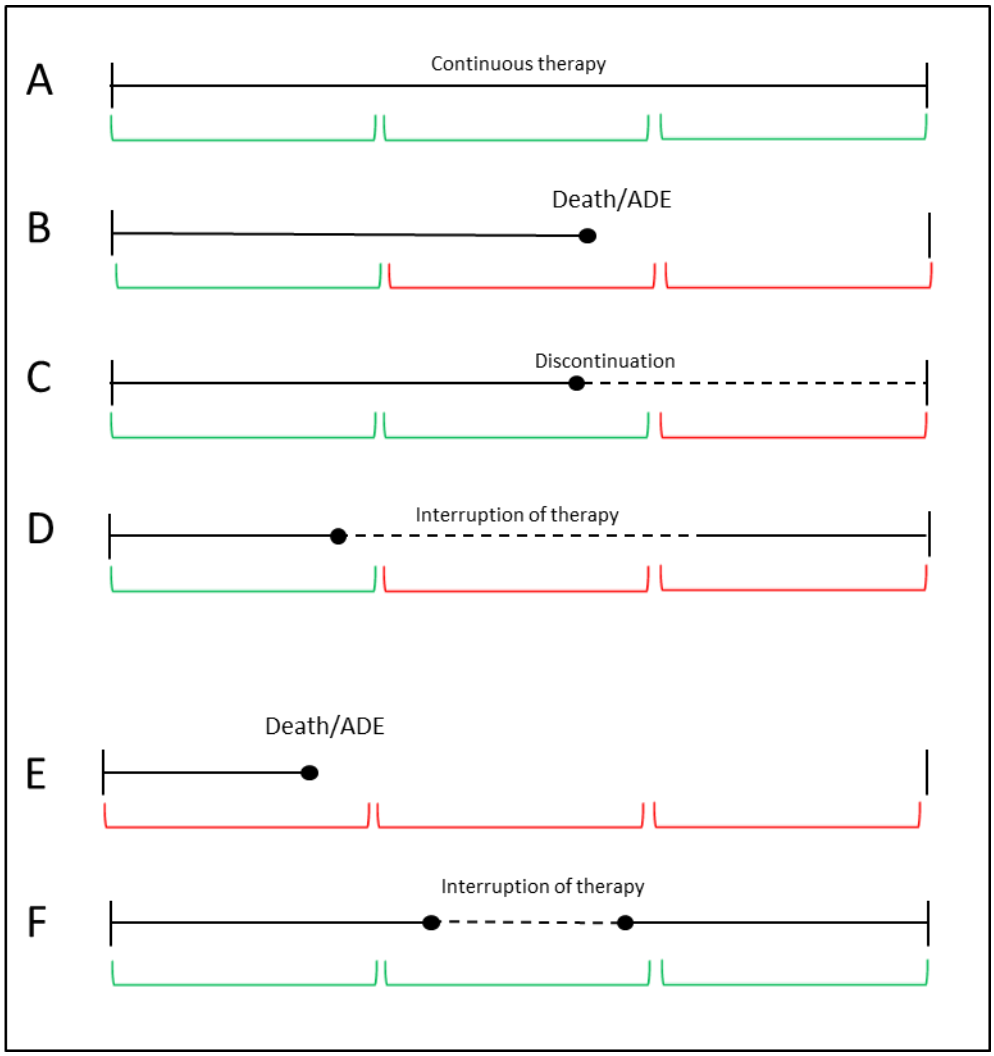
Scenario I represents a patient who is established on continuing therapy, with observed time starting at index+90. Scenario II is a patient who died or had a record of ADE before day 90 of therapy and was excluded as they have not had

long-term therapy. Scenario III is a case where therapy of the study medication carries on past day 90, but is discontinued before day 270. This patient was included in the study and observed for the first monitoring interval as they were still alive and eligible for monitoring in the first interval. Scenario IV is a patient who had a censoring event (death or ADE) between days 90 and 270. They were excluded on the basis of insufficiently long follow-up period to adequately observe monitoring exposure.

The aim of this study was to estimate the effect of frequency of monitoring on the risk of ADEs. To calculate the frequency, the number of tests and length of therapy were required. However, the length of therapy may vary across patients as some patients were discontinued or exited the dataset before the end of the observation period. In order to incorporate variable length of observation, the follow-up period was split into a several intervals, with the length of each interval determined by the recommended frequency of testing for each medication in the study. Chapter 3 reported that guidelines recommend U&E monitoring, including serum potassium, creatinine and/or sodium, once every 12 months during ongoing ACEI/loop diuretic therapy and thyroid function tests once every 6 months during ongoing amiodarone therapy. Each monitoring interval was intentionally made longer than the interval listed in guidelines by 25% in order to allow for missed appointments. Observation periods were thus set as 15 months for ACEI and loop diuretics and 7.5 months for amiodarone. Thus, the periods were defined as follows (all values expressed as number of days elapsed since the index date): $90 \leq p_1 < 540$, $540 \leq p_2 < 990$, $990 \leq p_3 < 1440$ for ACEi and diuretics; $90 \leq p_1 < 315$, $315 \leq p_2 < 540$, $540 \leq p_3 < 765$, $765 \leq p_4 < 990$ for amiodarone, where p_1, p_2, p_3, p_4 are study dates in periods 1-4.

For the purposes of analysis in this study, a patient was assumed to be on active therapy during a monitoring cycle in the presence of at least one recorded prescription of the drug during the period. For instance, in the absence of a record of prescription of an ACE inhibitor between days 540 and 990 (second monitoring cycle), therapy was assumed to have been discontinued for that particular patient. If a patient incurred a censoring event (death or ADE), monitoring was not observed during the period when the event occurred, as it was unknown whether monitoring was precluded by the event. Figure 4.4 illustrates the complexity of observing persistence of therapy and exposure to monitoring in the cohort.

Fig. 4.4. Description of the approach to medication persistence



The above figure demonstrates several hypothetical cases of patients in the ACEI and loop diuretic cohorts. Each line represents the observed time, which begins at index+90 and ends at the end of the analysis period (index+1440, or 3.75 years of observed time). Uninterrupted sections of the line represent active drug therapy with evidence of repeat prescriptions, dotted line represents the time when a patient is in contact with primary care (prescription or consultations) but is not on active therapy with the study medication. The brackets beneath each line show the monitoring intervals, each equalling 15 months, which are denoted by colour (green=on active therapy; red=discontinued). Exposure was not observed in an interval which contains censoring event (B), but was observed if it contains the last observed prescription (discontinuation or interruption of therapy) (C and D). If no prescription of the study medication occurs in one interval, therapy was assumed to be discontinued and exposure was not be observed in subsequent periods, as any resumption of therapy was assumed to be a new course of therapy (D).

Two further possible scenarios warrant interpretation. If a patient was censored out within the first period, they were excluded from the study as the event precluded monitoring; it was thus not possible to observe which monitoring regimen was initially assigned to the patient (E). If therapy has been interrupted and resumed within the same interval, continuity of therapy was assumed (F).

4.8.2 Identifying monitoring regimen

In each medication cohort, patients were categorised into four comparison groups according to the monitoring regimen assigned to them. Frequency of monitoring was categorised as follows: $n=0$ for absence of monitoring, $0<n<1$ for monitoring conducted less frequently than recommended in guidelines, $1\leq n\leq 1.5$ for monitoring conducted at recommended frequency, $n>1.5$ for monitoring conducted more frequently than recommended, where n= number of "test days" per period of continuing therapy. This design allowed for a comparative analysis of alternative monitoring regimens against a baseline group assigned monitoring in accordance with guidelines. The reason that recommended frequency group was defined in such a way, rather than selecting patients with $n=1$ was due to the added 25% leeway which may have resulted in two monitoring tests in different years falling within the same interval. Having $n=1$ as the requirement would have therefore classified such cases as "more frequent", when in actual fact they formed part of a monitoring regimen which strictly adhered to guidelines.

ACEI and loop diuretics

Chapter 3 specified that guidelines for follow-up monitoring of ACE inhibitors and loop diuretics recommended urea & electrolyte (U&E) monitoring once every 12 months. U&E tests are composed of serum sodium, potassium, urea and creatinine tests, which are recorded individually in CPRD laboratory test records. The presence of monitoring was defined as a record of at least two of sodium, potassium or creatinine tests on any given day. The follow-up period was split into three monitoring intervals in order to observe the frequency of monitoring assigned to each patient. The total number of "test days" recorded within the observed period of continuing therapy was divided by the number of periods in which evidence of continuing therapy was found (presence of at least one prescription of the target medication).

Amiodarone

Chapter 3 specified that guidelines for follow-up monitoring of amiodarone recommended thyroid function tests (TFT) to be conducted once every 6 months. TFT tests are split into TSH, free T3 and T4 tests. A record of one or more of the above during a given day of the observed period counted as one "test day". The total follow-up period was split into 4 monitoring intervals. The total number of "test days" recorded within the observed period of ongoing therapy was divided by the number of periods in which evidence of continuing therapy was found (presence of at least one prescription of the target medication).

4.9 Identifying outcomes

4.9.1 General information

The outcomes in this study included specific ADEs linked to the study medications observed during the study period. The list of outcomes in each medication cohort and criteria used to identify these outcomes from CPRD and HES records is included in Table A6 in the Appendix. ADEs were identified using a search of Read codes for a diagnosis of the condition or an abnormal laboratory test result. The first day of diagnosis or recorded abnormal test result since the index date was considered the date of onset of the ADE.

Several methods were used to identify ADEs in the cohort study. CPRD Read-coded diagnoses and laboratory tests were used to identify ADEs treated in primary care. Primary ICD-10 codes were used to identify ADE-related hospital admissions.

During the process of designing the study there were several alternative methods of identifying ADEs through test results: use of reference ranges used by the GP which are recorded in the laboratory test sheet of CPRD or applying fixed thresholds across the whole dataset, for each outcome in the study.

4.9.2 ACE inhibitors

The primary outcomes assessed in the study of ACE inhibitor monitoring were hyperkalaemia and acute kidney injury (AKI). ADE cases treated both in primary care and the hospital setting were included in the analysis.

Diagnoses

CPRD clinical file was queried for Read codes corresponding to hyperkalaemia and acute kidney injury (AKI).

Laboratory tests

CPRD test file was queried for Read codes corresponding to serum potassium and creatinine level tests. Following the removal of missing and extreme values as described in Section 4.7.6, hyperkalaemia was identified as serum potassium >6mmol/L. A patient was deemed to have AKI in case of a rise of serum creatinine of ≥26 μmol/L or ≥1.5 fold within a 7 day period, in accordance with the guidelines for the diagnosis of AKI (Ftouh, Thomas et al. 2013). The dates of abnormal results were recorded and reconciled with dates found in the clinical file to identify the first occurrence of the ADE, which was subsequently used as the censoring event in the time-to-event model.

Hospital admissions

Day cases and hospital admissions with a primary ICD-10 code corresponding to hyperkalaemia and AKI were identified. These were reconciled with ADE diagnoses and abnormal test results identified from CPRD records to identify the censoring ADE event in the time-to-event analyses of hospital outcomes using data from GP practices with a link to HES.

4.9.3 Loop diuretics:

The primary outcomes assessed in the study of loop diuretic monitoring were hypokalaemia and hyponatraemia. Additional outcomes assessed included falls as a consequence of underlying drug-induced hypokalaemia and hyponatraemia. ADE cases treated both in primary care and hospital settings were included in the analysis.

Diagnoses

CPRD and HES diagnostic records were queried for Read codes and ICD-10 codes corresponding to hypokalaemia and hyponatraemia. Fall-related injuries were identified from CPRD diagnosis records using a list of Read codes corresponding to an accidental fall and an injury/fracture. HES records were searched for hospitalisation with a primary diagnosis of injury or fracture. In order to establish that an injury was caused by an accidental fall, the presence of a Read code for

an accidental fall and Read/ICD-10 code for an injury or fracture within 30 days of each other were required to positively identify the ADE.

Laboratory tests

The CPRD laboratory test file was searched for valid test results. Mild hypokalaemia was identified as serum K<3.5mmol/L; hypokalaemia was identified as <2.5mmol/L, according to standard thresholds reported in studies of hypokalaemia (Oram, McDonald et al. 2013). Mild hyponatraemia was identified as serum Na<135mmol/L; severe hyponatraemia was <120mmol/L (Craig 2010). Cases of mild hypokalaemia and hyponatraemia were identified in order to correct the length of observed monitoring in the cohort. A finding of mild electrolyte imbalance is likely to induce more intensive monitoring, which may be erroneously classified as regular monitoring if it is included. Severe hyponatraemia and hypokalaemia identified through laboratory test results were classified as ADE in the analysis of primary care outcomes.

4.9.4 Amiodarone

Diagnoses

CPRD and HES diagnosis records were queried for Read codes corresponding to hypothyroidism (AIH) and thyrotoxicity (AIT). Major adverse cardiovascular events (MACE) were identified from HES data using ICD-10 codes corresponding to the conditions which are considered a MACE: heart failure, myocardial infarction, stroke, transient ischaemic attack, ventricular arrhythmia (Yiu, Jim et al. 2009).

Laboratory tests

The CPRD laboratory test file was searched for valid test results. Clinical AIH was identified as thyroid stimulating hormone (TSH) >10 mU/L, clinical AIT was identified as TSH < 0.2 mU/L, according to the thresholds for diagnosis listed in British Thyroid Association guidelines (British Thyroid Association 2006).

4.10 Comment on methods for identification of exposure and outcomes from retrospective data

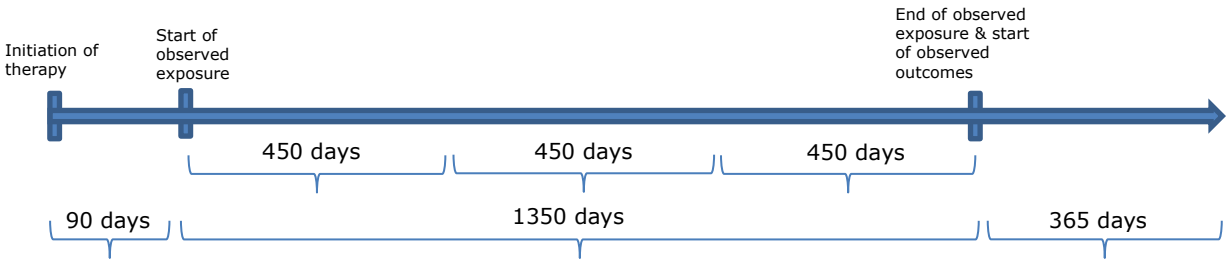
The time-to-event analysis in this study identified both the exposure in terms of the assigned monitoring regimen and outcomes in terms of diagnoses of ADE and abnormal test results by observing CPRD and HES records over the same time period (≥90 and <1440 days post-initiation in ACEI and loop diuretics, ≥90 and

<990 days post initiation in amiodarone). This approach is the most straightforward way of observing different strategies of monitoring and their effect in real practice, and has been the method of choice in previous observational studies which linked medication monitoring with data on ADEs (Raebel et al. 2010, McDowell et al. 2012). However, simultaneously observing both exposure and monitoring poses several methodological issues:

1. More frequent testing in practice may increase the likelihood of an ADE diagnosis, particularly in the context of ADEs identified using abnormal laboratory results, such as hyperkalaemia, hyponatraemia or thyrotoxicity. Patients may be diagnosed with these conditions simply by virtue of being in close contact with primary care.
2. Patients who have not experienced an ADE in the first few months of therapy may be moved to a less frequent regimen as the perceived risk of ADE is lower.

The above issues may introduce bias into the analysis and underestimate the protective effect of monitoring. The observed test and ADE records in the ACEI records were examined in Sections 6.11, 6.12 and 6.13 in order to assess the potential effect of these issues on the analysis results. An alternative method which separates the period of observed exposure and outcomes was investigated to establish whether it is possible to resolve these issues with the available data extracted from CPRD. An additional 12 month time interval following the end of the exposure period (1440 days post-initiation) was explored in order to identify ADEs. The original follow-up period was used to determine monitoring group assignment. The alternative design is described in Fig. 4.5.

Fig 4.5: Alternative method of identifying monitoring exposure and ADE outcomes from CPRD



The observed test and ADE records in the ACEI cohort will be examined in Section 6.12 in order to assess the potential effect of these issues on the analysis results. The results of the application of the alternative method in the ACEI cohort are outlined in Section 6.13.

4.11 Identifying covariates

4.11.1 General information

Each medication cohort included a number of covariates in order to control for potential confounding factors which affect both the exposure (monitoring regimen assigned) and the outcome (risk of medication-specific ADEs). In the ACE inhibitor and loop diuretics cohort, the covariates included: age, sex, comorbidity, polypharmacy, co-prescription of other cardiovascular medications which require routine monitoring of serum potassium sodium and creatinine, including ACE inhibitors, ARBs, loop and thiazide diuretics, presence of diabetes or chronic kidney disease (CKD) and year of initiation. The amiodarone cohort included the following covariates: age, sex, comorbidity, polypharmacy, year of initiation. Additional covariates were used to control for confounding when specifying analyses of amiodarone-induced MACE: prior diagnoses of hypertension, diabetes, hyperlipidaemia, ischaemic heart disease and left ventricular dysfunction at baseline. Please refer to Chapters 2 and 3 for a summary of the evidence on patient factors associated with ADEs and monitoring.

4.11.2 Demographic factors

Age was calculated at initiation of therapy using the date of birth and the index date. Additional variables extracted from the "patient" sheet were sex and practice number, which corresponds to the last 3 digits of the CPRD patient ID. Information on the date of death was obtained from the "patient" sheet and linked ONS mortality data.

4.11.3 Multimorbidity

A previously validated co-morbidity score which consists of 40 chronic disease categories developed by Barnett et al was used as a measure of multimorbidity (Barnett, Mercer et al. 2012). The rationale for using this measure to control for multimorbidity in the cohort study was the inclusion of a large number of chronic disease categories which may influence the decision to assign a patient to more intensive monitoring regimen. In addition, the authors of the index published detailed methods of calculating the measure from routinely collected diagnosis and prescription records. An algorithm was created in Stata 11 in order to adapt the comorbidity categories to Read-coded diagnosis and prescription codes used in CPRD. An additional motivation for using a more inclusive measure of morbidity was the fact that the Charlson index, as the most commonly used alternative

measure of comorbidity, was primarily designed to predict mortality (Charlson, Pompei et al. 1987). A multimorbidity measure which is associated with other outcomes, such as conduct of monitoring and the risk of ADE, was considered more appropriate.

The multimorbidity score was calculated at baseline and was constructed from information on the diagnoses and prescribed medications in the 12 month period preceding the index date. The criteria used to identify comorbidities are listed in Table A7 in the Appendix.

4.11.4 Polypharmacy and co-prescribed cardiovascular medications

A polypharmacy measure was designed by compiling a list of CPRD product codes based on the chapters of the BNF. The extracted prescription records were assessed to calculate the number of medications prescribed from different BNF chapters for each patient. Concomitant prescriptions with another cardiovascular medication (such as diuretics and/or ARB during ACEI therapy) which could potentially confound monitoring assignment were identified from the CPRD prescription records.

4.11.5 Specific conditions

Chapter 2 identified a number of conditions which have been previously found to be associated with the risk of ADEs linked to the study medications. These included diabetes mellitus and chronic kidney disease (CKD) associated with ACEI or loop diuretic-induced electrolyte disorders. Diagnoses of diabetes mellitus (both Type 1 and Type 2) and CKD were identified from CPRD diagnosis records during the 12 months preceding initiation to determine the presence of these conditions in patients. In the case of CKD, the stage of disease was identified, which ranged from 1 to 5 based on the Kidney Disease Outcomes Quality Initiative classification (The Renal Association). Due to possible progression or regression of kidney function over time, the highest CKD stage attained during the 12 month period was selected.

4.11.6 Other factors

The year of initiation was included as a confounder in the analysis in order to control for secular trends in the practice of monitoring during the calendar years in the study. The length of observed time (in terms of number of monitoring

intervals with continuing drug therapy) for each patient was also included as a covariate in order to control for unobserved factors associated with early discontinuation of therapy.

4.12 Discussion

The use of large databases of EHR for medication safety research offered several methodological advantages for this study. These include access to large, representative samples, lengthy follow-up and systematically coded diagnoses, prescriptions and clinical events. Using EHRs facilitates observation of both monitoring practice and its effects in terms of incidence of ADEs in real-life settings. However, the use of large databases comes with certain limitations: it is often difficult to link electronic data across different services, and as in the case of primary care in England, across competing electronic record systems within the same sector. The use of this source of data for observational research carries with it the issue of non-random treatment assignment, a design limitation shared with other types of observational data which needs to be handled independently using statistical tools designed to control confounding in observational research.

This study employed CPRD as its main source of EHRs, which is a representative collection of primary care records from over 600 practices in England. Patient records contained in CPRD have previously been validated against alternative sources of data and undergo periodic data checks to ensure the completeness and high quality of information of recorded data. It is partially linked to the HES database, which allowed the current study to examine the effect of monitoring on ADEs treated both in primary and secondary care. This study identified key parameters from routinely recorded data recorded on CPRD, including exposure to drug therapy from prescription files, monitoring regimen from laboratory test files and ADE outcomes using Read-coded diagnoses in general practices and ICD-10 coded hospital admissions in hospitals.

A cohort study design was chosen in order to answer the research questions set out in this programme of research. In order to achieve this, each medication cohort was divided into three comparison groups according to their assigned monitoring frequency, which was observed during a follow-up period of 1350 days in ACEI and loop diuretics and 900 in amiodarone. The length of follow-up reflected the recommended interval of monitoring in each medication cohort, and comprised three intervals in ACEI and four in amiodarone in order to observe regimens with different frequency over a sufficiently long period, whilst

minimising attrition arising from discontinuation of therapy which inevitably occurs over time. Appropriate methods were applied in order to identify patients who are on long-term therapy and thus eligible for follow-up monitoring and distinguish them from those discontinued or censored at an early stage. The number of tests observed during the follow-up period was used to identify the assigned regimen. The outcomes assessed in the study were cases of ADE specific to each medication cohort, identified from diagnoses in GP and hospital records, as well as abnormal laboratory tests.

Chapter 3 identified several design features which are important in a study of monitoring, including methods to correctly identify patients who are exposed to drug therapy, measuring frequency of monitoring and identifying ADE outcomes. The CPRD and HES databases provide a rich source of patient-level data which allowed the investigator to address each of these issues. Continuity of drug therapy was ensured by identifying periods of discontinuation, monitoring regimens were identified by observing the timing of laboratory tests and cases of ADE were correctly captured using a systematic search of diagnoses and abnormal laboratory test results. Detailed records of diagnoses and prescriptions on CPRD were used to construct multimorbidity and polypharmacy measures, as well as covariates representing the presence of specific conditions and co-prescriptions which may influence the decision to conduct monitoring. These covariates were used to construct propensity score weights to control for selection bias, which is outlined in Section 5.5.2 in the next chapter.

The use of retrospectively collected data from EHRs presented certain methodological issues relevant to the analysis of monitoring outlined in this chapter. The first issue stems from the nature of some of the ADE outcomes which were identified from abnormal test results and required laboratory tests in order to be diagnosed. The inclusion of patients with no recorded monitoring may have resulted in an underestimate of the rate of ADEs. The second issue was caused by identifying monitoring exposure and ADE outcomes from the same time interval in the data, which could potentially underestimate the protective effect of monitoring in the time-to-event analysis. The implications of these issues in the context of this study are examined in Section 6.12 and an alternative analysis was conducted, the results of which are outlined in Section 6.13.

The finalised datasets containing the exposures, outcomes and covariates for each included patient in the study were used in statistical analyses to investigate the effectiveness and cost-effectiveness of monitoring. A detailed description of

how this patient-level data was applied in the time-to-event analysis is described in Chapter 5. Chapter 6 summarises the results of the time-to-event analysis, whilst Chapter 7 describes the methods and results of the economic analysis.

Chapter 5

Statistical methods

5.1 Introduction to the chapter

This chapter describes the methods of an analysis of the effect of different frequencies of follow-up laboratory monitoring on the risk of specific ADEs linked to ACE inhibitors, loop diuretics and amiodarone. To begin, the aims and objectives of the chapter are listed, which is followed by a description of the methods of controlling for non-random treatment assignment in observational research and a description of methods used to estimate propensity scores using diagnosis and prescription records from CPRD. This is followed by a discussion of the key analysis design decisions and how these methods were used to estimate the effect of regular monitoring of the study medications on the risk of ADE using retrospectively collected data from CPRD and HES. A comment is made on how the results of this analysis were used to populate a model to estimate the cost-effectiveness of thyroid function monitoring in amiodarone therapy, which is the focus of Chapter 7 of this thesis.

5.2 Chapter objectives

i. Describe the rationale for the chosen analysis design.

ii. Describe the specific exposures, outcomes and covariates and how they were used in the time-to-event analysis.

iii. Outline the steps taken to control confounding arising from non-random assignment into comparison groups with alternative monitoring regimens.

iv. Describe methods of time-to-event analysis to estimate the effect of monitoring on the hazard rate of specific ADEs linked to study medications and provide justification for the chosen method.

5.3 Set up of the analysis

5.3.1 Aims and objectives of analysis

The aim of the analysis described in this chapter was to investigate the effect of conducting routine monitoring of medications during established therapy with ACE inhibitors, loop diuretics and amiodarone in terms of the risk of specific adverse drug events (ADEs) linked to these medications.

The following objectives were set for the analysis of monitoring and ADEs:

1. Apply statistical methods to control for non-random assignment into comparator groups using covariates derived from CPRD
2. Specify a time-to-event model in the form of Weibull regression and estimate the effect of monitoring frequency on the following outcomes:
    - Hyperkalaemia, AKI during ACEI therapy
    - Hypokalaemia, hyponatraemia and fall-related injuries during loop diuretic therapy
    - AIH, AIT and MACE during amiodarone therapy

5.3.2 Patient inclusion and cohort selection

The sources of data used to populate this analysis were the CPRD and HES databases. Separate cohorts of patients prescribed each of the study medications were created. Separate analyses were carried out to estimate the effect of monitoring within each medication cohort due to differences in monitoring guidelines and the outcomes across study medications. A full description of inclusion criteria, data cleaning and cohort selection steps are included in Chapter 4.

5.3.3 Exposure

A separate analysis was conducted for each medication cohort; therefore, separate methods to identify exposure to monitoring were developed in the case of each study medication. In the ACE inhibitor and loop diuretic datasets exposure to monitoring was determined by the presence of a urea & electrolyte (U&E) test during the analysis period. In the amiodarone dataset, presence of monitoring was defined as the presence of a thyroid function test (TFT). Three types of tests are carried out as part of thyroid function monitoring: thyroid-stimulating hormone (TSH), triiodothyronine (T3) and thyroxine (T4). The

methods of identifying exposure in the study for all three datasets are described in chapter 4.

5.3.4 Outcomes

The outcome of interest in this study was the risk of specific ADEs during the course of long-term therapy. ADE outcomes were tailored to the specific medicines and medicine groups under investigation in this study and were as follows:

i. Hyperkalaemia linked to ACEI treated in primary care, identified from CPRD diagnosis and laboratory test records; hyperkalaemia treated in secondary care, identified from HES primary diagnosis data.
ii. Acute kidney injury (AKI) linked to ACEI treated in primary care, identified from CPRD diagnosis and laboratory test records; AKI treated in secondary care, identified from HES primary diagnosis data.
iii. Severe hypokalaemia or hyponatraemia linked to loop diuretics treated in primary care, identified from CPRD laboratory test records; severe hypokalaemia or hyponatraemia treated in secondary care, identified from HES primary diagnosis data.
iv. Injuries incurred as a result of an accidental fall linked to a loop diuretic which were treated in secondary care, identified using Read codes for a fall from CPRD and HES diagnoses for admission or day case treatment of an injury.
v. Hypothyroidism (AIH) linked to amiodarone, treated in primary care, identified from CPRD diagnosis and laboratory test records.
vi. Thyrotoxicity (AIT) linked to amiodarone, treated in primary care, identified from CPRD diagnosis and laboratory test records.
vii. Emergency hospital admission for major adverse cardiovascular event (MACE) linked to amiodarone, identified using HES primary diagnoses.

The rationale for the inclusion of the above outcomes, specific ranges used to identify abnormal laboratory tests and other information can be found in Chapter 4. The full list of Read and ICD-10 codes used to identify relevant records from CPRD and HES, respectively, is included in Table A6 in the Appendix.

5.3.5 Patient covariates

The literature review in Chapter 3 identified a number of patient characteristics may be associated with both the conduct of monitoring and the risk of ADEs. GPs are more likely to carry out laboratory tests in patients with known risk of factors for specific ADEs, resulting in more frequent monitoring of certain patient groups. In particular, previous studies of monitoring have found that older patients are more likely to be monitored appropriately (Raebel, McClure et al. 2007, Akbarov, Kontopantelis et al. 2015), as are patients with multiple comorbid conditions (Simon, Andrade et al. 2005) and concomitant drug treatments (Akbarov, Kontopantelis et al. 2015). Other factors which may influence the decision to conduct regular monitoring are presence of specific conditions such as CKD, diabetes and heart failure and co-prescription of certain drugs, including diuretics (Raebel, McClure et al. 2007). In the context of this study, presence of specific conditions, such as type I and II diabetes and chronic kidney disease were risk factors for hyperkalaemia and hyponatraemia in ACE inhibitor or diuretic therapy, meaning that patients with these conditions are more likely to be carefully observed using laboratory tests during therapy. Therefore, a number of pre-specified patient and treatment-specific covariates were included in the analysis in order to control for factors which may confound the relationship between monitoring and risk of ADE in the study cohorts. All covariates were measured at baseline in order to estimate their potential effect on assignment to a particular monitoring regimen, as well as the risk of ADE, apart from co-prescription of other medications which are likely to prompt more frequent monitoring, which were assessed during the follow-up period.

Covariates extracted directly from CPRD included demographic factors (age at index date and sex) and the year of initiation of therapy. Additional patient characteristics were derived from CPRD prescription and clinical records. This included multimorbidity (measured as the number of chronic conditions diagnosed at baseline), polypharmacy (number of medications prescribed from different BNF chapters). In the ACE inhibitor (ACEi) and loop diuretics cohorts, the presence of specific conditions such as diabetes and CKD were assessed as well as concomitant prescription with other antihypertensive medications (such as loop diuretics and/or ARB during ACEi therapy). A detailed description of the methods for constructing covariates for analysis from CPRD records can be found in Section 4.11.

Lifestyle factors, such as smoking, BMI and alcohol consumption are known to have a significant impact on patient health and the propensity to develop chronic

and acute disease. The literature review in Chapter 3 examined patient factors which affect monitoring assignment in observational studies of monitoring. The review did not find evidence that patient lifestyle factors significantly contribute to decision-making in monitoring, based on several large-scale studies (Simon, Andrade et al. 2005) (Raebel, McClure et al. 2007) (Akbarov, Kontopantelis et al. 2015). Monitoring assignment is generally affected by current treatment (co-prescriptions) and prevalent conditions (e.g. diabetes, CKD), but not by factors which may determine health status in the future. The addition of lifestyle factors from CPRD requires identifying a search of clinical records using a validated code list to identify smoking, BMI and alcohol consumption recorded during GP consultations. Accurate identification of these covariates is difficult due to their dynamic nature (patients may stop and start smoking or drinking, lose or gain weight over time), and the quality of electronic recording of lifestyle information in general practices is uncertain. Lifestyle factors were not examined in this studied as their addition would not benefit the design of the time-to-event analysis.

5.4 Estimating treatment effect from observational data

5.4.1 Rationale for the choice of observational cohort design

In any investigation of cause and effect, both the exposure or intervention in question and the chosen outcome are likely to be affected by other variables. The effect of an intervention on the health of patients or the cost of treatment may be affected by setting-specific (e.g. clinic, GP practice, hospital), provider-specific (the healthcare professional delivering the intervention) or patient-specific (e.g. adherence, accessibility) factors which may affect the uptake of the intervention. Similarly, provider or patient factors may affect the outcome of the intervention: older patients with pre-existing chronic conditions, those in more deprived areas are more likely to have worse outcomes. An investigator needs to find a way of isolating the causal effect of the intervention from other variables which influence the exposure or outcome, or both, in order to find an unbiased measure of the average treatment effect.

This analytical problem can be framed in terms of the causal model formulated by Donald D. Rubin (Rubin 1974). The most accurate way to measure the causal effect of the intervention is to measure the outcomes of control and intervention in each participant simultaneously. In reality, it is not possible to measure both counterfactuals and the investigator is instead presented with an array of

alternative methods which approximate this perfect experimental design to estimate an unbiased estimate of causal effect.

The simplest extension of the hypothetical counterfactual case described above is a repeated trial. This involves applying each treatment in turn to the same participant and measuring the difference in outcome. This design may produce a biased measure of causal effect as the characteristics of the trial or its participants may change in subsequent observations. For example, a repeated trial to estimate the effect of monitoring may be compromised by attrition due to ADEs or discontinuation of therapy, changing conditions as patients age or develop new medical conditions.

Another option is to measure the average treatment effect by comparing two or more comparison groups, rather than by observing the treatment effect within each individual. Assignment of participants into comparison groups using a random mechanism may balance the mean values of characteristics of the participants across studies. Random assignment, in theory, removes systematic differences in patient characteristics across groups and any differences in baseline covariates are assumed to be random. Once this has been achieved, the true causal effect of the intervention can be estimated using statistical tools. For these reasons the randomised controlled trial (RCT) is often considered superior to other methods of causal inference (D'Agostino and Kwan 1995) and ranks at the top of the hierarchy of published evidence in terms of minimisation of bias alongside meta-analyses and systematic review of RCTs, according to NICE (NICE 2006).

Despite the clear methodological advantage of random treatment assignment, an RCT does not guarantee against systematic bias across comparators. Internal validity may be compromised if the behaviour of study participants is affected by knowledge of treatment assignment (e.g. decreased adherence in control group by patients, differences in care towards patients with treatment and control by a GP), which may produce a biased estimate of the causal effect. In such cases the typical solution is blinding, which involves making either the investigator or participant (or both) unaware of which treatment is delivered to individual cases, which should preclude systematic behavioural differences.

A randomised study design may additionally compromise generalizability. It has been observed that clinical researchers are different from clinicians who do not participate in research and may have a systematically different approaches to delivering an intervention compared to usual practice (Sarles, Camatte et al.

1977, Black 1996). Recruitment of participants into a trial is typically done on a voluntary basis, with patients who are older, more ill and poorer under-represented in RCTs, meaning that the recruited sample may not be representative of the population (McKee, Britton et al. 1999). Over-recruitment from these under-represented population subgroups may remedy this issue. Other potential factors which could reduce external validity in RCTs is the application of strict protocols for the delivery of intervention, exclusion of patients based on a high risk of ADE or certain characteristics such as old age, co-prescriptions, previous failed treatment, leading to the exclusion of patient groups which may be of interest to the study.

Other potential disadvantages of RCTs compared to an observational cohort design is the substantial cost of designing and implementing an RCT, which are often labour intensive in terms of design, data collection and ensuring compliance with study and ethical protocols. By contrast, an observational cohort analysis using either manual collection of data in a prospective format or retrospectively using routinely collected data and electronic health records may prove a more economical alternative. In addition, the cost and time constraints imposed on investigators result in relatively smaller sample sizes in RCTs, as compared to the potential size of datasets available from routinely collected data. This is particularly useful in an investigation of rare outcomes such as adverse drug events. In an observational study, a patient who drops out of treatment may still be observed and the reasons and consequences of treatment discontinuation may be better informed. Routinely collected data typically offers longer periods of observed data, which is useful in cases when the outcomes of an intervention may take a number of years to be fully realised.

5.5 Methods for controlling non-random assignment

Section 4.5.2 demonstrated a justification for the use of an observational study design in the analysis of the effect of monitoring for the prevention of ADEs in primary care. The main methodological limitation of using observational data is the potential bias arising from non-random assignment into treatment groups, known as selection bias. This type of bias arises from differences in the distribution of baseline covariates across comparison groups which may confound the estimation of the causal effect of the treatment. Non-random treatment assignment holds relevance in the context of this study, as the decision to conduct monitoring at a certain frequency during the course of therapy with ACEI,

loop diuretics or amiodarone is likely to be affected by the characteristics of the patient and the presence of known risk factors.

Two competing methods have typically been used to control for confounding in epidemiological research: regression and propensity scores. This section provides a description of both of these methods, their merits and limitations and describes the rationale for the use of propensity score weighting in this study.

5.5.1 Statistical regression analysis

The most common method of controlling for known confounding factors in observational studies is regression analysis. It involves specifying a regression equation which relates a response variable to some linear or non-linear combination of known covariates. Different types of regression analysis are applied depending on the type of response variable. The exposure of interest is included as a covariate in the model and treatment effect is estimated through interpretation of the coefficient corresponding to the exposure variable. This can be the marginal treatment effect (in case of linear regression) or can be interpreted as an odds ratio (logistic regression) or hazard ratio (time-to-event model). One of the advantages of this method include the ability to assess the fit of the model to the data using metrics such as the R squared, which estimates the proportion of the variation in the response variable which is explained by the model or likelihood-ratio test, which estimates the probability of achieving the result relative to a null model with no included covariates. These tools are used to assess how reliable the model is at drawing inferences from the observed data. Another advantage is the ease of interpretation of the results, as the estimated coefficients represent the marginal effect of an explanatory variable on the outcome (e.g. ordinary least squares) or change in terms of the risk or hazard ratio (logistic or time-to-event, respectively). A regression model may achieve a poor fit in case of incorrect selection of explanatory variables or if comparison groups don't overlap to a large degree in terms of baseline characteristics. Another potential disadvantage of regression analysis is the risk of collinearity between variables which may produce biased coefficient estimates or inflated standard errors.

5.5.2 Propensity scores

When are propensity scores used?

Propensity scores may be applied to deal with selection bias issues in observational research (Rosenbaum 1983). A propensity score is the probability of treatment assignment conditional on observed baseline covariates (Rosenbaum 1983). Balancing the group assignment in a cohort study based on this score draws the distribution of baseline characteristics of the treatment and control groups closer together. This reduces the risk of confounding from systematic differences in baseline characteristics which influence both treatment group membership and the outcome of interest. The propensity score is estimated by fitting a multivariate logit or probit model with treatment assignment as the response variable and a number of covariates which are believed to be linked to treatment allocation. There are several methods used for balancing treatment assignment based on the propensity score, including matching, stratification, inverse probability of treatment weighting (IPTW) and covariate adjustment (Austin 2011). The merits and limitations of each method are described below.

Estimating the propensity score

In the context of this analysis, propensity scores were estimated using an ordered probit model. This choice of method is justified by the need to calculate propensity scores for each of the three alternative monitoring regimens. The exposure used in this study was ordinal, in the sense that the categories were in increasing order to monitoring intensity, starting with infrequent monitoring, followed by recommended and frequent testing. In this case an ordered probit was the most appropriate type of regression.

Matching

Matching on the propensity score involves bringing together sets of treated and untreated patients with similar probability of treatment assignment. Once a matched cohort is established, estimates of treatment effect can be derived directly just as in any experimental format. Patients in the control and intervention groups may be matched on a one-to-one basis, with one control assigned to each treatment unit. In some cases matching multiple controls to each participant may be superior in order to avoid loss of information.

The performance of a propensity score matching method can be appraised by measuring two parameters: distance and balance. Distance is the difference in the propensity score between matched units; balance measures the difference in

distribution of baseline covariates between the resulting treatment and control groups (D'Agostino 1998).

A matching algorithm brings together subjects by minimising the distance between treated patients and members of the control pool. The decision regarding which algorithm to apply depends on several factors, such as the number of covariates available for analysis (Cochran 1968, ROSENBAUM and RUBIN 1983, D'Agostino 1998, Austin 2011). Two alternative algorithms are most widely used in the epidemiological literature: one which attaches each treated unit to one or more control units with the closest score (by minimising the "distance", which is the difference in propensity scores), known as the greedy algorithm, and another algorithm which minimises the total distance (known as optimal matching). Optimal matching works best at minimising distance in cases where the number of available units in the control pool (m) equals the number of treated units (n) (thus $r=m/n=1$) or in cases where r is small. Conversely, as $r \rightarrow \infty$ this difference tends to zero as further control units are available for matching, thus allowing the more straightforward "greedy" method to achieve a smaller total distance (Austin 2011).

Stratification

Stratification is a simple method used in comparisons across groups in the presence of systematic differences in baseline covariates. The cohort may be stratified according to one or more covariates and analyses subsequently conducted within each stratum in order to compare patients with similar characteristics. For instance, the cohort may be stratified according to age or the number comorbid conditions if these factors are expected to affect initial exposure and/or the outcome. Comparisons across patients of a similar age or health state assume that the treatment effect is less likely to be confounded by age and comorbidity. However, this method is not practical in the presence of multiple confounding factors as it requires a large number of strata, reducing the available sample size.

This problem may be addressed by stratifying the sample of subjects based on the estimated propensity score, which combines information on all confounding factors which affect exposure assignment. Direct comparisons across treated and control patients within each stratum may then be conducted. The appropriate number of strata will depend on the overall sample size and the effectiveness of the stratification process to remove bias. Generally, the percentage of bias removed will increase with the number of strata used (Cochran 1968). However,

as the number of subdivisions increases, the sample size of each stratum will decrease, affecting statistical power. The aggregate effect is estimated by calculating the average effect across all strata. This method has been shown to be effective at removing most of the selection bias (Cochran 1968, Rosenbaum and Rubin 1984). However, it has been noted that stratification on the propensity score may not perform well with a large number of covariates as this increases the number values across which matching is conducted (Cochran 1968). This issue may be circumvented by using applying categorical, rather than interval variables, to increase the chance that all patients in the intervention group are assigned a match from the control pool.

Another method which employs stratification using the propensity score is mean marginal weighting with stratification (MMWS), which is a variation of propensity score weighting. The merits and limitations of this method are discussed in the corresponding section below.

Using the propensity score as a covariate in a regression model

The estimated propensity score can also be used directly as a covariate in the regression model to estimate treatment effect. This method may offer certain advantages over matching and stratification as it does not involve loss of information from unmatched patients and is not hindered by issues arising from a large number of covariates. Adjustment can be applied directly to a regression model involving the whole cohort or in subgroup analysis by estimating the relationship between the propensity score and the outcome within separate exposure groups. A major limitation of this method is the possibility that the analysis may include patient subgroups which don't overlap in terms of their propensity scores, thus potentially drawing incorrect inferences from the data.

Inverse probability of treatment weighting (IPTW)

Another way of using the calculated propensity score is to compute weights corresponding to the inverse probability of a particular patient being assigned to a treatment group. These values may be used to re-weigh the sample in favour of patient groups which are under-represented in certain treatment groups. For example, patients who are monitored less frequently may be relatively younger and have fewer comorbid conditions compared with patients who are monitored more regularly. IPTW attaches a larger weight to patients with a higher age and comorbidity score in order to re-balance the distribution of baseline covariates, and a lower weight to those patients in the other groups, drawing the comparison

groups closer in terms of characteristics of included patients and rendering them more comparable.

The weights are calculated as follows (adapted from Robins et al) (Robins, Hernan et al. 2000):

$$\frac{Z_i}{e_i} + \frac{(1 - Z_i)}{1 - e_i}$$

Where Z_i represents a dummy for the assignment to the intervention of patient *i* (1=treated; 0=control) and is the propensity score of patient *i*. Essentially, the weight reduces down to $1/e_i$ for treated patients and $1(1\text{-}e_i)$ for non-treated patients, which re-weights the sample in favour of patients who are assigned the treatment but had a low propensity to be assigned (e.g. patients who are older and more ill but assigned to lower frequency monitoring as per previous example).

Marginal mean weighting through stratification

An alternative method of using propensity scores to control for non-treatment assignment was developed by Hong (Hong 2012). It involves combining IPTW with stratification on the propensity score to calculate weights based on probability of treatment assignment. It involves calculating propensity scores and stratifying the sample into quintiles based on each propensity score (applying multiple scores as the number of possible assignments>2). The calculated weight is conditional on the assigned treatment and the quintile in which the patient is placed. The weight is calculated using the following formula:

$$\frac{n_{s_\zeta} \cdot \Pr(Z = \zeta)}{n_{\zeta=\zeta, s_\zeta}}$$

Where n_{s_ζ} is the number of patients in quintile s_ζ , $\Pr(Z = \zeta)$ is the probability of being assigned to treatment ζ and $n_{\zeta=\zeta, s_\zeta}$ is the number of patients in quintile s_ζ actually assigned to treatment ζ .

Choice of method

There is a variety of ways in which statistical tools may be used to control for selection bias arising from non-random assignment. Both propensity score and standard regression techniques are valid methods used to adjust for confounding factors in observational research. Matching on the propensity score is applicable to binary exposure and was considered inappropriate in the context of this study due to the need to incorporate three levels monitoring frequency. MMWS and IPTW were considered as alternative methods of weighting on the propensity score with more than two exposures. A sensitivity analysis was conducted to test the performance of IPTW, MMWS and standard regression in order to determine the most appropriate method. The results of the time-to-event analysis of hyperkalaemia and AKI in the ACEI dataset were compared across the three methods and adjustment was tested by examining the covariate balance across comparison groups. MMWS was the method of choice of controlling confounding in the time-to-event analysis based on the results of the sensitivity analysis. The results of the sensitivity analysis are included in Section 6.14 and the potential implications for the study results are discussed in Section 6.15.2.

Whilst propensity score weighting may be used to handle confounding arising from systematic differences in known confounders, it is unable to control for unmeasured confounding. Appropriate methods were applied in order to derive values for relevant confounders which affect monitoring group assignment, such as a valid measure of multimorbidity and polypharmacy (as described in Chapter 4). However, it was not possible to guarantee that all confounding factors were present in the analysis. This remains a limitation of the data source used. In order to ensure comparability across the three monitoring frequency groups in the time-to-event analysis, patients had to have a similar probability of being assigned into each of the three monitoring regimens. This was done by estimating an area of common support, which was defined as the interval between the maximum of the lowest value and minimum of the highest value across the three propensity score distributions. This means that a patient with a propensity score in this range may be compared to a patient with a similar propensity score in either of the other two comparison groups. All patients with propensity scores which fall outside the interval of common support were assigned a propensity score weight of zero.

5.6 Methods for estimation of causal effect of monitoring

5.6.1 Survival analysis: rationale

Censoring in the context of time-series data analysis occurs when a patient exits the analysis before the planned follow-up date. In other words, study subjects may finish treatment, move to another catchment area or die before the end of the study, which upon retrospective review will appear as missing, or "censored" data. Censoring is a common feature of retrospective cohort studies, making it difficult to make comparisons across patients with different length of exposure. This poses a methodological issue in standard time-series analysis which assumes that all patients complete follow-up.

Standard regression analysis techniques estimate a relationship between an outcome variable and a number of covariates by assuming a constant probability of the event over time, otherwise known as the constant hazard condition. For example, the cumulative probability of cancer recurrence following chemotherapy over a 4 year period is assumed to be double the cumulative hazard over a 2 year period. This assumption may not always hold, i.e. the hazard of an event may be significantly higher soon after the event than after a certain amount of time has elapsed. Conversely, the hazard function, which represents the probability of the event as a function of time, may increase with time.

Time-to-event analysis, otherwise known as survival analysis, is a type of regression analysis which estimates the treatment effect in terms of the probability of occurrence at a given point in time (hazard rate), rather than an absolute value whose probability of occurrence is independent of the elapsed time interval over which outcomes were observed. It is methodologically flexible, in the sense that it can either explicitly assume a distribution for the hazard function (exponential, gamma or other) or it allows the option of not assuming an explicit distribution, so long as the hazard is proportional over time across comparison groups (Cox proportional hazard model). An additional advantage of time-to-event regression methods is the fact that it takes into account censored data by incorporating the length of exposure, rather than assuming equal exposure across all study subjects (Tabachnick 2007).

5.6.2 Preparing dataset for time-to-event analysis

The datasets for the time-to-event analyses were set up in such a way that all patient information was contained in one row. In other words, each row

corresponded to a patient, with individual columns representing exposure time, presence/absence of specific ADE outcome and the monitoring regimen assigned. Time-to-event analysis explicitly models time of exposure, requiring information on the length of time (in days) from start of analysis until exit. The length of the analysis period was 900 days in the amiodarone dataset and 1350 days in the ACE inhibitor dataset. Patients who were censored out (due to death, discontinuation or ADE) were identified, and the length of time from start of analysis until they exited the study was calculated from event dates in CPRD records.

The next step in data preparation is adjusting the resulting patient-time for immortal time bias. This type of bias is defined as the period of patient-time during which the event cannot occur – a patient is said to be in "immortal time" from the start of analysis until the start of the exposure period, as censoring out is precluded until the start date of exposure. Classifying this time period as exposed was likely to overestimate the effect of exposure on the risk of outcome and needed to be reclassified as unexposed to avoid bias.

In the case of this study, patients who have received monitoring were precluded from having an ADE in the time interval between study entry and the first monitoring test. If this time interval is included in the "monitored" patient-time in the survival analysis, this will invariably inflate the protective effect of the monitoring strategy. In order to avoid this error, the time interval prior to the first test in all monitored patients was subtracted from the total observed time prior to analysis.

The parameters of the Weibull model were estimated by fitting a time-to-event regression model using the `streg` command with option `dist` used to specify the type of distribution required using Stata 11.

5.6.3 Specifying the time-to-event regression model

In time-to-event analysis, several alternative methods are used to estimate the effect of treatment on the hazard rate (instantaneous probability) of a particular event. The most commonly used method is the semi-parametric Cox proportional hazards model, which fits a regression model of the log hazard rate of the response variable on the log baseline hazard and a set of covariates. The Cox regression model is a convenient choice as it does not require the investigator to make assumptions about the shape of the hazard function, only requiring that the hazard ratio between treated and control groups is constant with respect to time

(proportional hazards assumption). However, if prior knowledge of the process determining the rate of failure is sufficient to form an assumption of the shape of the hazard function, fitting a parametric model to the survival data may offer a better estimate of the treatment effect. Firstly, it allows better inferences to be made regarding survival times in treated and control groups; secondly, it allows extrapolation by predicting rates of failure beyond the length of the observed period, which is particularly useful when modelling risk of an event over extended periods, such as a patient's lifetime (Carroll 2003).

A variety of parametric time-to-event models have been developed to reflect the full range of assumed distributions of the failure process. For cases involving constant hazard, the exponential model may be used. However, this is a strong assumption, e.g. it is difficult to say with confidence that the risk of an ADE is the same 1 year into therapy as it is 3 years into therapy. Other models assume a variable hazard rate over time: these include the gamma distribution model, which assumes an asymptotically decreasing hazard rate over time, log-logistic and log-normal models, which assume that the hazard rate initially rises before falling asymptotically after hitting a peak. A specific distribution can be applied in a situation where an assumption can be made regarding the change of the hazard rate over time; i.e. whether it is likely to fall, rise or stay constant.

The alternative parametric approach involves assuming a Weibull distribution, which models the hazard function explicitly. It estimates two parameters: scale parameter λ and shape parameter γ. The advantage of the Weibull model is that it is applicable in cases where the hazard rate rises or falls over time, which is captured by γ. Where $\gamma=1$ the hazard rate is constant, and the hazard function is identical to the exponential model.

Using the probability density function of the Weibull distribution, we can define the parametric form of the Weibull model as follows:

$$h(t) = \gamma\lambda t^{\gamma-1}$$

$$H(t) = \lambda t^{\gamma}$$

$$S(t) = e^{-\lambda t^{\gamma}}$$

Where t is the time elapsed, $\lambda \geq 0$ and $\gamma > 0$.

Weibull regression estimates the shape and scale parameters of the hazard function by fitting a Weibull distribution to the time-to-event data. The Weibull parameters may be applied directly to a decision-analytic model in order to

estimate the probability of progression of ADE in a fixed time interval. These probabilities were used in a model to estimate the cost-effectiveness of thyroid function monitoring in amiodarone therapy, the methods and results of which are outlined in Chapter 7.

5.7 Discussion

This chapter laid out a framework for a time-to-event analysis of monitoring frequency of ACEI, loop diuretics and amiodarone using CPRD and HES data. In previous chapters of this thesis, the issue of non-random assignment into treatment groups in observational research was explored, highlighting the importance of dealing with confounding factors which may affect the decision to conduct regular monitoring in order to avoid selection bias. The merits of different methods of achieving this using retrospectively collected data were explored in this chapter. These included propensity score methods (matching, stratification, covariate adjustment, IPTW and MMWS) and standard regression methods. In the context of analysis in this study, MMWS, IPTW and standard regression were found to be potentially valid methods, particularly given the need to incorporate three levels of exposure. The choice of appropriate method was determined using a sensitivity analysis comparing all three methods, the results of which are discussed in Section 6.14.

A longitudinal analysis of monitoring cannot assume a constant length of follow-up across the entire cohort due to patients exiting the study dataset as a result of discontinuation of therapy, death or ADE. In order to incorporate censoring into the analysis, time-to-event regression was selected as the method of choice to estimate the effect of monitoring frequency on the risk of ADE in this study. Weibull regression was chosen as the functional form of time-to-event analysis. A Weibull model was suitable for this analysis due to its ability to estimate the parameters of the hazard function from observed data, which may be used to construct an economic model.

The methods outlined in this chapter were applied to the datasets derived from CPRD and HES databases described in Chapter 4. The next chapter outlines the descriptive analysis of monitoring, the process of estimating propensity scores and calibration of the design of time-to-event analyses in each medication-specific cohort. The results of the Weibull models are presented, followed by interpretation of the results and a comment on the way the results of this

analysis were applied to a model-based cost-effectiveness analysis of thyroid function monitoring of amiodarone therapy.

Chapter 6

Results of the time-to-event analysis of monitoring and ADEs linked to the study medications

6.1 Introduction to the chapter

This chapter summarises the results of the quantitative analysis of monitoring and ADEs linked to ACE inhibitors, loop diuretics and amiodarone, the methods for which have been described in Chapter 5 of this thesis. The details of patient numbers from the data extraction are summarised in Table A1 in the Appendix. The analyses were carried out using patient data from CPRD and HES, with the methods for cohort selection, identification of exposure, outcomes and covariates outlined in Chapter 4. Each section of this chapter concentrates on a separate medication cohort and is split as follows: descriptive statistics of cohort, results of probit model to estimate propensity scores used for adjustment of confounding, and the results of the time-to-event analyses which examine the relationship between the frequency of follow-up monitoring of the study medications and the rate of ADEs. At the end of the chapter, the numbers of observed ADEs, analysis of monitoring over calendar years in the study period and the common support procedure are summarised. Due to partial linkage between CPRD and HES data, separate cohorts were formed using records from HES-linked CPRD practices in order to carry out analyses of ADE which required hospital care. Sensitivity analyses testing the validity of study design decisions are outlined in Sections 6.10 to 6.14. At the end of the chapter, the results are given interpretation and in case of the amiodarone cohort, the implications of the results of the time-to-event analyses on the economic analysis.

6.2 ACE inhibitors: full-sample analysis

6.2.1 Descriptive analysis

The descriptive statistics of the cohort are summarised in Table 6.1. The final cohort of patients on established therapy with an ACE inhibitor included 306203 patients in 599 general practices. The mean number of patients per practice was 511 (SD 314). Mean age in the cohort was 63.2 (SD 13.4) and 46.2% were female. Multimorbidity was common, with 1 in 5 patients having 2 or more comorbid conditions at baseline. Polypharmacy was also common in the cohort,

with 36% of the cohort co-prescribed 5 or more medications from different BNF chapters during the exposure period, including the index prescription.

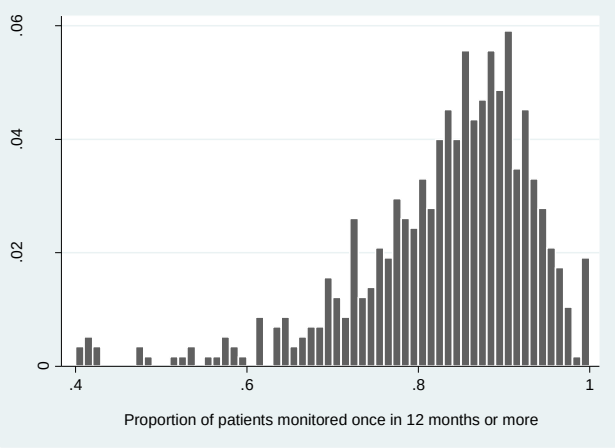
A high proportion of the cohort had records of U&E tests below the recommended frequency of once in 12 months during long-term therapy. Fifteen percent (45020/306203) did not have any record of monitoring during treatment and were thus excluded prior to the analysis, in accordance with the cohort selection methods outlined in Chapter 4. Over a third of patients were monitored in accordance with guidelines (35.4%, 108231/306203), whilst 36.2% (110705/306203) were monitored more frequently than once in 12 months. A subset of patients had a record of some monitoring, albeit below recommended frequency (13.8%, 42247/306203).

Table 6.1: Summary of patient characteristics, ACEI full cohort

Variable	All patients	No monitoring	Less frequent than recommended	Recommended frequency	More frequent than recommended
		$t^a=0$	0<t<1	1≤t≤1.5	t>1.5
Patient characteristics					
N (% of total)	306203	45020 (14.7)	42247 (13.8)	108231 (35.4)	110705 (36.2)
Mean age (SD)	63.2 (13.4)	61.1 (14.5)	59.1 (13.4)	62.5 (13.1)	66.3 (12.6)
Age category: n (%)					
<50	50770	10118 (19.9)	10479 (20.6)	18480 (36.4)	11693 (23.0)
50-59	70278	11007 (15.7)	11999 (17.1)	26604 (37.9)	20668 (29.4)
60-69	80409	10636 (13.2)	10059 (12.5)	29496 (36.7)	30218 (37.6)
70-79	67401	8022 (11.9)	6400 (9.5)	22151 (32.9)	30828 (45.7)
80-89	33402	4408 (13.2)	2932 (8.8)	10293 (30.8)	15769 (47.2)
≥90	3943	829 (21.0)	378 (9.6)	1207 (30.6)	1529 (38.8)
Female sex %	46.2	46.9	44.3	45.6	47.2
Mean number of co-morbid conditions (SD)	0.82 (1.01)	0.76 (0.98)	0.66 (0.90)	0.75 (0.95)	0.97 (1.09)
Number of comorbid conditions					
0 or 1	245398	36936 (15.1)	36009 (14.7)	89371 (36.4)	83082 (33.9)
2	40072	5361 (13.4)	4408 (11.0)	12889 (32.2)	17414 (43.5)
3	14012	1832 (13.1)	1290 (9.2)	4156 (29.7)	6734 (48.1)
4	4834	651 (13.5)	393 (8.1)	1346 (27.8)	2444 (50.6)
≥5	1887	240 (12.7)	147 (7.8)	469 (24.9)	1031 (54.6)
Mean no. of co-prescriptions (SD)	4.7 (2.5)	3.9 (2.4)	4.1 (2.3)	4.4 (2.4)	5.5 (2.6)
No. of co-prescriptions					
0	30160	7681 (25.5)	5425 (18.0)	11606 (38.5)	5448 (18.1)
1	38410	8044 (20.9)	6686 (17.4)	14790 (38.5)	8890 (23.2)
2	42677	7174 (16.8)	7055 (16.5)	16542 (38.8)	11906 (27.9)

3	43502	6141 (14.1)	6613 (15.2)	16491 (37.9)	14257 (32.8)
4	41258	5146 (12.5)	5508 (13.4)	14890 (36.1)	15714 (38.1)
5	35892	3966 (11.1)	4197 (11.7)	12454 (34.7)	15275 (42.6)
6	28691	2884 (10.1)	2952 (10.3)	9164 (31.9)	13691 (47.7)
7	20430	1946 (9.5)	1953 (9.6)	6063 (29.7)	10468 (51.2)
8	12879	1087 (8.4)	1038 (8.1)	3412 (26.5)	7342 (57.0)
≥9	12304	951 (7.7)	820 (6.7)	2819 (22.9)	7714 (62.7)
N (%) co-prescribed ARBs	52835 (17.3)	7049 (15.7)	3903 (9.2)	17605 (16.3)	24278 (21.9)
Co-prescribed loop or thiazide diuretics	128266 (41.9)	16255 (36.1)	15914 (37.7)	42728 (39.5)	53369 (48.2)
N (SD) with diagnosis of CKD at baseline	13790 (4.5)	1367 (3.0)	965 (2.3)	4072 (3.8)	7386 (6.7)
CKD stage (% of total in category)					
1	209	30 (14.4)	18 (8.6)	86 (41.2)	75 (35.9)
2	1351	120 (8.9)	142 (10.5)	482 (35.7)	607 (44.9)
3	11714	1141 (9.7)	776 (6.6)	3400 (29.0)	6397 (54.6)
4	461	56 (12.2)	25 (5.4)	91 (19.7)	289 (62.7)
5	55	20 (36.4)	4 (7.3)	13 (23.6)	18 (32.7)
N (%) with a diagnosis of diabetes at baseline	34430 (11.2)	2589 (5.8)	1755 (4.2)	9364 (8.7)	20722 (18.7)
Year of initiation					
2003	30902	4188 (13.6)	4358 (14.1)	10204 (33.0)	12152 (39.3)
2004	38042	4697 (12.4)	5250 (13.8)	12745 (33.5)	15350 (40.4)
2005	38805	4943 (12.7)	5325 (13.7)	13466 (34.7)	15071 (38.8)
2006	49367	6108 (12.4)	6867 (13.9)	17255 (35.0)	19137 (38.8)
2007	44619	5329 (11.9)	6393 (14.3)	15769 (35.3)	17128 (38.4)
2008	36636	4041 (11.0)	5253 (14.3)	13429 (36.7)	13913 (38.0)
≥2009	67832	15714 (23.2)	8801 (13.0)	25363 (37.4)	17954 (26.5)
[a] Number of U&E tests per monitoring period in continuous therapy					

Fig. 6.1: ACEI: proportion of patients appropriately monitored, by practice



A large amount of variation was observed when monitoring practice was split by individual practice. Most practices adhered to recommended monitoring of U&Es in 70-90% of patients, with noticeable negative skew arising from a number of

outlying practices with poor monitoring. Due to limited information regarding the arrangements and infrastructure available to individual GP practices, it was not possible to judge whether the absence or low proportion of monitoring in certain practices was due to inadequate level of care, alternative arrangement for laboratory monitoring (such as secondary care or other provider filling this role) or poor electronic recording of tests.

6.2.2 Estimation of propensity scores

Tables 6.2, 6.5, 6.8, 6.11 and 6.14 summarise the coefficients for each patient factor included in the propensity score models for each medication-specific cohort. The coefficients represent the effect of the covariate on the probability of assignment to one of the more frequent monitoring categories. The coefficients are related to probabilities through a cumulative standard normal distribution. A value >0 represented a positive effect on the probability, whilst <0 a negative effect.

All of the included covariates in the ordered probit model summarised in Table 6.2 had a statistically significant effect on the assigned frequency of monitoring. Patients in all age groups of ≥50 years old at baseline were more likely to have a record of U&E tests compared to patients aged 50 or younger. Female patients were less likely to monitored compared to male patients. Those with 2 comorbid chronic conditions or more were also more likely to be appropriately monitored, compared to patients with no conditions apart from the condition for which an ACEI was indicated. Polypharmacy was a predictor of monitoring, with all categories of patients co-prescribed at least one medication in addition to ACEI therapy more likely to have monitoring present during the observed period. Patients with CKD stage 2, 3 or 4 at baseline were more likely to be monitored compared to those with no recorded diagnosis of CKD. Diagnosis of type I or II diabetes at baseline, as well as co-prescription of either ARB or diuretic was also positively associated with presence of monitoring. Compared to patients who were initiated on ACEI therapy in 2003, those who began therapy in 2005 and 2006 were more likely to be monitored, whilst those first prescribed the drug in 2009 or later were less likely to have regular monitoring. Those patients who were on the drug for longer (2 or 3 valid periods of continuous therapy) were less likely to be monitored compared to those with only one valid period of therapy.

Table 6.2: Coefficient estimates of the probit model, ACEI full dataset

Category/variable name	Coefficient (SE)
Age category	
<50	0
50-59	0.178 (0.007) ***
60-69	0.347 (0.007) ***
70-79	0.482 (0.008) ***
80-89	0.488 (0.009) ***
≥90	0.314 (0.022) ***
Sex	
Male	0
Female	-0.088 (0.005) ***
No. of co-morbidities	
0	0
1	0.012 (0.007)
2	0.054 (0.011) ***
3	0.065 (0.019) **
≥4	0.094 (0.031) **
No. of co-prescribed medications	
0	0
1	0.060 (0.010) ***
2	0.133 (0.010) ***
3	0.216 (0.010) ***
4	0.312 (0.010) ***
5	0.390 (0.010) ***
6	0.488 (0.011) ***
7	0.553 (0.012) ***
8	0.668 (0.014) ***
≥9	0.796 (0.015) ***
CKD stage	
No CKD	0
1	0.109 (0.086)
2	0.146 (0.034) ***
3	0.336 (0.013) ***
4	0.527 (0.065) ***

5	0.042 (0.198)
Type I or II diabetes	
No	0
Yes	0.616 (0.008) ***
Co-prescriptions	
Not co-prescribed ARB	0
Co-prescribed ARB	0.135 (0.006) ***
Not co-prescribed diuretic	0
Co-prescribed diuretic	0.073 (0.005) ***
Year of initiation	
2003	0
2004	0.007 (0.010)
2005	0.021 (0.009) *
2006	0.021 (0.009) *
2007	-0.009 (0.009)
2008	0.005 (0.010)
≥2009	-0.135 (0.009) ***
No. of follow-up periods[1]	
1	0
2	-0.677 (0.008) ***
3	-0.501 (0.006) ***
$\boldsymbol{Likelihood\ ratio\ test}: \boldsymbol{\chi^2(35) = 40970, p < 0.01}$	
[1] Number of periods in continuous therapy until censor event or end of follow-up, cycle length of 900 days	
* p<0.05 ** p<0.01 *** p<0.001	

6.2.3 Time-to-event analysis

The rate of hyperkalaemia was significantly lower in the less frequent group compared to the recommended frequency group. No statistically significant effect of monitoring was found in any other categories. The full details of the hazard ratios and estimated Weibull parameters are outlined in Table 6.3.

Table 6.3: Hazard ratios in Weibull models, ACEI full cohort.

Frequency of monitoring	HR (95% C.I.) of hyperkalaemia in primary care	HR (95% C.I.) of AKI in primary care
Once every 12 months	1	1
Less than every 12 months	0.568 (0.418,0.773)***	0.628 (0.377,1.047)
More than every 12 months	1.003 (0.905,1.112)	0.896 (0.757,1.059)
Weibull parameters	Coefficient (95% C.I.) of hazard function of hyperkalaemia in control group	Coefficient (SE; 95% C.I.) of hazard function of AKI in control group
Scale parameter (ln(λ))	-18.23*** (-18.79,-17.68)	-18.63*** (-19.64,-17.62)
Shape parameter (γ)	1.95*** (1.87,2.03)	1.87*** (1.73,2.01)
HK model Wald test: $\boldsymbol{\chi^2(2) = 13.62, p < 0.01}$		
AKI model Wald test: $\boldsymbol{\chi^2(2) = 4.10, p = 0.129}$		
* p<0.05 ** p<0.01 *** p<0.001		

6.3 ACE inhibitors: subset analysis of HES-linked practices

6.3.1 Descriptive analysis

Around 60% of practices registered on CPRD are linked via the unique patient identifier with the HES database. As a result, it was necessary to produce a separate dataset with the aim of analysing the effect of monitoring on ADEs which required hospital treatment. Within the dataset comprised of data from practices linked to HES, 189251 patients on long-term therapy with an ACE inhibitor were identified from 349 general practices. The mean number of patients per practice was 315 (SD 181). Mean age in the cohort was 63.4 (SD 13.5) and 46.2% were female. Nearly one-fifth of the cohort (19.5%, 36949/189251) had two or more comorbidities at baseline. Polypharmacy was highly prevalent, with around 35% of the cohort (66331/189251) co-prescribed 5 or more medications alongside the index ACEI prescription.

A high proportion of the cohort had records of U&E tests below the recommended frequency of once in 12 months during long-term therapy. Twelve percent (23094/189251) did not have any record of monitoring during treatment and were thus excluded prior to the analysis, in accordance with the cohort selection methods outlined in Chapter 4. Over a third of patients were monitored in accordance with guidelines (36.9%, 69732/189251), whilst 37.6%

(71203/189251) were monitored more frequently than once in 12 months. A subset of patients had a record of some monitoring, albeit below recommended frequency (13.3%, 25222/189251).

Table 6.4: Summary of patient characteristics in ACEI cohort, HES-linked practices.

Variable	All patients	No monitoring	Less frequent than recommended	Recommended frequency	More frequent than recommended
		t=0	0<t<1	1≤t≤1.5	t>1.5
Patient characteristics					
N (% of total)	189251	23094 (12.2)	25222 (13.3)	69732 (36.9)	71203 (37.6)
Mean age (SD)	63.4 (13.5)	60.9 (14.7)	59.0 (13.5)	62.6 (13.1)	66.5 (12.6)
Age category: n (%)					
<50	31020	5388 (17.4)	6353 (20.5)	11842 (38.2)	7437 (24.0)
50-59	42940	5613 (13.1)	7209 (16.8)	16947 (39.5)	13171 (30.7)
60-69	49683	5360 (10.8)	5944 (12.0)	19190 (38.6)	19189 (38.6)
70-79	41921	3905 (9.3)	3709 (8.9)	14227 (33.9)	20080 (47.9)
80-89	21149	2347 (11.1)	1763 (8.3)	6742 (31.9)	10297 (48.7)
≥90	2538	481 (19.0)	244 (9.6)	784 (30.9)	1029 (40.5)
Female sex %	46.2	47.0	44.3	45.6	47.1
Mean number of co-morbid conditions (SD)	0.81 (1.00)	0.71 (0.96)	0.64 (0.88)	0.74 (0.94)	0.97 (1.09)
Number of comorbid conditions: n (%)					
0 or 1	152302	19266 (12.7)	21713 (14.3)	57849 (38.0)	53474 (35.1)
2	24529	2563 (10.5)	2511 (10.2)	8211 (33.5)	11244 (45.8)
3	8427	854 (10.1)	707 (8.4)	2555 (30.3)	4311 (51.2)
4	2883	303 (10.5)	208 (7.2)	837 (29.0)	1535 (53.2)
≥5	1110	108 (9.7)	83 (7.5)	280 (25.2)	639 (57.6)
Mean no. of co-prescriptions (SD)	4.6 (2.5)	3.6 (2.3)	3.9 (2.3)	4.3 (2.4)	5.5 (2.6)
No. of co-prescriptions: n (%)					
0	19303	4389 (22.7)	3486 (18.1)	7799 (40.4)	3629 (18.8)
1	24446	4549 (18.6)	4210 (17.2)	9802 (40.1)	5885 (24.1)
2	26692	3773 (14.1)	4376 (16.4)	10799 (40.5)	7744 (29.0)
3	27081	3148 (11.6)	3955 (14.6)	10688 (39.5)	9290 (34.3)
4	25398	2493 (9.8)	3215 (12.7)	9506 (37.4)	10184 (40.1)
5	21921	1836 (8.4)	2368 (10.8)	7915 (36.1)	9802 (44.7)
6	17321	1277 (7.4)	1643 (9.5)	5778 (33.4)	8623 (49.8)
7	12252	816 (6.7)	1051 (8.6)	3735 (30.5)	6650 (54.3)
8	7633	448 (5.9)	523 (6.9)	2055 (26.9)	4607 (60.4)
≥9	7204	365 (5.1)	395 (5.5)	1655 (23.0)	4789 (66.5)
N (%) co-prescribed ARBs	33713 (17.8)	3572 (15.5)	2549 (10.1)	11648 (16.7)	15944 (22.4)
N (%) co-prescribed loop or thiazide diuretics	75654 (40.0)	7642 (33.1)	8826 (34.5)	26228 (37.6)	32958 (46.3)
N (%) with diagnosis	8429 (4.5)	591 (2.6)	543 (2.2)	2476 (3.6)	4819 (6.8)

of CKD at baseline					
CKD stage (% of total in category)					
1	52	6 (11.5)	5 (9.6)	21 (40.4)	20 (38.5)
2	608	41 (6.7)	64 (10.5)	186 (30.6)	317 (52.1)
3	7432	505 (6.8)	455 (6.1)	2212 (29.8)	4260 (57.3)
4	307	26 (8.5)	15 (4.9)	54 (17.6)	212 (69.1)
5	30	13 (43.3)	4 (13.3)	3 (10.0)	10 (33.3)
N (%) with a diagnosis of diabetes at baseline	22092 (11.7)	1258 (5.4)	903 (3.6)	6123 (8.8)	13808 (19.4)
Year of initiation, n (%)					
2003	18904	2006 (10.6)	2466 (13.0)	6502 (34.4)	7930 (42.0)
2004	23076	2184 (9.5)	3034 (13.2)	8131 (35.2)	9727 (42.2)
2005	23001	2052 (8.9)	3090 (13.4)	8371 (36.4)	9488 (41.3)
2006	29907	2693 (9.0)	4141 (13.9)	10839 (36.2)	12234 (40.9)
2007	27746	2490 (9.0)	3671 (13.2)	10310 (37.2)	11275 (40.6)
2008	23081	2032 (8.8)	3170 (13.7)	8868 (38.4)	9011 (39.0)
≥2009	43536	9637 (22.1)	5650 (13.0)	16711 (38.4)	11538 (26.5)

6.3.2 Estimation of propensity scores

All of the included covariates in the ordered probit model summarised in Table 6.5 had a statistically significant effect on the assigned frequency of monitoring. Patients in all age groups of ≥50 years old at baseline were more likely to have a record of U&E tests compared to patients aged 50 or younger. Female patients were less likely to monitored compared to male patients. Those with 2 comorbid chronic conditions or more were also more likely to be appropriately monitored, compared to patients with no conditions apart from the condition for which an ACEI was indicated. Polypharmacy was a predictor of monitoring, with all categories of patients co-prescribed at least one medication in addition to ACEI therapy more likely to have monitoring present during the observed period. Patients with CKD stage 2, 3 or 4 at baseline were more likely to be monitored compared to those with no recorded diagnosis of CKD. Diagnosis of type I or II diabetes at baseline, as well as co-prescription of either ARB or diuretic was also positively associated with presence of monitoring. Compared to patients who were initiated on ACEI therapy in 2003, those who began therapy in 2004, 2008 and 2009 or later were less likely to be monitored. Those patients who were on the drug for longer (2 or 3 valid periods of continuous therapy) were less likely to be monitored compared to those with only one valid period of therapy.

Table 6.5: Coefficient estimates of the probit model, ACEI HES-linked dataset.

Category/variable name	Coefficient (SE)
Age category	
<50	0
50-59	0.186 (0.009) ***
60-69	0.352 (0.009) ***
70-79	0.503 (0.010) ***
80-89	0.498 (0.012) ***
≥90	0.317 (0.027) ***
Sex	
Male	0
Female	-0.097 (0.006) ***
No. of co-morbidities	
0	0
1	0.021 (0.009) *
2	0.069 (0.015) ***
3	0.076 (0.025) **
≥4	0.092 (0.040) *
No. of co-prescribed medications	
0	0
1	0.065 (0.012) ***
2	0.131 (0.012) ***
3	0.228 (0.012) ***
4	0.328 (0.012) ***
5	0.408 (0.013) ***
6	0.501 (0.014) ***
7	0.585 (0.015) ***
8	0.712 (0.018) ***
≥9	0.859 (0.019) ***
CKD stage	
No CKD	0
1	0.130 (0.172)
2	0.194 (0.051) ***
3	0.354 (0.016) ***
4	0.623 (0.081) ***
5	0.060 (0.293)
Type I or II diabetes	
No	0
Yes	0.652 (0.009) ***
Co-prescriptions	
Not co-prescribed ARB	0

Co-prescribed ARB	0.111 (0.008) ***
Not co-prescribed diuretic	0
Co-prescribed diuretic	0.083 (0.006) ***
Year of initiation	
2003	0
2004	-0.025 (0.012) *
2005	-0.009 (0.012)
2006	-0.007 (0.011)
2007	-0.022 (0.012)
2008	-0.033 (0.012) **
≥2009	-0.187 (0.011) ***
No. of follow-up periods[1]	
1	0
2	-0.688 (0.010) ***
3	-0.492 (0.008) ***
$\boldsymbol{Likelihood\ ratio\ test}$: $\boldsymbol{\chi^2(35) = 27801, p < 0.01}$	
[1] Number of periods in continuous therapy until censor event or end of follow-up, cycle length of 900 days	
* p<0.05 ** p<0.01 *** p<0.001	

6.3.3 Time-to-event analysis

The group of patients who were monitored less frequently than recommended had a lower hazard rate of hospital admission for hyperkalaemia (HR 0.13; 95% CI 0.02,0.95), after adjusting for monitoring group assignment. Patients with frequent monitoring had a significantly higher rate of hospital admission for AKI (HR 1.56, CI 1.23,1.99). The full details of the hazard ratios and estimated Weibull parameters are outlined in Table 6.6.

Table 6.6: Hazard ratios in Weibull models, ACEI HES-linked practices

Frequency of monitoring	HR (95% C.I.) of hospitalisation for hyperkalaemia	HR (95% C.I.) of hospitalisation for AKI
Once every 12 months	1	1
Less than every 12 months	0.127 (0.017,0.951)*	1.170 (0.609,2.249)
More than every 12 months	0.921 (0.492,1.723)	1.560 (1.225,1.987)***
Weibull parameters	Coefficient (SE; 95% C.I.) of hazard function of hyperkalaemia at baseline	Coefficient (SE; 95% C.I.) of hazard function of AKI at baseline
Scale parameter (ln(λ))	-19.69*** (-22.30,-17.10)	-15.21*** (-16.52,-13.89)
Shape parameter (γ)	1.70*** (1.37,2.12)	1.31*** (1.13,1.51)
HK model Wald test: $\boldsymbol{\chi^2(2) = 4.06, p = 0.131}$		
AKI model Wald test: $\boldsymbol{\chi^2(2) = 13.24, p < 0.01}$		
* p<0.05 ** p<0.01 *** p<0.001		

6.4 Loop diuretics: full cohort

6.4.1 Descriptive analysis

Within the full dataset of patients on long-term therapy with a loop diuretic, 82178 patients were identified from 615 general practices, as shown in Table 6.7. The mean number of patients per practice was 134 (SD 89). Mean age in the cohort was 72.3 (SD 13.4) and 55.8% were female. Multimorbidity was common, with nearly 1 in 3 patients having 2 or more comorbid conditions at baseline (31%, 25449/82178). The majority of patients in the cohort were co-prescribed 5 or more medications from different BNF chapters during the exposure period (61%, 50250/82178) and 1 in 10 (10%, 8173/82178) were co-prescribed an additional 9 medications or more.

A high proportion of the cohort had records of U&E tests below the recommended frequency of once in 12 months during long-term therapy. Nearly one in five (18.4%, 15098/82178) did not have any record of monitoring during treatment and were thus excluded prior to the analysis, in accordance with the cohort selection methods outlined in Chapter 4. Over a quarter of the cohort was monitored in accordance with guidelines (28.8%, 23628/82178), whilst 43.1% (35394/82178) were monitored more frequently than once in 12 months. Around

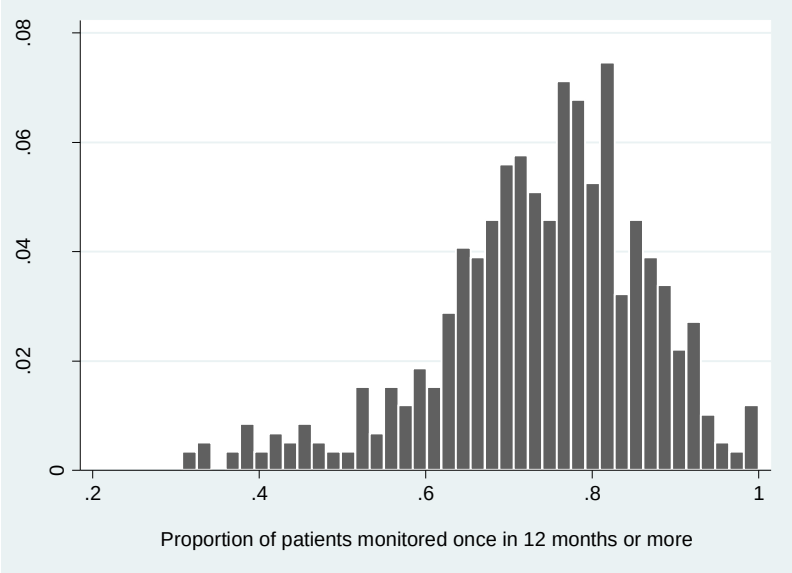
one in ten had a record of some monitoring, albeit below recommended frequency (9.8%, 8058/82178).

Table 6.7: Summary of patient characteristics, loop diuretic full cohort

Variable	All patients	No monitoring	Less frequent than recommended	Recommended frequency	More frequent than recommended
		t=0	0<t<1	1≤t≤1.5	t>1.5
Patient characteristics					
N (% of total)	82178	15098 (18.4)	8058 (9.8)	23628 (28.8)	35394 (43.1)
Mean age (SD)	72.3 (13.1)	71.3 (15.8)	70.9 (14.5)	72.4 (13.0)	73.0 (11.5)
Age category, n (%)					
<50	5350	1696 (31.7)	795 (14.9)	1444 (27.0)	1415 (26.5)
50-59	8106	1763 (21.8)	934 (11.5)	2389 (29.5)	3020 (37.3)
60-69	15967	2472 (15.5)	1568 (9.8)	4687 (29.4)	7240 (45.3)
70-79	25364	3627 (14.3)	2121 (8.4)	7078 (27.9)	12538 (49.4)
80-89	22600	4140 (18.3)	2126 (9.4)	6585 (29.1)	9749 (43.1)
≥90	4791	1400 (29.2)	514 (10.7)	1445 (30.2)	1432 (29.9)
Female sex %	55.8	60.8	62.4	57.1	51.2
Mean number of co-morbid conditions (SD)	1.12 (1.17)	1.07 (1.18)	0.99 (1.10)	1.08 (1.15)	1.20 (1.20)
Number of comorbid conditions, n (%)					
0 or 1	56729	10686 (18.8)	5926 (10.5)	16623 (29.3)	23494 (41.4)
2	15152	2588 (17.1)	1330 (8.8)	4261 (28.1)	6973 (46.0)
3	6752	1201 (17.8)	556 (8.2)	1804 (26.7)	3191 (47.3)
4	2536	419 (16.5)	172 (6.8)	714 (28.2)	1231 (48.5)
5	766	152 (19.8)	61 (8.0)	180 (23.5)	373 (48.7)
≥6	243	52 (21.4)	13 (5.4)	46 (18.9)	132 (54.3)
Mean no. of co-prescriptions (SD)	6.3 (2.5)	5.3 (2.5)	6.0 (2.4)	6.0 (2.4)	6.9 (2.4)
No. of co-prescriptions, n (%)					
0	1658	696 (42.0)	161 (9.7)	501 (30.2)	300 (18.1)
1	3855	1309 (34.0)	402 (10.4)	1226 (31.8)	918 (23.8)
2	6342	1810 (28.5)	712 (11.2)	2046 (32.3)	1774 (28.0)
3	9013	2121 (23.5)	1002 (11.1)	2850 (31.6)	3040 (33.7)
4	11060	2317 (21.0)	1171 (10.6)	3451 (31.2)	4121 (37.3)
5	12438	2089 (16.8)	1248 (10.0)	3713 (29.9)	5388 (43.3)
6	12041	1796 (14.9)	1171 (9.7)	3396 (28.2)	5678 (47.2)
7	10149	1275 (12.6)	914 (9.0)	2779 (27.4)	5181 (51.1)
8	7449	864 (11.6)	655 (8.8)	1853 (24.9)	4077 (54.7)
≥9	8173	821 (10.1)	622 (7.6)	1813 (22.2)	4917 (60.2)
N (%) co-prescribed ACEI	41860 (50.9)	5095 (33.7)	3223 (40.0)	11669 (49.4)	21873 (61.8)
N (%) co-prescribed ARBs	16818 (20.5)	1850 (12.3)	1235 (15.3)	4435 (18.8)	9298 (26.3)
N (%) co-prescribed thiazide diuretics	9643 (11.7)	1400 (9.3)	662 (8.2)	2726 (11.5)	4855 (13.7)
N (%) with diagnosis of CKD at baseline	4070 (5.0)	511 (3.4)	229 (2.8)	979 (4.1)	2351 (6.6)

CKD stage (% of total in category)					
1	34	7 (20.6)	4 (11.8)	10 (29.4)	13 (38.2)
2	316	32 (10.1)	24 (7.6)	97 (30.7)	163 (51.6)
3	3124	373 (11.9)	160 (5.1)	768 (24.6)	1823 (58.4)
4	491	66 (13.4)	30 (6.1)	81 (16.5)	314 (64.0)
5	105	33 (31.4)	11 (10.5)	23 (21.9)	38 (36.2)
N (%) with a diagnosis of diabetes at baseline	8867 (10.8)	676 (4.5)	290 (3.6)	1976 (8.4)	5925 (16.7)
Year of initiation, n (%)					
2003	11019	2200 (20.0)	1394 (12.7)	3126 (28.4)	4299 (39.0)
2004	10923	1859 (17.0)	1217 (11.1)	3086 (28.2)	4762 (43.6)
2005	10566	1773 (16.8)	1108 (10.5)	2902 (27.5)	4783 (45.3)
2006	10700	1688 (15.8)	1054 (9.9)	2936 (27.4)	5022 (46.9)
2007	9818	1488 (15.2)	957 (9.8)	2671 (27.2)	4702 (47.9)
2008	9946	1453 (14.6)	931 (9.4)	2769 (27.8)	4793 (48.2)
≥2009	19206	4637 (24.1)	1397 (7.3)	6139 (32.0)	7033 (36.6)

Fig. 6.2: Loop diuretics: proportion of patients appropriately monitored, by practice



A large amount of variation was observed when monitoring compliance was split by individual practice. Most practices adhered to recommended monitoring of U&Es in 60-90% of patients, with noticeable negative skew arising from a number of outlying practices with poor monitoring.

6.4.2 Estimation of propensity scores

All of the included covariates in the ordered probit model summarised in Table 6.8 had a statistically significant effect on the assigned frequency of monitoring. Patients in all age groups of ≥50 years old at baseline were more likely to have a record of U&E tests compared to patients aged 50 or younger. Female patients

were less likely to monitored compared to male patients. Those with 4 comorbid chronic conditions or more were also more likely to be appropriately monitored, compared to patients with no conditions apart from the condition for which a loop diuretic was indicated. Polypharmacy was a predictor of monitoring, with all categories of patients co-prescribed at least two medication in addition to loop diuretic therapy more likely to have monitoring present during the observed period. Patients with CKD stage 3 or 4 at baseline were more likely to be monitored compared to those with no recorded diagnosis of CKD. Diagnosis of type I or II diabetes at baseline, as well as co-prescription of ACE, ARB or thiazide diuretic was also positively associated with presence of monitoring. Compared to patients who were initiated on loop diuretic therapy in 2003, those who began therapy in 2004 or later were more likely to be monitored. Those patients who were on the drug for longer (2 or 3 valid periods of continuous therapy) were less likely to be monitored compared to those with only one valid period of therapy.

Table 6.8: Coefficient estimates of the probit model, loop diuretics full cohort

Category/variable name	Coefficient (SE)
Age category	
<50	0
50-59	0.198 (0.024) ***
60-69	0.329 (0.022) ***
70-79	0.441 (0.021) ***
80-89	0.380 (0.021) ***
≥90	0.221 (0.027) ***
Sex	
Male	0
Female	-0.142 (0.009) ***
No. of co-morbidities	
0	0
1	0.014 (0.012)
2	0.032 (0.017)
3	0.039 (0.027)
4	0.082 (0.050) *
≥5	0.287 (0.094) **
No. of co-prescribed medications	
0	0
1	0.079 (0.043)
2	0.140 (0.040) ***
3	0.243 (0.039) ***

4	0.322 (0.038) ***
5	0.439 (0.038) ***
6	0.512 (0.038) ***
7	0.579 (0.039) ***
8	0.660 (0.040) ***
≥9	0.793 (0.040) ***
CKD stage	
No CKD	0
1	-0.279 (0.222)
2	0.018 (0.072)
3	0.264 (0.025) ***
4	0.451 (0.065) ***
5	-0.092 (0.137)
Type I or II diabetes	
No	0
Yes	0.435 (0.016) ***
Co-prescriptions	
Not co-prescribed ACEI	0
Co-prescribed ACEI	0.428 (0.010) ***
Not co-prescribed ARB	0
Co-prescribed ARB	0.366 (0.012) ***
Not co-prescribed thiazide diuretic	0
Co-prescribed thiazide diuretic	0.052 (0.015) ***
Year of initiation	
2003	0
2004	0.078 (0.018) ***
2005	0.117 (0.018) ***
2006	0.142 (0.018) ***
2007	0.116 (0.018) ***
2008	0.154 (0.018) ***
≥2009	0.055 (0.016) **
No. of follow-up periods[1]	
1	0
2	-0.582 (0.012) ***
3	-0.441 (0.011) ***
$\boldsymbol{Likelihood\ ratio\ test}$: $\boldsymbol{\chi^2(35) = 9524, p < 0.01}$	

[1] Number of periods in continuous therapy until censor event or end of follow-up, cycle length of 900 days

* p<0.05 ** p<0.01 *** p<0.001

6.4.3 Time-to-event analysis

The rate of hyponatraemia or hypokalaemia diagnosed in primary care was lower in patients monitored less frequently (HR 0.17, C.I. 0.06,0.45) and more frequently (HR 0.53, C.I. 0.35,0.82) compared to patients monitored according to guidelines. The full details of the hazard ratios and estimated Weibull parameters are outlined in Table 6.9.

Table 6.9: Hazard ratios in Weibull models, loop diuretic full cohort

Frequency of monitoring	HR (95% C.I.) of diagnosis of hyponatraemia or hypokalaemia
Once every 12 months	1
Less than every 12 months	0.167 (0.063,0.445)***
More than every 12 months	0.533 (0.347,0.817)**
Weibull parameters	**Coefficient (95% C.I.) of hazard function of hyperkalaemia at baseline**
Scale parameter (ln(λ))	-17.93*** (-19.75,-16.11)
Shape parameter (γ)	1.76*** (1.52,2.03)
Model Wald test: $\chi^2(2) = 16.82, p < 0.001$	
* p<0.05 ** p<0.01 *** p<0.001	

6.5 Loop diuretics: subset analysis of HES-linked practices

6.5.1 Descriptive analysis

Around 60% of practices registered on CPRD are linked via the unique patient identifier with the HES database. As a result, it was necessary to produce a separate dataset with the aim of analysing the effect of monitoring on ADEs which required hospital treatment. Within the dataset from practices linked to the HES database, 49072 patients on long-term therapy with an ACE inhibitor were identified from 357 general practices. The mean number of patients per practice was 137 (SD 84). The descriptive statistics of this cohort are presented in Table 6.10 below. Mean age in the cohort was 72.6 (SD 11.4) and 55.7% were female. Multimorbidity was common, with 30% of patients (14759/49072) having 2 or more comorbid conditions at baseline. Polypharmacy was also common in the cohort, with 60% of the cohort co-prescribed 5 or more medications from different BNF chapters during the exposure period (29370/49072), including the

index prescription. Nearly one in five (18%, 8842/49072) were co-prescribed 8 or more different medications.

A high proportion of the cohort had records of U&E tests below the recommended frequency of once in 12 months during long-term therapy. Sixteen percent (7875/49072) did not have any record of monitoring during treatment and were thus excluded prior to the analysis, in accordance with the cohort selection methods outlined in Chapter 4. Around 30% of patients were monitored in accordance with guidelines (14763/49072), whilst 44.5% (21842/49072) were monitored more frequently than once in 12 months. A subset of patients had a record of some monitoring, albeit below recommended frequency (9.4%, 4592/49072).

Table 6.10: Summary of patient characteristics in loop diuretic cohort, HES-linked practices

Variable	All patients	No monitoring	Less frequent than recommended	Recommended frequency	More frequent than recommended
		t=0	0<t<1	1≤t≤1.5	t>1.5
Patient characteristics					
N (% of total)	49072	7875 (16.1)	4592 (9.4)	14763 (30.1)	21842 (44.5)
Mean age (SD)	72.6 (11.4)	71.7 (15.8)	71.2 (14.8)	72.6 (13.0)	73.2 (11.4)
Age category: n (%)					
<50	3004	867 (28.9)	455 (15.2)	862 (28.7)	820 (27.3)
50-59	4746	896 (18.9)	530 (11.2)	1491 (31.4)	1829 (38.5)
60-69	9470	1242 (13.1)	858 (9.1)	2925 (30.9)	4445 (46.9)
70-79	15103	1824 (12.1)	1165 (7.7)	4397 (29.1)	7717 (51.1)
80-89	13803	2259 (16.4)	1254 (9.1)	4135 (30.0)	6155 (44.6)
≥90	2946	787 (26.7)	330 (11.2)	953 (32.3)	876 (29.7)
Female sex %	55.7	61.2	63.0	57.2	51.2
Mean number of co-morbid conditions (SD)	1.10 (1.16)	1.01 (1.15)	0.96 (1.10)	1.05 (1.13)	1.19 (1.19)
Number of comorbid conditions, n (%)					
0 or 1	34313	5730 (16.7)	3430 (10.0)	10488 (30.6)	14665 (42.7)
2	8829	1267 (14.4)	723 (8.2)	2630 (29.8)	4209 (47.7)
3	3913	583 (14.9)	306 (7.8)	1106 (28.3)	1918 (49.0)
4	1463	208 (14.2)	89 (6.1)	416 (28.4)	750 (51.3)
5	413	67 (16.2)	36 (8.7)	95 (23.0)	215 (52.1)
≥6	141	20 (14.2)	8 (5.7)	28 (19.9)	85 (60.3)
Mean no. of co-prescriptions (SD)	6.2 (2.5)	2.1 (2.4)	5.9 (2.4)	5.9 (2.4)	6.8 (2.4)
No. of co-prescriptions, n (%)					
0	1063	429 (40.4)	91 (8.6)	339 (31.9)	204 (19.2)
1	2419	771 (31.9)	262 (10.8)	779 (32.2)	607 (25.1)

2	3930	1026 (26.1)	451 (11.5)	1341 (34.1)	1112 (28.3)
3	5552	1210 (21.8)	596 (10.7)	1838 (33.1)	1908 (34.4)
4	6738	1202 (17.8)	694 (10.3)	2228 (33.1)	2614 (38.8)
5	7385	1039 (14.1)	673 (9.1)	2316 (31.4)	3357 (45.5)
6	7221	919 (12.7)	664 (9.2)	2100 (29.1)	3538 (49.0)
7	5922	605 (10.2)	532 (9.0)	1663 (28.1)	3122 (52.7)
8	4311	375 (8.7)	330 (7.7)	1099 (25.5)	2507 (58.2)
≥9	4531	299 (6.6)	299 (6.6)	1060 (23.4)	2873 (63.4)
N (%) co-prescribed ACEI	25199 (51.4)	2528 (32.1)	1777 (38.7)	7352 (49.8)	13542 (62.0)
N (%) co-prescribed ARBs	10241 (20.9)	891 (11.3)	658 (14.3)	2804 (19.0)	5888 (27.0)
N (%) co-prescribed thiazide diuretics	5818 (11.9)	700 (8.9)	354 (7.7)	1687 (11.4)	3077 (14.1)
N (%) with diagnosis of CKD at baseline	2391 (4.9)	232 (2.9)	105 (2.3)	596 (4.0)	1458 (6.7)
CKD stage (% of total in category)					
1	15	1 (6.7)	2 (13.3)	4 (26.7)	8 (53.3)
2	159	16 (10.1)	9 (5.7)	51 (32.1)	83 (52.2)
3	1867	163 (8.7)	76 (4.1)	485 (26.0)	1143 (61.2)
4	294	33 (11.2)	13 (4.4)	46 (15.7)	202 (68.7)
5	56	19 (33.9)	5 (8.9)	10 (17.9)	22 (39.3)
N (%) with a diagnosis of diabetes at baseline	5465 (11.1)	308 (3.9)	113 (2.5)	1211 (8.2)	3833 (17.5)
Year of initiation, n (%)					
2003	6582	1135 (17.2)	772 (11.7)	1944 (29.5)	2731 (41.5)
2004	6549	948 (14.5)	709 (10.8)	1956 (29.9)	2936 (44.8)
2005	6049	810 (13.4)	624 (10.3)	1742 (28.8)	2873 (47.5)
2006	6301	800 (12.7)	564 (9.0)	1830 (29.0)	3107 (49.3)
2007	5815	657 (11.3)	513 (8.8)	1702 (29.3)	2943 (50.6)
2008	5901	720 (12.2)	537 (9.1)	1733 (29.4)	2911 (49.3)
≥2009	11875	2805 (23.6)	873 (7.4)	3856 (32.5)	4341 (36.6)

6.5.2 Estimation of propensity scores

All of the included covariates in the ordered probit model summarised in Table 6.11 had a statistically significant effect on the assigned frequency of monitoring. Patients in all age groups of ≥50 years old at baseline were more likely to have a record of U&E tests compared to patients aged 50 or younger. Female patients were less likely to be monitored compared to male patients. Those with 5 comorbid chronic conditions or more were also more likely to be appropriately monitored, compared to patients with no conditions apart from the condition for which a loop diuretic was indicated. Polypharmacy was a predictor of monitoring, with all categories of patients co-prescribed at least three medications in addition

to loop diuretic therapy more likely to have monitoring present during the observed period. Patients with CKD stage 3 or 4 at baseline were more likely to be monitored compared to those with no recorded diagnosis of CKD, as were patients with a diagnosis of type I or II diabetes at baseline. Patients co-prescribed ACEI, ARB or thiazide diuretic was also more likely to receive regular monitoring. Compared to patients who were initiated on loop diuretic therapy in 2003, those who began therapy in 2005 and onwards were more likely to be monitored. Those patients who were on the drug for longer (2 or 3 valid periods of continuous therapy) were less likely to be monitored compared to those with only one valid period of therapy.

Table 6.11: Coefficient estimates of the probit model, loop diuretics HES-linked dataset

Category/variable name	Coefficient (SE)
Age category	
<50	0
50-59	0.215 (0.031) ***
60-69	0.358 (0.028) ***
70-79	0.468 (0.027) ***
80-89	0.408 (0.027) ***
≥90	0.205 (0.035) ***
Sex	
Male	0
Female	-0.149 (0.012) ***
No. of co-morbidities	
0	0
1	0.009 (0.016)
2	0.023 (0.023)
3	0.061 (0.036)
4	0.099 (0.068)
≥5	0.278 (0.119) *
No. of co-prescribed medications	
0	0
1	0.044 (0.053)
2	0.082 (0.049)
3	0.202 (0.048) ***
4	0.292 (0.047) ***
5	0.426 (0.047) ***
6	0.486 (0.047) ***
7	0.531 (0.048) ***

8	0.663 (0.049) ***
≥9	0.786 (0.049) ***
CKD stage	
No CKD	0
1	-0.205 (0.314)
2	0.010 (0.103)
3	0.295 (0.033) ***
4	0.591 (0.086) ***
5	0.122 (0.197)
Type I or II diabetes	
No	0
Yes	0.497 (0.020) ***
Co-prescriptions	
Not co-prescribed ACEI	0
Co-prescribed ACEI	0.438 (0.012) ***
Not co-prescribed ARB	0
Co-prescribed ARB	0.392 (0.015) ***
Not co-prescribed thiazide diuretic	0
Co-prescribed thiazide diuretic	0.071 (0.019) ***
Year of initiation	
2003	0
2004	0.035 (0.023)
2005	0.079 (0.023) **
2006	0.116 (0.023) ***
2007	0.096 (0.024) ***
2008	0.103 (0.023) ***
≥2009	0.002 (0.021)
No. of follow-up periods[1]	
1	0
2	-0.577 (0.016) ***
3	-0.413 (0.014) ***
$\boldsymbol{Likelihood\ ratio\ test}: \boldsymbol{\chi^2(35) = 6391, p < 0.01}$	

[1] Number of periods in continuous therapy until censor event or end of follow-up, cycle length of 900 days

* p<0.05 ** p<0.01 *** p<0.001

6.5.3 Time-to-event analysis

The group of patients monitored more frequently had a significantly lower rate of hospital admission for hyponatraemia or hypokalaemia (HR 0.48, C.I. 0.32,0.73) compared to patients monitored according to guidelines. The full details of the hazard ratios and estimated Weibull parameters are outlined in Table 6.12.

Table 6.12: Hazard ratios in Weibull model, loop diuretics HES-linked practices

Frequency of monitoring	HR (95% C.I.) of hospitalisation for hyponatraemia or hypokalaemia	HR (SE; 95% C.I.) of hospitalisation for fall-related injury
Once every 6 months	1	1
Less than every 6 months	0.252 (0.063,1.005)	0.739 (0.417,1.308)
More than every 6 months	0.479 (0.315,0.727)**	0.980 (0.754,1.273)
Weibull parameters	Coefficient (95% C.I.) of hazard function of hyponatraemia or hypokalaemia in control group	Coefficient (95% C.I.) of hazard function of fall-related injury in control group
Scale parameter (ln(λ))	-16.91*** (-18.44,-15.37)	-14.10*** (-15.14,-13.05)
Shape parameter (γ)	1.69*** (1.48,1.92)	1.36*** (1.22,1.52)
HN/HK model Wald test: $\boldsymbol{\chi^2(2) = 13.66, p < 0.01}$ Falls model Wald test: $\boldsymbol{\chi^2(2) = 1.11, p = 0.574}$ * p<0.05 ** p<0.01 *** p<0.001		

6.6 Amiodarone cohort

6.6.1 Descriptive analysis

A total of 8349 patients on long-term therapy with amiodarone were identified from 602 general practices. The mean number of patients per practice was 14 (SD 10). The descriptive statistics of this cohort is summarised in Table 6.13. Mean age in the cohort was 71 (SD 11) and 38.5% were female. A large proportion of patients had multiple co-morbidities: 47% (3905/8349) had 2 or more comorbidities. Polypharmacy was also highly prevalent, with 38% (3136/8349) of the cohort with 5 or more co-prescribed medications from different BNF chapters alongside the index amiodarone prescription.

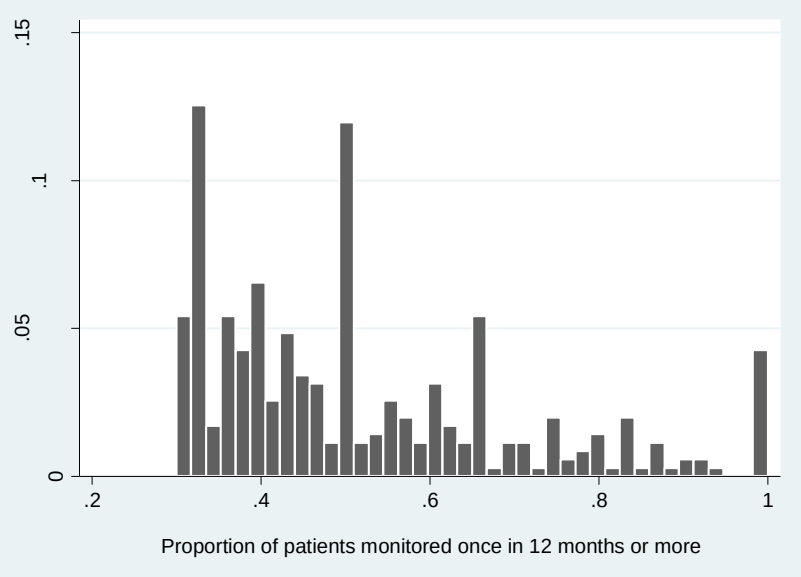
Over a third of the patients had no record of any thyroid function monitoring during the observed period and were excluded from the time-to-event analysis (33.6%, 2804/8349). Just 3116 patients (37.3%) were monitored at least once every 6 months, with TFT tests carried out once every 6 months in 25% (2095/8349) of patients. The conduct of regular monitoring varied across subsets of patients with different characteristics, as demonstrated in Table 6.13. One example is age: amongst the groups aged under 50 and 50-59 the proportion monitored every 6 months or more was 31.6% and 30.9%, respectively. In comparison, patients aged 60-69 and 70-79 were monitored in accordance with guidelines or more intensively 37.8% and 39.9% of the time, respectively. In older population this proportion fell, notably for patients aged 90 or above the proportion was 34.2%. The proportion monitored every 6 months or more was positively related to the number of comorbidities and co-prescriptions: 35.6% of patients with 1 condition (no comorbidities) compared to 41.1% of patients with 5 conditions or more; 27.7% of patients with no co-prescriptions alongside amiodarone compared to 41.7% in patients with 5 or more concurrently prescribed medications. The proportion monitored in accordance with guidelines steadily increased over the calendar years in the study, rising from 32.2% in 2003 to 42.8% in 2008 and onwards.

Table 6.13: Summary of patient characteristics, amiodarone cohort

Variable	All patients	No monitoring	Less frequent than recommended	Recommended frequency	More frequent than recommended
		t=0	0<t<1	1≤t≤1.5	t>1.5
Patient characteristics					
N (% of total)	8349	2804 (33.6)	2429 (29.1)	2095 (25.1)	1021 (12.2)
Mean age (SD)	71.3 (11.4)	70.8 (12.1)	70.2 (11)	71.4 (11.2)	72.5 (9.9)
					Age category; n (%)
<50	352	152 (43.2)	89 (25.3)	83 (23.6)	28 (8)
50-59	877	325 (37.1)	281 (32)	198 (22.6)	73 (8.3)
60-69	2031	668 (32.9)	594 (29.2)	527 (25.9)	242 (11.9)
70-79	3029	958 (31.6)	864 (28.5)	787 (26)	420 (13.9)
80-89	1832	609 (33.2)	543 (29.6)	435 (23.7)	245 (13.4)
≥90	228	92 (40.4)	58 (25.4)	65 (28.5)	13 (5.7)
Female sex %	3215 (38.5)	989 (35.3)	876 (36.1)	859 (41)	491 (48.1)
Mean comorbidity score (SD)	1.63 (1.26)	1.57 (1.24)	1.58 (1.24)	1.66 (1.26)	1.82 (1.32)
					Comorbidity score
1	4444	1516 (34.1)	1347 (30.3)	1112 (25)	469 (10.6)
2	2122	732 (34.5)	595 (28)	511 (24.1)	284 (13.4)
3	1108	347 (31.3)	313 (28.2)	288 (26)	160 (14.4)

4	439	134 (30.5)	110 (25.1)	126 (28.7)	44 (10)
≥5	236	75 (31.8)	64 (27.1)	58 (24.6)	39 (16.5)
Mean no. of co-prescribed medications (SD)	4 (2.1)	3.9 (2.1)	3.9 (2.1)	4 (2.1)	4.4 (2.1)
					No. of co-prescriptions
1	927	365 (39.4)	305 (32.9)	188 (20.3)	69 (7.4)
2	1349	454 (33.7)	409 (30.3)	349 (25.9)	137 (10.2)
3	1498	524 (35)	395 (26.4)	401 (26.8)	178 (11.9)
4	1439	467 (32.5)	439 (30.5)	363 (25.2)	170 (11.8)
5	1189	366 (30.8)	349 (29.4)	313 (26.3)	161 (13.5)
6	863	273 (31.6)	255 (29.5)	209 (24.2)	126 (14.6)
≥7	1084	355 (32.7)	277 (25.6)	272 (25.1)	180 (16.6)
					Year of initiation
2003	1415	521 (36.8)	438 (31)	324 (22.9)	132 (9.3)
2004	1405	494 (35.2)	440 (31.3)	308 (21.9)	163 (11.6)
2005	1211	424 (35)	356 (29.4)	293 (24.2)	138 (11.4)
2006	1162	374 (32.2)	347 (32.2)	286 (24.6)	155 (13.3)
2007	1060	340 (32.1)	299 (28.2)	287 (27.1)	134 (12.6)
≥2008	2096	651 (31.1)	549 (26.2)	597 (28.5)	299 (14.3)

Fig. 6.3: Amiodarone: proportion of patients appropriately monitored, by practice



A large amount of variation was observed when monitoring practice was split by individual practice. Adherence to thyroid function monitoring ranged from 30-90% of patients, with noticeable positive skew arising from a number of outlying practices with better-than-average monitoring. This finding suggests that there is significant disagreement across individual GPs on the correct approach to monitoring thyroid function in amiodarone therapy.

6.6.2 Estimation of propensity scores

The probit model identified a number of patient characteristics associated with more frequent monitoring of thyroid function during amiodarone therapy. Older and female patients were more likely to be closely monitored. In terms of multimorbidity, the results were mixed: compared with patients with no comorbidities, those with 1 or 3 comorbid conditions were more likely to be monitored in accordance with guidelines, but no effect was found in the other comorbidity groups. Unlike in the case of ACE inhibitor monitoring, co-prescribed medications were not found to be significantly associated with more intensive monitoring of thyroid function in amiodarone. The year of prescription was also independently associated, with patients who were first prescribed amiodarone in 2006 or later more likely to receive regular monitoring. The number of follow-up periods was positively associated (opposite of the result in the ACE inhibitor cohort), suggesting that the longer a patient is on the drug, the more likely they are to be monitored in line with guidelines.

Table 6.14: Coefficient estimates in probit model, amiodarone cohort

Category/variable name	Coefficient (SE)
Age	0.004 (0.001)**
Sex	
Male	0
Female	0.192 (0.026) ***
No. of co-morbidities	
0	0
1	0.093 (0.035) **
2	0.046 (0.039)
3	0.128 (0.046) **
4	0.104 (0.063)
≥5	0.040 (0.081)
No. of co-prescribed medications	
0	0
1	-0.007 (0.208)
2	0.171 (0.207)
3	0.178 (0.207)
4	0.218 (0.207)
5	0.250 (0.207)
6	0.259 (0.208)
7	0.284 (0.210)
≥8	0.284 (0.211)
Year of initiation	

2003	0
2004	0.022 (0.041)
2005	0.012 (0.043)
2006	0.108 (0.044)*
2007	0.106 (0.045)*
≥2008	0.186 (0.038)***
No. of follow-up periods[1]	
1	0
2	0.136 (0.036) ***
3	0.281 (0.40) ***
4	0.126 (0.029)***
Presence of risk factors for MACE	
Hypertension	-0.050 (0.053)
Diabetes	0.104 (0.065)
Hyperlipidemia	0.103 (0.109)
Ischaemic heart disease	-0.002 (0.048)
Ventricular arrhythmia	0.059 (0.048)
$\boldsymbol{Likelihood\ ratio\ test}: \boldsymbol{\chi^2(35) = 6391, p < 0.01}$	

[1] Number of periods in continuous therapy until censor event or end of follow-up, cycle length of 900 days

* p<0.05 ** p<0.01 *** p<0.001

6.6.3 Time-to-event analysis

The weighted Weibull models estimating the hazard ratio of AIH, AIT and MACE across comparison groups with different frequency of monitoring have demonstrated statistically significant effects in a number of categories. In patients monitored less frequently than 6 months compared with those monitored once in 6 months, the risk of both AIH (HR 2.79; C.I. 1.42,5.74), AIT (2.57; C.I. 1.40,4.74) and MACE (2.13, C.I. 1.31,3.45) was higher. Monitoring above the recommended frequency resulted in an increase in the rate of AIH diagnosed in primary care (HR 1.53, C.I. 1.08,2.17).

Table 6.15: Coefficient estimates in Weibull models, amiodarone cohort

Frequency of monitoring	HR (95% C.I.) of diagnosis of AIH	HR (95% C.I.) of diagnosis of AIT	HR (95% C.I.) of emergency admission for MACE
Once every 6 months	1	1	1
Less than every 6 months	2.785 (1.418,5.470)**	2.573 (1.397,4.740)**	2.125 (1.308,3.454)**
More than every 6 months	1.530 (1.079,2.168)*	1.231 (0.884,1.714)	1.454 (0.793,2.665)
Weibull parameters	**Coefficient (95% C.I.) of hazard function of AIH**	**Coefficient (95% C.I.) of hazard function AIT**	**Coefficient (95% C.I.) of hazard function of MACE**
Scale parameter (ln(λ))	-14.952 (-16.994,-12.910)***	-18.980 (-21.199,-16.760)***	-14.886 (-18.397,-11.376)***
Shape parameter (γ)	1.801 (1.519,2.135)***	2.451 (2.139,2.808)***	1.656 (1.200,2.286)***
AIH model Wald test: $\boldsymbol{\chi^2(2) = 11.72, p < 0.01}$			
AIT model Wald test: $\boldsymbol{\chi^2(2) = 9.48, p < 0.01}$			
MACE model Wald test: $\boldsymbol{\chi^2(2) = 9.26, p < 0.01}$			
* p<0.05 ** p<0.01 *** p<0.001			

6.7 Interval of common support

Each cohort was examined in order to ensure that the propensity scores of included patients are within the interval of common support to ensure comparability across comparison groups. The distributions of propensity scores in each medication cohort are shown in Figures 6.4 to 6.8 Analytic weights were assigned to all patients whose propensity scores were within the interval of common support using the MMWS method described in Chapter 5. Any patients outside this interval were assigned a weight of zero. The weights were assigned prior to running each Weibull regression used in the time-to-event analysis. Details are included in Table 6.16.

Fig. 6.4: Distribution of propensity scores, ACEI full cohort

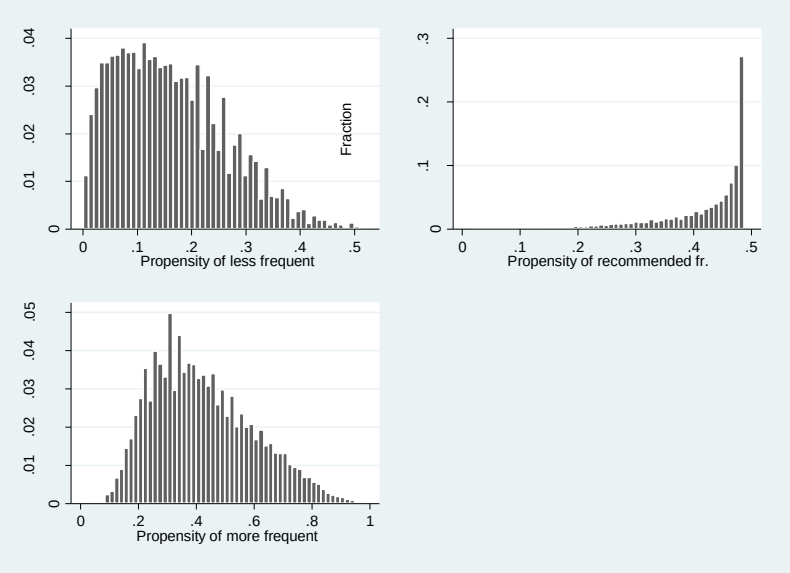


Fig. 6.5: Distribution of propensity scores, ACEI subset of HES-linked practices

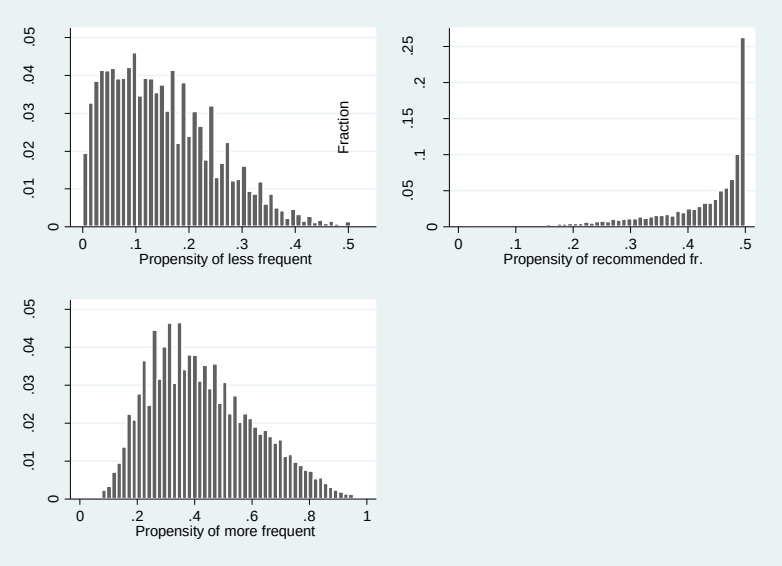


Fig. 6.6: Distribution of propensity scores, loop diuretic full cohort

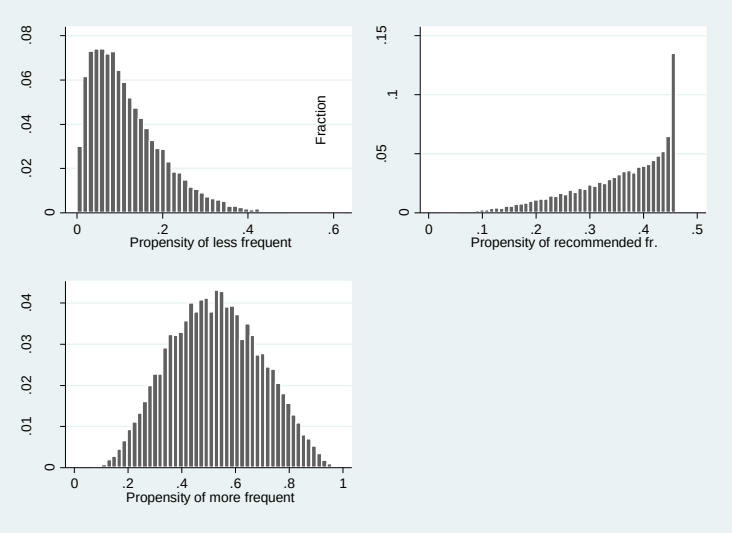


Fig. 6.7: Distribution of propensity scores, loop diuretic HES subset

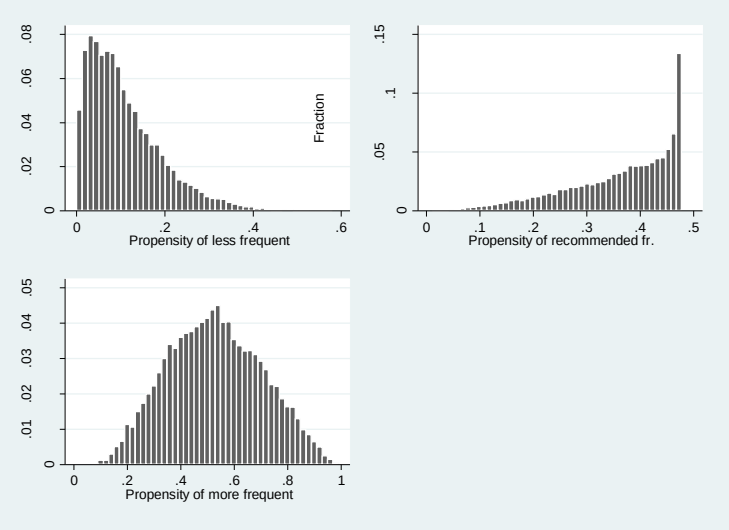


Fig. 6.8: Distribution of propensity scores, amiodarone cohort

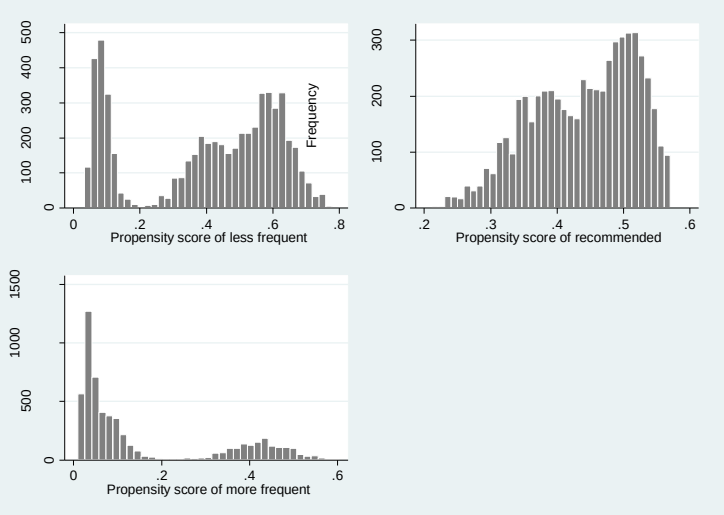


Table 6.16: Exclusion of patients with propensity scores outside common support

Category	ACEI	loop diuretics	amiodarone (AIH/AIT)	amiodarone (MACE)
	N (%)	N (%)	N (%)	N (%)
Original	261183	66924	5545	5760
Excluded: outside common support	932 (0.4)	38 (0.1)	745 (13.4)	1649 (28.6)
Included	260251 (99.6)	66924 (99.9)	4800 (86.6)	4111 (71.4)

6.8 Analysis of monitoring over calendar-years in the study

Fig. 6.9: Analysis over calendar-years in the study

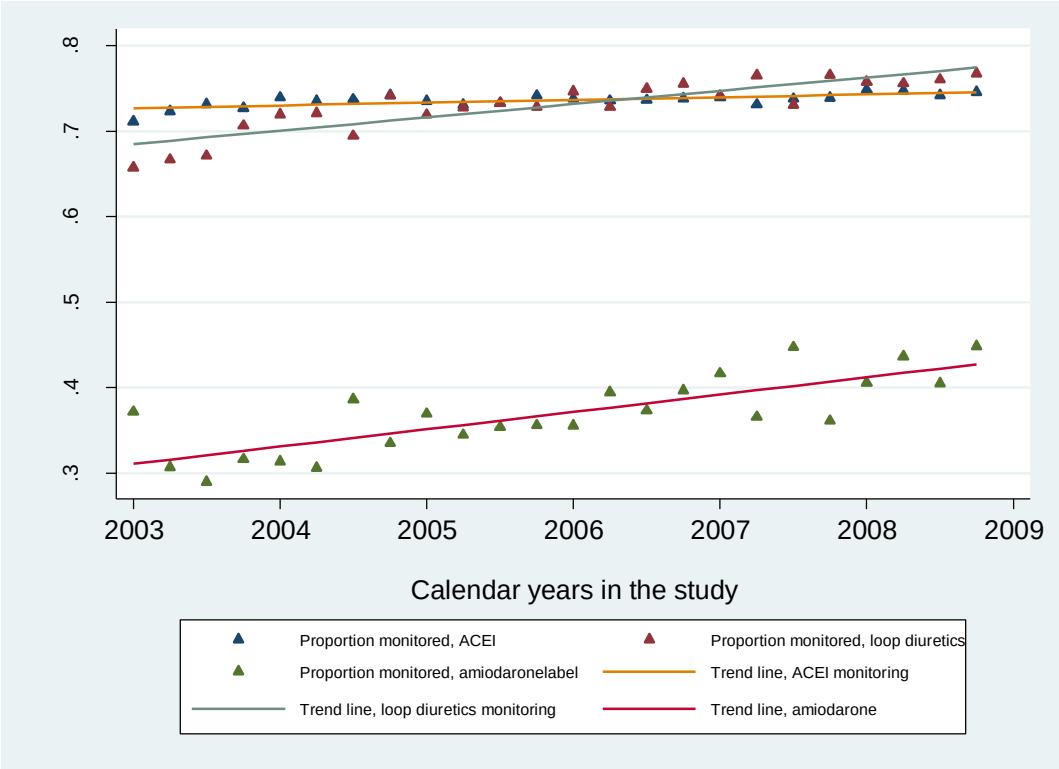


Figure 6.9 represents the completion of monitoring at recommended frequency or above, split by 3-month time intervals for patients initiated on the respective medication therapy from 2003 to 2009. It can be observed that the proportion of patients with appropriate frequency of monitoring of thyroid function in amiodarone therapy was significantly lower than in the case of the other two medication cohorts. In order to observe the trend across calendar years without short-term variation, a linear regression was fitted across the observations and a line of best fit plotted atop the scatter graph. There was a much more pronounced positive trend in the conduct of monitoring in the amiodarone cohort. Monitoring did not appear to increase in the ACEI cohort, whilst completion improved slightly in the loop diuretics cohort over the calendar years in the study.

6.9 Observed ADEs during the study period

Table 6.17: Summary of observed ADEs across medication-specific cohorts

Drug/Outcome	All groups	Less frequent	Recommended	More frequent
ACEI (N full/subsample)	261183/151145	42247/22647	108231/63962	110705/64536
HK (pr. care) N (%)	1826 (0.70)	87 (0.21)	597 (0.55)	1142 (1.03)
AKI (pr. care)	679 (0.26)	32 (0.08)	224 (0.21)	423 (0.38)
HK (admission)	48 (0.03)	1 (<0.01)	16 (0.03)	31 (0.05)
AKI (admission)	402 (0.27)	17 (0.08)	101 (0.16)	284 (0.44)
Loop diuretics (N full/subsample)	66962/41125	8107/4564	23581/14739	35364/21822
Hypon./hypok. (pr. care)	96 (0.14)	5 (0.06)	38 (0.16)	53 (0.15)
Hypon./hypok. (admission)	99 (0.24)	3 (0.07)	41 (0.28)	55 (0.25)
Fall-related injury (admission)	297 (0.72)	25 (0.55)	89 (0.60)	183 (0.84)
Amiodarone (N full/subsample)	5545/3581	2429/1475	2095/1573	1021/533
AIH (pr. care)	314 (5.66)	80 (3.29)	125 (5.97)	109 (10.68)
AIT (pr. care)	311 (5.61)	107 (4.41)	124 (5.92)	80 (7.84)
MACE (admission)	148 (4.13)	34 (2.31)	61 (3.88)	53 (9.94)

HK = hyperkalaemia; AKI = acute kidney injury; hypon. = hyponatraemia; hypok. = hypokalaemia; AIH = hypothyroidism
AIT = thyrotoxicity; MACE = major adverse cardiovascular event

6.10 Additional analysis 1: an investigation of outcomes in the "no monitoring group" in the ACEI cohort

Hyperkalaemia and AKI identified from CPRD records is normally diagnosed from laboratory tests, which may preclude patients with no record of monitoring during the follow-up period from having these ADE outcomes. Table 6.18 presents numbers of cases of hyperkalaemia and AKI identified in each comparison group in the ACEI cohort:

Table 6.18: Number of ADEs in each comparison group, ACEI cohort

ADE	No monitoring	Less frequent	Recommended	More frequent
N	45020	42247	108231	110705
HK (pr care)	118 (0.26)	87 (0.21)	597 (0.55)	1142 (1.03)
AKI (pr care)	53 (0.12)	32 (0.08)	224 (0.21)	423 (0.38)

There were 118 cases of hyperkalaemia and 53 cases of AKI in the cohort of patients with no monitoring. The proportion of patients with ADE was comparable to the one found in the group with sparse monitoring, and lower when compared against those monitored regularly. All cases of ADE in this group were identified

from CPRD diagnosis file, as they had no recorded monitoring tests. It is unknown how these diagnoses were established, whether their potassium and creatinine values were measured in secondary care or the laboratory tests were not recorded after being carried out. Given that 66.5% of cases of hyperkalaemia and AKI in the whole ACEI cohort were identified through a review of CPRD laboratory test results, it is likely that the actual number of ADEs was considerably higher. This is a potential source of bias as it is unknown how many cases of ADE were missed due to inability to measure serum potassium and creatinine.

A sensitivity analysis was conducted to examine the impact of inclusion of patients with no monitoring in the time-to-event analyses of hyperkalaemia and AKI in the ACEI cohort. An ordered probit model was re-fitted to accommodate the additional exposure category and was used to compute new propensity score weights using MMWS. The weights were applied to a Weibull model which examined each comparator (no monitoring, less frequent monitoring, more frequent monitoring) against a baseline group monitored at recommended frequency. The results are presented in Table 6.19. The analysis demonstrated that a strategy of no regular or sparse monitoring was associated with a lower rate of both hyperkalaemia and AKI, compared to recommended frequency monitoring. Patients monitored more frequently had a higher rate of both hyperkalaemia and AKI. The implications of this sensitivity analysis in the context of this project are discussed in Section 6.15.2.

Table 6.19: Results of sensitivity analysis of inclusion of no monitoring group, compared to original findings

Frequency of monitoring	HR (95% C.I.) in original analysis	HR (95% C.I.) in sensitivity analysis
Hyperkalaemia		
Once every 12 months	1	1
No monitoring	-	0.347 (0.277,0.435)***
Less than every 12 months	0.568 (0.418,0.773)***	0.513 (0.406,0.648)***
More than every 12 months	1.003 (0.905,1.112)	1.387 (1.252,1.537)***
AKI		
Once every 12 months	1	1
No monitoring	-	0.567 (0.408,0.788)**
Less than every 12 months	0.628 (0.377,1.047)	0.494 (0.335,0.729)***
More than every 12 months	0.896 (0.757,1.059)	1.436 (1.216,1.695)***
Original HK model Wald test: $\chi^2(2) = 13.62, p < 0.01$		
Original AKI model Wald test: $\chi^2(2) = 4.10, p = 0.129$		
Sensitivity HK model Wald test: $\chi^2(2) = 218.04, p < 0.001$		
Sensitivity AKI model Wald test: $\chi^2(2) = 64.28, p < 0.001$		
* p<0.05 ** p<0.01 *** p<0.001		

6.11 Additional analysis 2: an investigation of ADEs potentially driven by frequent monitoring in the ACEI cohort

This analysis examined the incidence of hyperkalaemia and AKI across patient groups with different frequency of monitoring in order to determine whether the outcomes may have been driven by frequent monitoring alone, rather than the underlying risk posed by the drug. The results are presented in terms of the raw numbers of patients presenting with ADEs, unadjusted risk ratios and risk ratios adjusted using MMWS weights estimated in Section 6.2.2. The results are presented in Table 6.20:

Table 6.20: Unadjusted and adjusted relative risk of hyperkalaemia and AKI, by monitoring frequency

Monitoring category	N	%	Hyperkalaemia cases (% of N)	AKI cases (% of N)	unadjusted RR (95% C.I.) for HK	unadjusted RR (95% C.I.) for AKI	adjusted RR (95% C.I.) for HK	adjusted RR (95% C.I.) for AKI
Recommended	108231	41.4	597 (0.55)	224 (0.21)	2.68 (2.14-3.35)	2.73 (1.89-3.96)	2.00 (1.67-2.38)	1.79 (1.37-2.36)
Less frequent	42247	16.2	87 (0.21)	32 (0.08)				
More frequent	110705	42.4	1142 (1.03)	423 (0.38)	1.87 (1.69-2.06)	1.85 (1.57-2.17)	1.36 (1.26-1.48)	1.21 (1.06-1.37)

Table 6.20 demonstrates that both the unadjusted and adjusted risk of hyperkalaemia and AKI was positively related to monitoring frequency. This is not surprising as more intensive monitoring groups contain a larger proportion of patients with old age, presence of comorbidities, multiple co-prescribed medications, as well as conditions such as CKD and diabetes which predispose them to a higher risk of ADE. The interpretation of this result is included in Section 6.15.2.

6.12 Additional analysis 3: an investigation of potential change in assigned monitoring regimen during the exposure period

An observational study of monitoring using retrospective data may not assume that patients assigned to a monitoring regimen will remain in the same regimen for the remainder of the follow-up period. Patients who have not experience an ADE in the initial months are perceived to be at lower risk of experiencing an ADE in the future and may be reclassified to less frequent monitoring by their GP. This analysis investigated the potential impact of this "reassurance" effect on the results of the time-to-event analysis in the ACEI cohort.

The analysis examined patients in the ACEI cohort who were monitored in accordance with guidelines in the first monitoring cycle and remained in the dataset for a minimum of one further cycle. Just 33621 out of 261183 patients were classified in this way (12.9% of total) (Table 6.21). There was a large increase in the proportion of patients monitored below recommended frequency (17.2% in the less frequent group and 49.9% with no further recorded tests) compared to the whole dataset, which is consistent with the "reassurance effect" hypothesis. Despite this, around one third of the cohort received further monitoring in subsequent cycles. Table 6.22 compares the characteristics of patients who were monitored in the first period and carried on to be on the drug for at least one further period to the whole cohort. The patients in this subgroup were younger, had fewer comorbidities and co-prescriptions, and less likely to have diabetes. There was a larger proportion of patients with CKD stage 3 or

higher. In addition, they were less likely to be censored from the study due to discontinuation, death or ADE, and more likely to complete the full period of follow-up. This group was less likely to be closely monitored due to lower prevalence of risk factors for hyperkalaemia and AKI.

Table 6.21: Breakdown of patients by monitoring frequency

Monitoring category	N	%
No monitoring	16789	49.9
Recommended	7846	23.3
Less frequent	5778	17.2
More frequent	3206	9.5
Total	33621	

Table 6.22: Baseline patient characteristics

Covariate	All patients	Monitored in first period
N	261183	33619
Mean age (SD)	63.6	62.2 (13.3)
Female %	46.1	45
Mean comorbidity score (SD)	0.83 (1.01)	0.74 (0.91)
Mean no. of coprescriptions (SD)	4.84 (2.54)	4.72 (2.57)
Diabetes %	12.2	10.3
CKD stage 3 or above %	4.2	5.5
Patients completing follow-up	62.3	72.6

6.13 Additional analysis 4: separating the time periods used to identify monitoring exposure and ADE outcomes using CPRD data

An additional analysis was conducted in order to examine an alternative method of identifying exposure and outcomes in the time-to-event analysis. Instead of identifying ADEs within the same time period as exposure, an additional 12 month period was added at the end of the follow-up period for each patient during which presence of ADEs was ascertained. Table 6.23 summarises the characteristics of the patients included in this analysis, compared to the full cohort:

Table 6.23: Descriptive statistics of patients who completed follow-up, compared to full cohort

Covariate	Full cohort	Patients completed follow-up
N	261183	149019
Mean age (SD)	63.6 (13.2)	62.5 (12.7)
% female	46.1	44.3
Mean no. of comobidities (SD)	0.83 (1.01)	0.79 (0.98)
Mean no. of co-prescriptions	4.84 (2.54)	4.84 (2.51)
% with CKD stage 3 or above	4.2	3.8
% with diabetes	12.2	12.3

The included cohort was different in terms of baseline characteristics compared to the full cohort. The included patients tended to be younger, healthier (based on mean number of comorbidities) and a lower proportion had CKD stage 3 or above. These differences are explained by the exclusion of patients for whom outcomes could not be ascertained due to early discontinuation of therapy, death or ADE.

The results of the time-to-event analysis with MMWS using propensity scores to control for confounding are listed in Table 6.24.

Table 6.24: Results of time-to-event analysis of ACEI monitoring, alternative exposure method

Frequency of monitoring	HR (95% C.I.) of hyperkalaemia	HR (95% C.I.) of AIK
Once every 12 months	1	1
Less than every 12 months	0.871 (0.663, 1.142)	0.839 (0.539, 1.307)
More than every 12 months	1.631 (1.352, 1.968)***	1.716 (1.274, 2.311)***
Weibull parameters		
Scale parameter (ln(λ))	-11.514 (-12.008, -11.020)***	-13.578 (-14.540,-12.6141)***
Shape parameter (γ)	1.002 (0.926,1.084)	1.193 (1.047-1.369)***
HK model Wald test:	$X^2(2) = 41.83, p < 0.001$	
AIK model Wald test:	$X^2(2) = 20.89, p < 0.001$	
*** p<0.05 ** p<0.01 *** p<0.001**		

The alternative method produced different results compared to the original analysis. The main difference is the apparent positive effect in the frequent monitoring group for the hazard rate of both hyperkalaemia and AKI, which wasn't observed in the main analysis. However, given the differences in the

baseline characteristics of patients in this analysis and the main analysis, the two sets of results are not directly comparable.

There were several issues with the design of the alternative method. Firstly, identifying outcomes at the end of the follow-up period required patients to still be in the dataset, e.g. it excluded all patients who have been censored or discontinued from therapy. This was a major threat to external validity as the resulting dataset was no longer representative of the population of patients on long-term ACEI therapy. The remaining patients were those in whom the monitoring strategy has been successful, or those who weren't monitored due to a low risk of ADE. The excluded patients potentially included cases where monitoring did not prevent ADE, or inadequate monitoring contributing to ADE, which were of direct interest to this study. A solution to this issue was to keep patients who have exited the dataset. However, this would have produced flawed results as a) no drug-ADE relationship can be demonstrated in patients for whom therapy had been discontinued; b) it is unclear how monitoring may affect the likelihood of a second ADE in patients who were censored out of the dataset due to a first ADE. This becomes an investigation of the effect of reactive, rather than proactive monitoring, on the risk of subsequent ADEs, which is beyond the scope of this work; and c) the outcomes in patients who died during the 4-year follow-up period would remain unknown.

6.14 Additional analysis 5: performance of alternative methods of controlling for non-random assignment to monitoring groups

Section 5.5 outlined several options for controlling confounding arising from non-random treatment assignment in observational studies. This analysis examined the impact of alternative methods to control confounding on the results of the time-to-event analysis of monitoring in the ACEI dataset. The method of choice in the analysis (MMWS) was compared to IPTW and Weibull regression which included all patient covariates. The latter method employed backward covariate selection. Starting from the model with all the original covariates listed in Section 6.2.2, non-significant ($p<0.05$) covariates were iteratively removed, starting with the covariate with the highest p value. The results for each alternative method are listed in Table 6.25. The performance of each method was additionally assessed in terms of covariate balance by examining the mean values of key parameters across comparison groups in Table 6.26 and Figures 6.10 to 6.12.

Table 6.25: Results of time-to-event analysis of ACEI monitoring, by adjustment method

Comparison	HR (95% CI), MMWS weights	HR (95% CI), IPTW weights	HR (95% CI), regression
ACEI and HK (primary care)			
Less frequent vs. recommended	0.568 (0.418-0.773)***	0.440 (0.334-0.580)***	1.623 (1.289-2.043)***
More frequent vs. recommended	1.003 (0.905-1.112)	0.866 (0.778-0.964)**	0.791 (0.711-0.879)***
ACEI and AKI (primary care)			
Less frequent vs. recommended	0.628 (0.377-1.047)	0.431 (0.279-0.666)***	1.638 (1.126-2.383)***
More frequent vs. recommended	0.896 (0.757-1.059)	0.729 (0.612-0.869)***	0.693 (0.583-0.825)***
ACEI and HK admission			
Less frequent vs. recommended	0.127 (0.017-0.951)*	0.145 (0.019-1.093)	0.452 (0.059-3.463)
More frequent vs. recommended	0.921 (0.492-1.723)	0.811 (0.420-1.566)	0.902 (0.477-1.705)
ACEI and AKI admission			
Less frequent vs. recommended	1.170 (0.609-2.249)	0.747 (0.427-1.310)	1.120 (0.666-1.882)
More frequent vs. recommended	1.560 (1.225-1.987)***	1.391 (1.086-1.783)**	1.328 (1.042-1.694)*

The aim of propensity score weighting as a method to control confounding in observational data is to balance the distributions of the baseline covariates across comparison groups in the analysis in order to render them more comparable. The performance of alternative methods for propensity score weighting described in Section 5.5 was assessed by comparing patient characteristics in the three monitoring groups. Pseudo-datasets were generated to reflect weighting using MMWS and IPTW by creating duplicate records for patients according to their weight. The original weights were multiplied by 10 to take into account the large number of weights with a value smaller than 1, particularly in MMWS weights. The distributions of patients with selected characteristics were compared to the original unadjusted cohort. The following characteristics were used to check covariate balance: mean age, proportion of female patients, patients with multimorbidity (≥1 comorbidities), ≥4 comorbidities, ≥4 co-prescriptions, ≥7 co-prescriptions, CKD and diabetes.

Table 6.26: Mean values of covariates across comparison groups, split by adjustment method

Variable	Original			MMWS			IPTW		
	Less frequent	Recom-mended	More frequent	Less frequent	Recom-mended	More frequent	Less frequent	Recom-mended	More frequent
Mean age (SD)	59.1 (13.4)	62.5 (13.1)	66.3 (12.6)	63.7 (13.5)	63.1 (13.1)	63.8 (13.1)	63.3 (13.5)	63.5 (13.2)	63.9 (13.1)
% Female	44.3	45.6	47.2	45.6	46	46	45.5	46.1	46.1
≥1 comorbidities	14.4	17	24	22	18	19.9	21.1	19	20
≥4 comorbidities	0.3	0.4	1.2	0.8	0.5	0.6	0.8	0.6	0.6
≥4 coprescriptions	29	33.7	40.4	36.7	35	36.6	36.2	35.1	37
≥7 coprescriptions	9	11.4	23.1	22	12.7	16.6	20.3	14.4	16.6
CKD	2.3	3.8	6.7	5.4	4.2	4.8	4.9	4.7	4.8
Diabetes	4.2	8.7	18.7	14.5	10	12.2	13.3	12	12.2

Fig. 6.10: Mean age across comparison groups, split by adjustment method

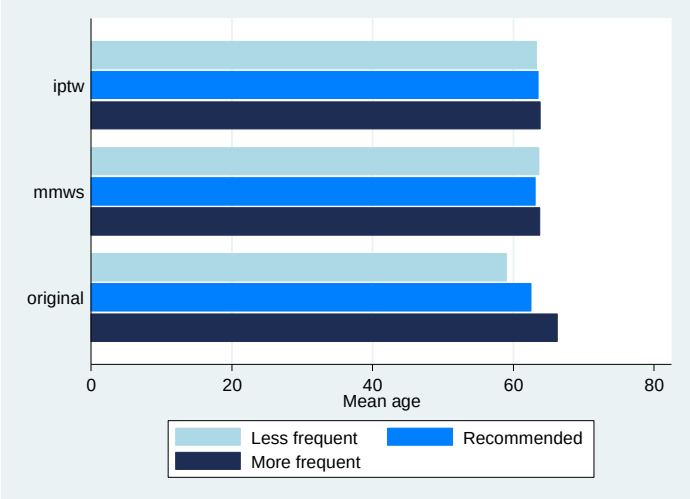


Fig. 6.11: Proportion of patients with ≥1 comorbid condition, split by adjustment method

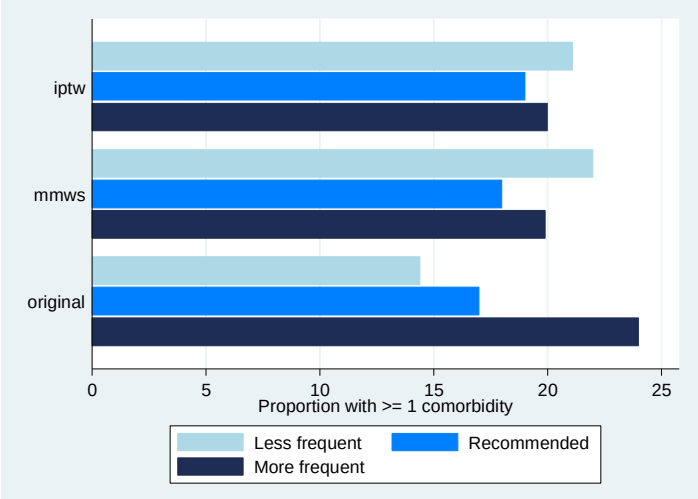
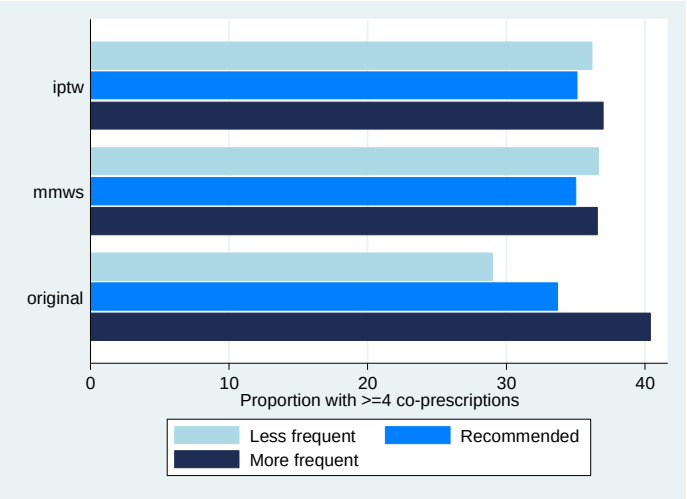


Fig. 6.12: Proportion of patients with ≥4 co-prescriptions, split by adjustment method



It is evident from Table 6.26 that both MMWS and IPTW performed well at balancing the comparison groups in terms of their baseline characteristics. For instance, the difference in mean age between the oldest and youngest group was reduced from 7.2 years in the original data to 0.7 using MMWS and 0.6 using IPTW. The comparison groups were more balanced in terms of the proportion of patients with key predictors of monitoring and ADE, such as multimorbidity and polypharmacy, as demonstrated in Table 6.26 and Figures 6.10 to 6.12.

There were differences in the results of the time-to-event analysis across the compared methods (statistically significant at $p<0.05$). The most notable difference was in the effect of monitoring frequency on hyperkalaemia and AKI. The original analysis which used MMWS found that monitoring less frequently than recommended was associated with a reduction in the rate of hyperkalaemia (HR 0.57) compared to the recommended monitoring group. IPTW method generated the same result for hyperkalaemia, and additionally found a reduced rate of AKI in the less frequent group (HR 0.43). Another finding in the IPTW analysis found a reduced rate of both hyperkalaemia (HR 0.87) and AKI (HR 0.73) in the more frequent group. By contrast, standard regression method found that the risk of both hyperkalaemia (HR 1.62) and AKI (HR 1.64) was found to be higher in the less frequent group. There were fewer notable differences in results in hospital outcomes (admission for hyperkalaemia or AKI) between standard regression and propensity score weighting methods. The rate of hospital admission for hyperkalaemia was higher in the less frequent group under standard regression, which caused the p value cross the 5% threshold for statistical significance.

6.15 Discussion

6.15.1 Discussion of results of time-to-event analysis

This chapter outlined the results of a quantitative analysis of regular monitoring of ACEI, loop diuretics and amiodarone and the effect of different frequencies of monitoring long-term therapy on the risk of ADEs linked to these medications. There were several notable similarities between the separate medication-specific cohorts. Multimorbidity and polypharmacy were common across all three cohorts, with the proportion of patients classed as multimorbid (2 or more co-morbid chronic conditions) varying between one in five and one in every two patients in the cohort. A majority of patients had multiple co-prescribed medications, with sizeable numbers of patients experiencing polypharmacy. There were also differences: patients in the ACEI cohort tended to be younger and healthier (in terms of the mean comorbidity score) and had fewer co-prescriptions than patients in the loop diuretic and amiodarone cohorts. In addition, female patients were a minority in the ACEI cohort, whilst in the other two cohorts they constituted a majority of the sample. A large number of patients were co-prescribed both ACEI and loop diuretics, as well as other medications which are likely to lead to both more intensive monitoring and influence electrolyte levels, such as ARBs and thiazide diuretics, suggesting that combined therapy was commonly prevalent amongst patients included in the analysis.

The descriptive analysis of monitoring frequency across the three medication-specific cohorts found that a large proportion of patients weren't monitored in accordance with the frequencies set within published guidelines for regular monitoring. Moreover, the approach to monitoring was quite evenly split across alternative frequencies. For instance, in the full sample ACEI analysis, 35% of patients had records of U&E testing at the recommended frequency of once in 12 months. Roughly the same proportion of patients was monitored more frequently, whilst 14% were monitored at a lower frequency (excluding patients who had no recorded tests at all). The proportions were roughly similar in the loop diuretics cohort. In the case of amiodarone, compliance to recommended monitoring was much poorer: 29% of the cohort received some degree of monitoring, but at a frequency lower than that which is listed in guidelines, whilst 34% had no record of a test at all. As a result, patients who had TFT tests once in 6 months or more frequently constituted a minority of all patients on long-term amiodarone therapy. These findings are consistent with previous studies of monitoring which were reviewed in Chapter 3 of this thesis. Compliance to recommendations for follow-up monitoring was found to be poor in several observational and

randomised studies (Raebel, McClure et al. 2007, McDowell, Coleman et al. 2010, Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015). In addition, the relatively poor conduct of follow-up monitoring of thyroid function in amiodarone therapy was also observed in previous studies (Tjia, Fischer et al. 2011, Avery, Rodgers et al. 2012, Akbarov, Kontopantelis et al. 2015). The key difference between the study described in this chapter and previous evidence on monitoring is that the current study observed frequency of monitoring over a long period of observation (4 years for ACEI and loop diuretics and 3.75 years for amiodarone), whilst all previous studies identified in the review in Chapter 3 looked at monitoring as a binary outcome over a shorter time interval (typically 12 months).

The discussion in Chapter 5 made a strong case for developing and applying methods to control for confounding arising from patient factors which influence a GP's decision to conduct regular monitoring at a particular frequency. An ordered probit model was chosen in order to gauge the effect of patient characteristics on the monitoring frequency to which they have been assigned. Several statistically significant effects were observed: patients who were older, have more comorbidities, co-prescribed medications from different BNF chapters, as well as those diagnosed with CKD and diabetes at baseline were more likely to be assigned to a more intensive monitoring group. These effects were consistently observed across all cohorts, meaning that they were not unique to a single medication, but rather were more likely to be a reflection of the overall approach to monitoring in primary care. This suggests that GPs engaged in risk-profiling of patients to a certain degree when deciding how often to conduct regular monitoring of medications in individual cases. These findings are similar to ones found in previous studies of monitoring: Simon et al also found that monitoring of loop diuretics was more likely to be conducted in male patients, whilst amiodarone monitoring was more likely to be conducted in female patients (Simon, Andrade et al. 2005). Other studies have found that female patients were less likely to receive regular U&E monitoring in ACEI therapy, which contrasts with this study (Raebel, McClure et al. 2007). The effect of age, number of co-prescriptions and comorbidities on monitoring was unsurprising as it has been previously observed, as was the positive effect on monitoring conduct of presence of CKD, diabetes and co-prescribed diuretics (Simon, Andrade et al. 2005, Raebel, McClure et al. 2007, Akbarov, Kontopantelis et al. 2015). Another possible explanation for these observations is that patients who have a larger number of concomitant conditions and treatments tend to have more contacts with primary care services than average, leading to more frequent laboratory

investigations. Unfortunately, it was not possible to establish the indication for individual laboratory tests in order to distinguish which tests were carried out for the purpose of monitoring of the study medications (which is within the remit of this study) and those investigations which were carried out for other reasons. It is a feature of EHRs from large databases and remains a limitation of this study.

Although patients with conditions and characteristics which have previously been found to be associated with risk of ADE were more likely to be monitored, the analysis found that many patients with an apparently elevated predisposition towards an ADE were monitored inadequately (according to published guidelines). For instance, 2022 patients with CKD stage 3 or higher who were on long-term ACEI therapy (16.5% of all patients in this condition) and 4344 (12.6%) were monitored less frequently than recommended, despite the fact that NICE guidelines for treatment of ACEI specify the need for more intensive monitoring in these at-risk patient subgroups. A quarter of patients co-prescribed a loop diuretic during long-term ACEI therapy failed to receive monitoring at recommended frequency, although combined therapy is likely to increase the risk of serum potassium imbalance. In the amiodarone cohort, poor compliance to monitoring guidelines persisted in patient groups which are likely to be frail and ill: the majority of long-term users of amiodarone aged 80 years or older, those with 5 comorbidities or more or those with 7 co-prescribed medications or more had TFT tests carried out less frequently than once in 6 months.

Substantial variation in the conduct of regular monitoring was observed across general practices. In the case of U&E monitoring of ACEI and loop diuretics, a negative skew was observed, which was caused by the presence of practices with low compliance to monitoring guidelines. The distribution was different in the case of thyroid function monitoring of amiodarone, with a more even spread of practices according to proportion of patients with completed monitoring. This suggests that there was less agreement regarding the need for regular monitoring of amiodarone compared to ACEI and diuretics in primary care practice in England. When looking at monitoring practice across calendar-years in the study, there was an improvement in the proportion of patients who were monitored in accordance with guidelines in the loop diuretic and amiodarone cohort, but not in the ACEI cohort (Fig. 6.9). It was not possible to rule out that the increase in observed laboratory tests was due to improvements in recording of data, rather than a shift in the approach towards monitoring by GPs, which has been described previously within large database research (Wu, Jen et al. 2010). Isolating these two factors would require a separate investigation of data

recorded in CPRD, which was outside the scope of this study. However, the trends of improvement in recorded tests were different for each medication, which may suggest that at least part of the rise can be attributed to a change in practice over time.

The performance of MMWS as a tool to adjust for confounding was examined using the mean values of patient covariates in each comparison group in the weighted dataset in Section 6.14. It was found that MMWS adequately balanced the values of covariates when compared to the original unadjusted data and performed similarly to IPTW.

The time-to-event analyses linking monitoring frequency to the hazard rate of ADEs linked to the study medications found statistically significant effects in some of the outcomes assessed. Monitoring urea & electrolytes less frequently than recommended was associated with a reduced rate of hyperkalaemia linked to ACEI treated both in primary & secondary care and diuretic-induced hypokalaemia/hyponatraemia in primary care. Monitoring more frequently than recommended was associated with an increase in the rate of hospital admission for AKI linked to ACEI, but a reduction in the rate of hypokalaemia/hyponatraemia linked to loop diuretics treated in both primary and secondary care. Monitoring thyroid function during amiodarone therapy below the recommended frequency of once every 6 months was linked to an increase in the hazard rate of hypothyroidism, thyrotoxicity and MACE compared to a strategy involving monitoring in line with published guidelines. Monitoring TFT above recommended frequency was associated with an increase in the rate of AIH diagnosed in primary care.

It was difficult to draw comparisons to previous evidence as very few previously published studies have linked follow-up monitoring to ADEs, and those few studies which were identified in the literature review in Chapter 3 different in terms of design and population. A US-based study which employed a retrospective cohort design to estimate the effect of regular monitoring of U&E on the risk of ACEI-induced hyperkalaemia in patients with diabetes found that regular monitoring reduced the risk of ADE (RR 0.5, C.I. 0.37,0.66) (Raebel, Ross et al. 2010). The effect of monitoring was larger in the subgroup with CKD (RR 0.29), but not in patients without CKD and <50 years of age. The results of the current study found the opposite effect: less frequent monitoring was associated with an apparent reduction in the risk of hyperkalaemia and AKI. A UK-based study which used CPRD data to investigate regular monitoring of long-term ACEI

and diuretic therapy found that recommended monitoring of U&Es was associated with an increased risk of hospital admission (HR 1.37, 1.21,1.55). This was a similar finding to the time-to-event analysis of ACEI monitoring, which found an increase in the hazard rate of hospital admission for AKI amongst patients monitored above recommended frequency. The authors supposed that the increased risk reported in their study was due to an increased propensity to conduct investigations in patients who were admitted to hospital, suggesting that the direction of causality was from hospital admission to monitoring, rather than the opposite. The current study applied propensity score adjustment to minimise the risk of systematic differences in the propensity of hospital admission across patient groups, although it was not possible to rule out that increased contact with health services may have led to increased monitoring. Based on the literature review in Chapter 3, this was the first study which investigated the effect of frequency of thyroid function monitoring on the risk of amiodarone-induced ADEs.

6.15.2 Discussion of additional analyses to test methodological assumptions

Sections 4.4 and 4.11 highlighted a number of methodological issues linked to the design of the observational study. In particular, the exclusion of patients with no recorded monitoring and simultaneous identification of monitoring exposure and ADE outcomes introduced a risk of underestimating the protective effect of monitoring. Additional exploratory and sensitivity analyses were conducted in order to examine these issues and their potential impact on the internal validity of the study.

Inclusion of patients with no monitoring

The analysis of ADE cases in the group of patients with no recorded monitoring demonstrated that the proportion of patients with ADE in this group was similar to the group with sparse monitoring. However, given that two thirds of cases of hyperkalaemia and AKI were identified from abnormal laboratory test results, this figure may not be representative of the true rate of ADEs in this group. Laboratory monitoring and test results collected in this study were limited to primary care. It is likely that many of the included patients were monitored in hospital, particularly during inpatient or outpatient episodes. It is possible that patients who appear to be unmonitored based on primary care data were actually placed on a monitoring regimen in secondary care. This is another potential

source of bias, as inclusion of patients monitored in secondary care in the unmonitored group is likely to underestimate the effect of monitoring.

The above case is supported by the observed baseline characteristics of the "no monitoring" group. In the ACEI cohort, frequency of monitoring was correlated with certain patient characteristics: those in receipt of more frequent testing tend to be older, with more comorbidities, co-prescribed medications and other pre-existing risk factors. The group with no monitoring went against this trend: they tended to be older (61.1 vs. 59.9), had more comorbid conditions (0.76 vs. 0.66), were more likely to have pre-existing CKD (3% vs. 2.3%) and diabetes (5.8% vs. 4.2%) compared to the group with low frequency monitoring. Certain patient subgroups which are likely to include the frail and elderly were heavily represented in the no monitoring group: 14% of those aged 80 and over compared, 12.7% with ≥5 comorbidities, 14.7% with severe CKD. These figures contradict the trend that older, sicker patients are monitored more closely which is expected based on the findings of previous studies and the observed data from the rest of the cohort. However, the frail condition of these patients means that it was likely that they were monitored in secondary, rather than primary care, and the presence of these patients in the "no monitoring" group is explained by missing data on tests conducted in the hospital setting.

The likely explanation for these observations is that the group of patients with no observed monitoring was not homogeneous: it included patients who were perceived to be low risk (younger and healthier), those with a higher risk profile for whom drug therapy was discontinued after one monitoring cycle (which limited the opportunity to monitor) and frail elderly patients who were monitored in the secondary care setting. Due to the nature of the data source (no information on the indication for monitoring, no access to case notes), there was no reliable method of separating these groups of patients. A decision was made not to include patients with no observed monitoring during follow-up due to the risk of introducing bias into the time-to-event analyses.

A sensitivity analysis was conducted in order to gauge the effect on the time-to-event analysis in the ACEI cohort of inclusion of the no monitoring group. The new analysis with four exposures (no monitoring, less frequent, more frequent compared against a baseline group with recommended frequency), compared to the original analysis with 3 exposures produced different results in a number of categories. Absence of monitoring was associated with a statistically significant reduction in the rate of both hyperkalaemia (HR 0.35) and AKI (HR 0.57). The

major difference in the original results was the apparent increased risk of hyperkalaemia (HR 1.39) and AKI (HR 1.44) in the group with more frequent monitoring (both $p<0.001$), whilst the original analysis found no effect in the more frequent group. These results may have been skewed by misclassification of exposure in the "no monitoring group" due to missing hospital monitoring records. The results of the time-to-event analysis were driven by the propensity score weights which re-balanced the cohort in favour of patients with a smaller propensity of being assigned to a particular monitoring regimen. Patients who were monitored outside primary care were classed as not monitored in the analysis, which produced biased propensity scores and weights, potentially underestimating the hazard rate of hyperkalaemia and AKI in this group.

Abnormal laboratory test results driven by intensive monitoring

The exploratory analysis in Section 6.12 found that the rate of both hyperkalaemia and AKI increased monotonically with frequency of monitoring. The unadjusted RR for ADEs in monitored vs. monitor error group listed in Table 6.20 is consistent with the findings from previous studies: 1.36 (95% C.I. 1.21-1.52) in McDowell et al 2012 and 1.38 (0.9-2.14) in Raebel et al 2010. Given that Raebel et al study was conducted in a diabetic population, it was likely to underestimate the difference in risk between monitoring groups as the differences in baseline patient characteristics across comparison groups were likely to be smaller in a diabetic population, compared to the general population of patients on ACEI therapy.

It must be noted that relative risk figures based on the total number of ADE cases may not reflect the actual likelihood of occurrence of ADEs during long-term therapy under different monitoring regimens. Patients who are at lower risk of ADE are likely to have a longer follow-up time due to a lower rate of discontinuation or censoring event. Therefore, adjusted RR figures in Table 6.20 overestimate the true risk of ADE. The time-to-event analysis outlined in Section 5.6 was designed to take into account exposure time by estimating hazard ratios of ADE across comparison groups, which were a more accurate reflection of the true difference in the risk of ADE across comparison groups.

The differences in the risk of ADE were a reflection of the differences in baseline characteristics across comparison groups. Patients assigned to a more frequent monitoring regimen were at higher risk of ADE, and there is no indication that increased frequency of monitoring resulted in a higher rate of identified ADEs.

However, it was not possible to completely discount outcomes which were driven to some degree by intensity of laboratory testing. Frequent testing may produce abnormal test results driven by natural variation in serum potassium levels across time. Patients with a history of HK or AKI are more likely to access healthcare services, prompting regular testing and increasing the propensity of repeat diagnosis. Due to the absence of information on the indication of test or case notes for patients with ADE it was difficult to establish the direction of the causal effect in these cases.

Reduction in frequency of monitoring during follow-up

Based on the figures presented in Tables 6.21 and 6.22, there was an apparent "reassurance" effect in the monitoring exposure data. Patients who have progressed through one cycle of therapy without an ADE were less likely to be maintained on regular monitoring. This issue must be considered in the context of GP behaviour during therapy. Upon initiation, there is uncertainty regarding a patient's risk of ADE and the decision to monitor is based on the presence of known risk factors. After observing the patient for 1 or 2 years, the GP has collected enough data to make a more informed decision. In the absence of ADE, the case for regular monitoring is weaker and the strategy is changed to reflect this. This change in the approach to regular monitoring part way through follow-up potentially introduces a risk of bias into the analysis due to the fact that exposure is determined not only by the baseline characteristics (which may be controlled by propensity score weighting), but additionally by the absence of ADE in the first monitoring cycle, which is harder to control using statistical methods. However, the relatively small proportion of the dataset to which this applies was a factor which limited the potential of this issue to influence the results of the time-to-event analysis. Nonetheless, this effect was likely to be present and was noted as a limitation of the analysis design.

Alternative method of identifying exposure and outcomes

Retrospective analyses of monitoring may be limited by certain methodological difficulties associated with identifying exposure and outcomes in the same time period. Diagnoses of ADE may be driven by frequent testing, rather than the underlying reaction to the drug. Patients who do not present with an ADE may be reclassified to less intensive monitoring once GPs revise down the perceived risk of future ADEs. Upon investigation of the original data described in Section 6.12,

it was apparent that these issues were unlikely to be a determining factor in the analysis for a number of reasons, but it was not possible to discount the potential bias that may have affected the time-to-event analysis. An additional analysis was conducted in order to examine the merits of an alternative method which separated the time periods used to classify monitoring exposure and ADE outcomes. The alternative method produced different results to those from the original analysis, although the comparability of the two sets of results was questioned due to differences in the distributions of baseline characteristics in the two datasets. In particular, the alternative method required the exclusion of a large proportion of patients from the original dataset due to early withdrawal from the drug and censoring, which compromised its external validity. This resulted in the exclusion of patients who were at higher risk of ADE who benefit most from monitoring. The resulting dataset was therefore insufficiently representative of the study population to answer the main research questions set out in Chapter 1 of this thesis.

Comparing different methods of adjusting for confounding from non-random treatment assignment

Section 6.14 tested alternative methods of controlling confounding arising from non-random assignment of patients into monitoring groups for analysis. It was found that standard regression produced conflicting results in two out of four of the outcomes included in the study. Given that the alternative analyses were conducted on the same dataset of long-term ACEI users and the same set of baseline covariates were used to adjust for confounding, the differences in the resulting hazard ratios are explained by the different ways in which these covariates were used to isolate the causal effect of the monitoring regimen assigned to individual patients. Regression analysis identifies the amount of variation in the outcome (in this case the hazard rate of ADE) which is explained by changes in exposure (frequency of monitoring) and other covariates (baseline patient characteristics). This approach is a suitable method of adjusting for the confounding effect of patient characteristics if the exposure and other covariates are statistically independent. In the case of follow-up monitoring, assignment to a particular frequency of monitoring is determined by baseline characteristics of the patient which are also predictors of the outcome. In addition, covariates may be affected by the assigned monitoring regimen and the perceived risk of ADE at baseline, such as co-prescription of diuretics or ARBs in patients on long-term ACEI therapy which may predispose the patient to the risk of ADE. Therefore, standard regression may produce biased results and a method which takes into

account the patient factors which determine the eventual monitoring group assignment is required. Balancing the comparison groups using the probability of treatment allocation estimated from baseline covariates (IPTW and MMWS) can achieve this.

Another advantage of using propensity score methods over conventional regression methods is the possibility of adjusting for selection bias at the design stage of the model. Similarly to an RCT, which incorporates a random assignment mechanism to ensure balance across comparison groups before proceeding to analysis, these propensity score methods allow an investigator to assemble a matched or weighted cohort and ensure that selection bias is accounted for before estimating the treatment effect. After carrying out the adjustment, comparability across treatment groups may be assessed and the design can be calibrated to balance the trade-off between sample size and bias reduction. This feature makes matching or weighting methods advantageous over direct adjustment of regression using the propensity score, which ignores this issue altogether and may result in an analysis of treatment effect of the intervention across patient groups which are too different in terms of their baseline characteristics to be comparable. If the overlap in the distributions of baseline covariates is insufficient, the regression model may produce biased results of treatment effect.

As outlined in Section 5.5.2, IPTW estimates weights as the reciprocal of the propensity score. The estimated weight gets infinitely large in case of a propensity score close to zero. This may skew the results when a large proportion of estimated scores are close to zero (Kurth, Walker et al. 2006). In addition, IPTW method is prone to errors arising from misspecification of the model used to estimate propensity scores as the weights are directly related to its output. MMWS is advantageous as it computes a weight which reflects the subject's representation with a stratum of the propensity score, rather than transposing the score directly to compute weights. The method of computing MMWS weights limits the effect of outliers in the propensity score distributions.

The results of the time-to-event analysis of amiodarone described in Section 6.6.3 support current guidelines for thyroid function monitoring during amiodarone therapy. In order to assess the economic effect of monitoring in this medication cohort, the parameters of the fitted Weibull regression model were applied in a model-based cost-effectiveness analysis. These include the hazard ratios of AIH, AIT and MACE as measures of the effect of monitoring on the

probability of occurrence of ADE and the shape and scale parameters which defined the progression of the hazard rate over time. This information was used to apply the initial transition probabilities in the model to investigate whether the reduction in the risk of ADE translates into reduction in cost and improvement in utility in patients. The methods and results of this analysis are the focus of Chapter 7.

Chapter 7

Methods and results of the cost-effectiveness analysis of amiodarone monitoring

7.1 Introduction to the chapter

The findings from the time-to-event analysis of amiodarone monitoring using EHRs from CPRD and HES suggest that monitoring at the recommended 6 month frequency during long-term therapy may be effective at preventing AIH and AIT. In the interest of examining the effect of these results on the cost of treatment of ADEs in these patients and the long-term outcomes in terms of patient utility, a cost-effectiveness modelling analysis of TFT monitoring of amiodarone therapy was carried out by combining information on transition through a hypothetical disease pathway with cost and utility effects of the different outcomes of amiodarone therapy.

This analysis builds on the time-to-event analysis of amiodarone monitoring outlined in Section 6.6.3. The methods and results of this analysis have been described in detail in Chapters 4, 5 and 6 of this thesis. This chapter describes the model which compared the costs of treatment of amiodarone ADEs from a UK NHS perspective in terms of 2014 price levels and outcomes in terms of QALYs accumulated over a 5 year study horizon. The analysis involved comparison of expected values of incremental cost and QALYs, with one-way and probabilistic sensitivity analyses used to assess the effect of uncertainty of the true values of model parameters. Finally, the chapter concludes by discussing the results and comparing with existing evidence in the literature.

7.2 Aims and objectives of the cost-effectiveness analysis

The aim of this analysis is to compare recommended frequency of regular thyroid function monitoring during amiodarone therapy compared to alternative strategies involving monitoring lower and higher than the recommended frequency in terms of incremental cost and utility gained by a hypothetical cohort of patients on regular amiodarone therapy in a typical treatment pathway.

The following objectives were set:

i. Describe the clinical pathway of thyroid ADEs linked to amiodarone (including hypothyroidism, thyrotoxicity and major adverse cardiovascular events) and their management in primary and secondary care by representing the progression of a hypothetical cohort of patients through a state transition model.
ii. Identify the resource use and cost associated with the management of thyroid ADEs linked to amiodarone.
iii. Value the impact of thyroid ADEs on patient health in terms of quality-adjusted life-years (QALYs).
iv. Assess the costs and outcomes in patients monitored at recommended frequency, compared to those who received monitoring at frequencies above or below this level.
v. Conduct an incremental economic analysis of cost per QALY gained over a 5 year study period .
vi. Identify key areas of uncertainty using a one-way sensitivity analysis and the overall effect of uncertainty in all model parameters using a probabilistic sensitivity analysis.

7.3 Methods

7.3.1 Study approval

The study protocol was reviewed by the Independent Scientific Advisory Committee (ISAC) for Medicines and Healthcare products Regulatory Agency in line with requirements for the extraction and dissemination of CPRD data. Linked data from the HES database was additionally requested. The protocol was approved in December 2015 (ISAC protocol no. 15_241). A copy of the approved protocol is available upon request from the author.

7.3.2 Model population

The source population which was used to estimate the probabilities of progression of ADEs in the preliminary time-to-event analysis included adult patients (≥18 years old) newly initiated on a course of amiodarone therapy between 2003 and 2010, with a record of regular amiodarone prescriptions 9 months after initiation. The full details of cohort selection and study design for the time-to-event analysis are included in Chapter 4.

7.3.3 Setting and perspective

The model simulated the costs and outcomes of alternative strategies of monitoring long-term amiodarone therapy in general practices in the UK. Outcomes were assessed in terms of ADEs treated both in primary and secondary care. Costs were estimated from the perspective of the NHS.

7.3.4 Model comparators.

The model compared alternative frequencies of monitoring of thyroid function (through TSH laboratory testing) in patients on stable therapy with amiodarone in primary care. The frequency recommended in guidelines (once in 6 months) was used as the baseline comparison group. This was compared against monitoring at a frequency below the recommended level (above zero but below once every 6 months) and monitoring above recommended frequency (above once every 6 months). The choice of comparators was a reflection of the design decisions for the time-to-event analysis of monitoring described in Section 5.3.

7.3.5 Model horizon and discounting

The model was run at the 5 year horizon. As the preliminary cohort study had a follow-up length of 2.5 years, cost and effectiveness for the subsequent 2.5 years involved extrapolation based on the parameters of the survival analysis computed from the cohort study. The NICE guidance for methodology used in health technology appraisal recommend using a lifetime horizon to accurately estimate the full economic effect of comparators (NICE 2004). However, the application of lifetime horizon to this study would have involved a substantial amount of extrapolation with parametric assumptions based on results inferred from 2.5 years of follow-up. Extrapolating survival analysis results beyond this observed period within a modelling analysis may compromise the internal validity of a cost-effectiveness analysis (Davies, Briggs et al. 2013). Two potential sources of error involve assumptions of proportionality of hazard function, which may not hold in reality as the relative effect of a strategy over its alternative may not necessarily remain constant over time following implementation; the second issue involves the behaviour of hazard functions over time: the parameters estimated in the time-to-event regression determined the shape and scale of the hazard function, which were assumed to be constant. It is unclear whether this may be applicable to an extrapolation to lifetime horizon as the rate at which instantaneous hazard of ADE changed in subsequent years of therapy is unknown. In the context of this

analysis, it is clear that both of these considerations may have affected the reliability of the analysis results: it is reasonable to assume that amiodarone therapy is used in chronic heart conditions and may persist over a period of many years; it is difficult to verify whether the relative effect of monitoring frequency on the risk of ADE stays constant over this long period due to insufficient data from long-term observational studies. Due to these considerations, a 5 year horizon was deemed more appropriate.

All future cost and QALY payoffs were discounted according to the HM Treasury's benchmark annual discounting rate of 3.5% (HM Treasury 2011), as the rate of time preference in relation to all costs and outcomes was assumed to be positive.

7.3.6 Measure of effectiveness

The effect of frequency of monitoring in this analysis was assessed via its impact on the probability of developing thyroid ADEs during amiodarone therapy and in the 12 months following its cessation. The specific ADEs included in the analysis were overt hypothyroidism (AIH), thyrotoxicity (AIT) and major adverse cardiovascular event (MACE). Hypothyroidism (abnormally low thyroid function) is diagnosed as TSH>10mU/L. Thyrotoxicity (overactive thyroid) is diagnosed as TSH<0.1 mU/L. MACE is a composite outcome and is made up of several cardiac events which require emergency hospital admission: heart failure, myocardial infarction, stroke or ventricular arrhythmia (Yiu, Jim et al. 2009). The presence of one of the above events was used to identify MACE in the study patients. A brief description of these ADEs are included in the table below, the full specification of ADE outcomes and justification for their inclusion in this study are listed in Chapter 4.

The transition probabilities for the progression to AIH, AIT or MACE were calculated using parameters of the Weibull cumulative hazard function estimated in the time-to-event analysis, the results of which were outlined in Section 6.6.3. Hazard ratios estimated in the individual Weibull regression models were used to represent the effect of monitoring frequency on the transition of patients to ADE states. Transition probabilities which reflect the treatment pathway and mortality at each stage of treatment are explored in more depth in Section 7.3.7 for each individual transition in the model.

Table 7.1: Definitions of ADEs used in this programme of research.

ADE outcome	Biochemical marker	Treatment	Cessation of therapy
Hypothyroidism (overt)	TSH >10mU/L	Thyroxine	No
Thyrotoxicity (overt)	TSH <0.1 mU/L	Treatment strategy depending on case	Yes
Major cardiovascular event (MACE)	N/A	Treatment strategy depending on case	Yes

Table 7.2: Hazard ratios from the time-to-event analysis.

Model	Comparison	HR	95% CI
AIH	Less frequent Vs recommended	2.78	1.42-5.47
	More frequent Vs. recommended	1.53	1.08-2.17
AIT	Less frequent Vs recommended	2.57	1.40-4.74
	More frequent Vs. recommended	1.23	0.88-1.71
MACE	Less frequent Vs recommended	2.13	1.31-3.45
	More frequent Vs. recommended	1.45	0.79-2.67

7.3.7 Decision-analytic model

Model structure

A state transition Markov model was designed in order to reflect the progression of amiodarone-related thyroid dysfunction, events which impact on resource use and patient utility in the first five years of stable amiodarone therapy and the typical treatment strategy for ADEs included in the study. The model estimated the mean cost and effect as hypothetical patients progress through the model states until they reach an absorbing state (death) or a terminating condition is met (maximum number of cycles reached). Utility was generated in terms of quality-adjusted life-years (QALYs) in each cycle the patient is alive, as by definition the number of QALYs gained in each alive state is strictly above zero. For cycles spent in ill health (e.g. suffering the consequences of an ADE), the QALY payoffs were discounted according to published evidence, reflecting the adverse impact of the ADE on patient health. Cost was estimated in GBP at 2014 price levels from the perspective of the UK NHS, total cost was assessed by multiplying resource-use associated with each state by unit costs from published UK sources. A Markov process design was used to model the costs and

effectiveness of each alternative strategy. The Markov chain design assumes that all participants move through a series of finite health states until a terminating condition, such as death or a certain proportion of the cohort reaching a certain state. The progression through the model reflects the natural progression of a disease and the standard treatment pathway. Cost and utility collected during time spent in each health state is aggregated and presented after running the model to completion. This modelling method is essentially a cohort simulation method and was chosen due to the ability to apply estimates of the hazard function of ADE estimated in the preliminary cohort analysis to estimate the transition probabilities in the model. This is a computationally efficient method of answering a decision question using both deterministic and probabilistic scenarios. A limitation of this method is its "memoryless" nature; meaning that probability of progression and the cost and utility of a particular health state is independent of time spent in either the current or past states. However, this may be circumvented using extensions, such as temporary tunnel states, which allow a model to keep track of the length of time spent by each patient in a particular state.

Amiodarone side effects involve a number of organ systems other than the thyroid gland, including respiratory, dermatological and other systems. This analysis focused explicitly on the unintended thyroid effects of amiodarone and assumed that the other adverse effects of the drug occur independently of TFT monitoring.

Patients entering the model were assumed to be on a stable course of therapy with amiodarone (e.g. they have gone through the initial monitoring stage) and remain on therapy for a minimum of 4 years. According to guidelines, regular monitoring of amiodarone needs to be conducted once in 6 months during long-term therapy and for 12 months after cessation (British Thyroid Association 2006). The initial distribution of patients across the treatment pathway assumed that all patients started in the "euthyroid" state. This was done in order to reflect the design of the observational study which was used to inform transition probabilities in the model, in which all patients were newly initiated with amiodarone and assumed to have no prior thyroid dysfunction specifically linked to the drug at baseline.

Patients in the model may progress to "untreated AIH", "untreated AIT" or "MACE" at any point during the study period.

Amiodarone-induced AIH and AIT may develop quite rapidly, with significant changes in TSH levels occurring over the course of several months (Basaria and Cooper 2005). The management of amiodarone-induced thyroid dysfunction also takes place over a relatively short period of time. Standard care for overt hypothyroidism involves thyroid replacement therapy in combination with continuing low-dose administration of amiodarone. Thyroid function was assumed to return to normal within 2-4 months (Martino, Aghini-Lombardi et al. 1987). A cycle length of 90 days was chosen in order to reflect the aetiology of disease and the time scale of clinical management of AIH and AIT in practice.

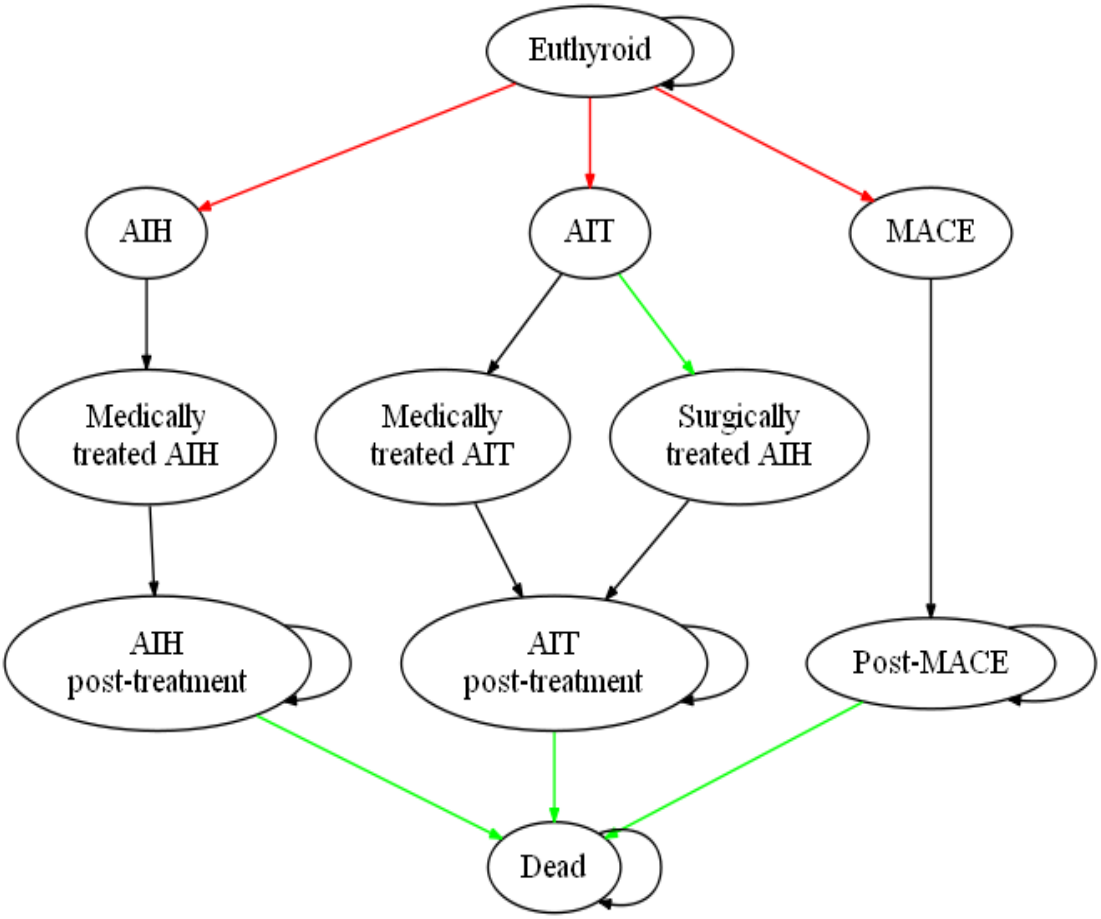
A discrete state transition model recalculates patient numbers after each cycle by assigning patients to health states at the end of the cycle according to the probabilities assigned in the model. However, in reality patients may move into another disease or treatment state at any point during the cycle. For example, assuming that all patients who die do so at the end of a cycle (as opposed to at some point during the cycle) overestimates survival and thus QALY accumulation. Half-cycle correction was applied to all cost and utility values attached to health states by assigning a half-payoff in the first cycle to adjust for the over-estimation. The half-payoff is reassigned after the last cycle if the patient is alive at the end of the last cycle. The state transition model used in the cost-effectiveness analysis is diagrammatically illustrated in Fig. 7.1. Table 7.3 lists all possible transitions within the model.

Table 7.3: Description of sources of data used to populate state transition probabilities

State	No symptoms	AIH	AIT	MACE	AIH med. tr.	AIT med. tr.	AIT sur. tr.	AIH post-tr.	AIT post-. tr.	Post-MACE	Dead
No symptoms											
AIH											
AIT											
MACE											
AIH med. tr.											
AIT med. tr.											
AIT sur. tr.											
AIH post-tr.											
AIT post- tr.											
Post-MACE											
Dead											

Weibull models from CPRD data
Literature
Complement

Fig. 7.1: Model diagram



Red arrows = transitions estimated from CPRD and HES data using Weibull models
Green arrows = transitions estimated from literature
Black arrows = complement of other transition probabilities
Arrows not shown: transitions from every state to "dead", which were estimated from literature
AIH=hypothyroidism; AIT=thyrotoxicity; MACE = major adverse cardiovascular event

State transitions

The model structure is split roughly into two parts: the initial transition from euthyroid state into one of the ADE states (AIH, AIT, MACE) and the subsequent treatment pathway which simulates the standard care for thyroid dysfunction or cardiovascular event. The initial transition probabilities from "euthyroid" to "AIH", "AIT" and "MACE" were calculated using the results of the preliminary time-to-event analysis which employed Weibull regression.

The preliminary time-to-event analysis used Weibull models to estimate the effect of monitoring frequency on the risk of AIH and AIT. Time-to-event models, which take into account censoring and explicitly estimate the hazard function, may be used to inform discrete state transition models as the rates of hazard of the event

of interest estimate in these models can be converted to probabilities of progression in a Markov-type framework. The method of estimation of transition probabilities from the Weibull pdf is described below (adapted from Briggs 2004) (Briggs 2004):

Given that $h(t) = \frac{f(t)}{S(t)}$

Where $h(t)$is the instantaneous hazard function, $f(t)$ is the probability density function (PDF) and $S(t)$ is the survivor function at time t

Applying the Weibull PDF:

$f(t) = \lambda\gamma t^{\gamma-1}\exp\{-\lambda t^{\gamma}\}$

Yields the following properties of the hazard of ADE:

$h(t) = \lambda\gamma t^{\gamma-1}$

$H(t) = \lambda t^{\gamma}$

Where λ≥0 and γ>0 determine the scale and shape of the hazard function at time t.

If the value of the cumulative hazard function of ADE at time *t-u* and *t* are known, the probability of an ADE occurring in the time interval *u* can be calculated as follows:

$tp(t_u) = 1 - \exp\{\lambda(t-u)^{\gamma} - \lambda t^{\gamma}\}$

The shape and scale parameters of the cumulative hazard function estimated using Weibull models are listed in Table 7.4. As the γ parameter was significantly different from 1 in each case ($p<0.05$), assumption of constant transition probabilities over time did not hold. Tunnel states were included into the model design in order to take account of this feature of the data.

Table 7.4: Shape parameter of hazard function of AIH and AIT

Model	γ estimate	95% confidence interval
AIH	1.80	1.52-2.14
AIT	2.45	2.14-2.81
MACE	1.66	1.20-2.29

The transition probability to ADE states in the comparator arms of the model (less and more frequent monitoring) were estimated by multiplying the cumulative hazard function using the hazard ratio estimated in Weibull regressions; the HR estimates are listed in Table 7.2. Details of the regression models used to derive HR estimates are described in Chapters 5 and 6. The full list of 90-day probabilities by treatment arm is included in Table 7.5 below.

Table 7.5: Baseline transition probabilities (for recommended monitoring group) from euthyroid to AIH, AIT and MACE, by strategy and cycle

Cycle no.	AIH		AIT		MACE	
	Mean	SE	mean	SE	Mean	SE
1	0.0024	0.0001	0.0010	0.0001	0.0067	0.0005
2	0.0067	0.0002	0.0047	0.0003	0.0171	0.0010
3	0.0108	0.0004	0.0102	0.0004	0.0257	0.0013
4	0.0147	0.0005	0.0168	0.0006	0.0329	0.0015
5	0.0185	0.0006	0.0244	0.0008	0.0393	0.0017
6	0.0221	0.0007	0.0326	0.0010	0.0450	0.0019
7	0.0256	0.0008	0.0412	0.0012	0.0502	0.0020
8	0.0289	0.0009	0.0502	0.0014	0.0549	0.0022
9	0.0322	0.0010	0.0593	0.0016	0.0593	0.0023
10	0.0354	0.0011	0.0685	0.0018	0.0635	0.0024
11	0.0385	0.0012	0.0777	0.0019	0.0673	0.0025
12	0.0415	0.0012	0.0870	0.0021	0.0710	0.0025
13	0.0444	0.0013	0.0961	0.0023	0.0744	0.0026
14	0.0472	0.0014	0.1053	0.0024	0.0776	0.0027
15	0.0500	0.0014	0.1143	0.0025	0.0807	0.0028
16	0.0527	0.0015	0.1232	0.0027	0.0837	0.0028
17	0.0554	0.0016	0.1321	0.0028	0.0865	0.0029
18	0.0580	0.0016	0.1408	0.0029	0.0892	0.0029
19	0.0605	0.0017	0.1493	0.0030	0.0917	0.0030
20	0.0630	0.0017	0.1578	0.0031	0.0942	0.0030

The shape parameter of the hazard functions were above 1 for both AIH and AIT, which suggests that the instantaneous hazard rate is on a rising trend over time. The longer a patient spent in euthyroid state, the larger their probability of incurring AIH or AIT. The 90-day probability rose from 0.2% in the first cycle to 6.3% in cycle 20 (the last cycle before the model reached its end). The probability of AIT rose from 0.1% to 15.8% over the same number of cycles. Despite the sudden rise, these probabilities are still within a clinically valid range. Previous observational studies of amiodarone have estimated that between 18% and 58% of patients on long-term therapy develop thyroid dysfunction (Martino,

Safran et al. 1984, Albert, Alves et al. 1987, Shukla, Jowett et al. 1994, Tsadok, Jackevicius et al. 2011).

The rising probability of AIH, AIT and MACE over time created the issue of probability coherence as the sum of probabilities from "euthyroid" to ADE and death states may rise above 1. To ensure coherence, probabilities were harmonised to ensure that the maximum sum of probabilities is 1 by applying the following rule:

If *pAIH* = probability of AIH; *pAIT* = probability of AIT; *pMACE* = probability of MACE; *pDeath* = probability of death in any given cycle, let the following hold:

$$(pAIH, pAIT, pMACE, pDeath) = \text{If } pTotal \begin{cases} >1 & \left(\frac{pAIH}{pTotal}, \frac{pAIT}{pTotal}, \frac{pMACE}{pTotal}, \frac{pDeath}{pTotal}\right) \\ \leq 1 & (pAIH, pAIT, pMACE, pDeath) \end{cases}$$

Where *pTotal* = *pAIH*+*pAIT*+*pMACE*+*pDeath*

Treatment pathway

The probability and outcomes of treatment pathways were informed by published studies of treatment of AIH and AIT in consultation with a GP with experience in management of patients on amiodarone therapy (Prof. Anthony Avery). The probabilities related to the outcome of treatment for AIH/AIT were assumed to be independent of the timing of diagnosis of ADE. In other words, the probability of survival following medical or surgical treatment was the same irrespective of whether the ADE was diagnosed early or later on during amiodarone therapy.

AIH

All surviving patients diagnosed with AIH were assumed to undergo medical treatment in the form of thyroid replacement therapy within one cycle of diagnosis. Normal thyroid function is typically restored within 2-4 months of treatment for AIH (Martino, Aghini-Lombardi et al. 1987), therefore the "medically treated AIH" was represented by a tunnel state, with all surviving patients progressing to the post-treatment phase within one 3-month cycle.

AIT

AIT can be split into two pathologically different categories: type I and type II AIT. The former is iodine-induced hyperthyroidism, which occurs due to underlying physiological factors. The latter is acquired from the pharmacological effect of a medication, which may develop spontaneously. Due to complexity

involved in both the diagnosis and treatment of AIT, there is no consensus on the correct treatment strategy for this condition. Thionamides are generally used as the first-line treatment for Type I AIT in the form of carbimazole, which is then converted to methimazole and blocks synthesis of thyroid hormones (Martino, Bartalena et al. 2001). The efficacy of carbimazole may be limited by large presence of iodine, in which case potassium perchlorate may be used (Reichert and de Rooy 1989). Glucocorticoids are used in the treatment of Type II AIT (Bogazzi, Tomisti et al. 2009). In case of non-response to pharmacological treatment, a thyroidectomy may be carried out to resolve AIT (Brennan, van Heerden et al. 1987). The probability of surgical management of AIT via thyroidectomy was assumed to be 8.3% (Tanda, Bogazzi et al. 2008). All surviving patients were assumed to revert to euthyroid state following 1 cycle of surgical treatment for AIT.

All surviving patients who have not undergone thyroidectomy were assumed to receive medical management for AIT within 1 cycle of diagnosis. Euthyroidism is typically restored within 1 or 2 months of medical treatment, therefore the "medically treated AIT" was represented by a tunnel state, with all surviving patients progressing to the post-treatment phase within one 3-month cycle. Discontinuation of amiodarone during treatment for AIT is recommended by the BNF (British Medical Association 2014). Therefore, the model assumed that amiodarone would be stopped during the treatment phase. Once euthyroid state is achieved, amiodarone therapy was assumed to resume.

MACE

Treatment for amiodarone-induced MACE involved emergency hospital admission for myocardial infarction, heart failure, stroke or ventricular arrhythmia. All surviving patients progressed to a post-MACE state, with subsequent mortality and outcomes derived from previous studies which examined patients following adverse cardiovascular events.

Probability of death:

Transition directly to the death state was assumed to be possible from every state in the model. The probability of death was estimated from age-adjusted ONS data (Office of National Statistics 2014), with adjustment made for excess mortality associated with the disease burden of the underlying arrhythmic condition within the study population and the morbidity associated with iatrogenic thyroid dysfunction. The probability of death is considerably higher in a

thyroidectomy and a separate probability was sourced from Houghton et al (Houghton, Farley et al. 2004). The 90-day probability of death from complication during surgery was assumed to be 0.11, based on previous evidence from a cohort of patients with AIT (Houghton, Farley et al. 2004). This was added to the baseline probability of death attached to the "no symptoms" state. The probability of death during and following medical treatment for AIH and AIT was assumed to be the same as the baseline risk of death in the "no symptoms" state as thyroid dysfunction was assumed to have been brought under control. This approach was taken in the only previously published economic model of amiodarone monitoring in the UK (Elliott, Putman et al. 2014). A breakdown of probability of death is included in Table 7.6.

The probability of death from MACE was calculated as the weighted probability of death from each type of MACE episode (heart failure, myocardial infarction, stroke and ventricular arrhythmia). The probability of death in the first 3 months following a diagnosis of heart failure was 0.11 (95% C.I. 0.08-0.14), derived from a study of all-cause mortality from heart failure in the UK (Mehta, Dubrey et al. 2009). The probability of death from acute myocardial infarction was 0.317, taken from a UK-based observational study which used HES data linked with ONS mortality data (the 2010 figures were used due to a downward trend in the incidence of myocardial infarction and associated mortality across the calendar years in the study, as noted by the authors) (Smolina, Wright et al. 2012). The probability of death from an initial inpatient spell due to stroke was 0.16, derived from NHS Scotland statistics on 30-day survival following a stroke in 2012/13 (NHS Scotland 2014). The probability of death from ventricular arrhythmia was assumed to be the same as in euthyroid state given that the included patients were on amiodarone, which is indicated for cardiac arrhythmia, and an adjustment for the increased mortality has already been applied. These probabilities were weighed according to the proportion of MACE events found in the HES database for patients in the cohort of long-term users of amiodarone identified in Chapter 6. The proportions were as follows: heart failure: 0.54; myocardial infarction: 0.25; stroke: 0.15; ventricular arrhythmia: 0.06. The resulting 90-day probability of death from MACE was 0.154.

Recognising that the life expectancy of patients with previously diagnosed heart failure is lower than the average across the model population, the risk of death post-inpatient treatment for heart failure was derived from a prospective screening study in the UK (Hobbs, Roalfe et al. 2007), which reported 5-year all-

cause mortality of 48% in patients diagnosed with heart failure. The 90-day mortality was calculated as 0.024.

The probability of death in patients who have survived the initial MI inpatient episode was derived from a retrospective cohort study of long-term survival post-MI, which found that the risk (HR) of death in 7 years following the event was double compared to the risk of death in the general population in England (Smolina, Wright et al. 2012). The probability of death following an MI in the model was thus assumed to be twice the age-standardised probability of death derived from ONS life tables.

The risk of death following stroke was adjusted using the findings from a retrospective cohort study of stroke survival which matched patients with and without a stroke in a population with atrial fibrillation, which reported a two-fold increase in the risk (HR) of death in the 6 years following a stroke, conditional on survival in the initial 60 days post-event (Fang, Go et al. 2014).

No additional adjustment was made for risk of death following an emergency admission for ventricular arrhythmia as the study population was arrhythmic from the start and the baseline probability of death was already adjusted for excess cardiac mortality attributed to cardiac arrhythmia.

The probabilities were weighed according to the proportion of MACE events. The 90-day probability of death was 0.025.

Table 7.6: Breakdown of death probabilities depending on condition

Probability	Description	Source
Baseline probability of death	Calculated from UK ONS all-cause mortality figures for deaths occurring in 2012, adjusted to 90-day probability assuming a constant daily mortality rate	(Office of National Statistics 2014)
Adjustment for underlying arrhythmic disorder	Based on evidence of increased risk of death from cardiac disrhythmia (RR 1.2 vs baseline)	(Osman, Gammage et al. 2002)
Adjustment for thyroid dysfunction	Risk of mortality assumed to double in the presence of thyroid dysfunction (RR 1.4 vs baseline)	(Yiu, Jim et al. 2009)
Probability of death during surgery	Based on a study of surgical management of AIT (0.11)	(Houghton, Farley et al. 2004)

The list of probabilities of death for each age category is included in Table 7.7. Linear interpolation was used to estimate probabilities for patents of specific age.

Table 7.7: Age-adjusted probability of death, split by ADE state

Age	baseline probability	adjusted for thyroid dysfunction	Adjusted for increased mortality post-MACE
70	0.0061	0.0071	0.0122
75	0.0103	0.0120	0.0205
80	0.0187	0.0218	0.0374
85	0.0342	0.0398	0.0685
90	0.0594	0.06810	0.1188
95	0.1086	0.12550	0.2171

7.3.8 Outcomes

The main outcome in the analysis was the number of quality-adjusted life-years (QALY) gained, which was estimated by multiplying number of years gained or lost as a result of monitoring multiplied by the utility weight corresponding to the health state of the patient ranging from 0 (death) to 1 (full health). QALY weights for each state in the model were sourced from published studies which reported quality of life of patients with thyroid dysfunction on a scale of 0 to 1.

The baseline utility weight which is attached to euthyroid states (no symptoms and post-treatment) was based on the utility estimate reported by Buxton et al for patients with atrial fibrillation (0.78, s.d. 0.21) (Buxton, Caine et al. 2006). Utility of newly diagnosed AIH and AIT was estimated using the utility weight reported by Sullivan et al (0.821 and 0.8 for AIH and AIT, respectively, multiplied by baseline utility) (Sullivan and Ghushchyan 2006). Utility during medical treatment was weighed using figures reported by Dosiou et al (2012) (Dosiou, Barnes et al. 2012) and Nolan et al (1985) (Nolan, Tarsa et al. 1985) (AIH (0.86) and AIT (0.98).

Health states for the MACE state were estimated using utility weights for the individual components of MACE sourced from published studies. The utility weight for heart failure was derived from a UK-based study of a pharmacist-led intervention in newly diagnosed HF patients (baseline EQ-5D utility 0.57) (Holland, Brooksby et al. 2007). The utility weight for stroke was 0.31 (Pickard, Johnson et al. 2005) and for myocardial infarction was sourced from a US-based collection of EQ-5D values for chronic conditions (0.704) (Sullivan and Ghushchyan 2006). In case of vascular arrhythmia, utility was assumed to be unchanged as patients' baseline utility already assumed presence of a cardiac arrhythmia.

Health state in post-treatment states was assumed to be the same as at baseline (Dosiou, Sanders et al. 2008). Utility during thyroidectomy was discounted using the weight reported by Esnaola et al (2001) (0.95) (Esnaola, Cantor et al. 2001). Utilities for each model state are listed in Table 7.8

Table 7.8: Utility weights

State (reference)	Utility weight (SE)	α, β
Baseline utility (euthyroid) (Buxton, Caine et al. 2006)	0.78 (0.04)	78, 22
Utility weight		low
AIH untreated (Sullivan and Ghushchyan 2006)	0.82 (0.04)	71.2, 15.5
AIT untreated (Sullivan 2006)	0.80 (0.11)	10.5, 2.6
AIH treated (Dosiou, Barnes et al. 2012)	0.86 (0.10)	10, 1.7
AIT treated (Nolan, Tarsa et al. 1985)	0.98 (no variance reported)	N/A
Post-treatment (Dosiou, Sanders et al. 2008)	1	N/A
Thyroid surgery (Esnaola, Cantor et al. 2001)	0.95 (no variance reported)	N/A
Heart failure (Holland, Brooksby et al. 2007)	0.58 (0.02)	374.9, 277.1
Myocardial infarction (Sullivan and Ghushchyan 2006)	0.70 (0.02)	259.3, 109
Stroke (Pickard, Johnson et al. 2005)	0.33 (0.04)	44.7, 99.5

7.3.9 Resource use and cost

Resource use assumptions were based on information from published studies on the standard practice in the treatment of AIH and AIT and the opinion of a GP with experience in management of thyroid dysfunction in primary care. Unit costs were derived from published UK sources of unit costs.

Table 7.9: Unit costs of care used in the model

Category	Item	Mean cost (£) (IQR)	Notes	Source
Primary care	GP consultation	46	Based on patient contact of 11.7 minutes	PSSRU 2014 (PSSRU 2014)
Laboratory tests	Thyroid function test	4.45	Cost of a thyroid function test	PINCER (Elliott, Putman et al. 2014)
Secondary care: outpatient	Follow-up cardiology consultation	123.49 (85.19-145.41)	Based on average cost of outpatient visit	NHS reference costs 2013/14 (Department of Health 2014)
	First endocrinology consultation	186.54 (133.54-	Based on average cost of outpatient visit	NHS reference costs 2013/14

		233.11)		
	Thyroid surgery	3172.76 (2469.94-3556.43)	Elective admission - thyroid procedure with CC score 0-1	NHS reference costs 2013/14
Secondary care: emergency inpatient	Heart failure	2771.87 (2039.18-3240.02)	HRG EB03, weighted average of non-elective cost across all CC score categories	NHS reference costs 2013/14
	Myocardial infarction	2243.86 (1673.25-2628.56)	HRG EB10, weighted average of non-elective cost across all CC score categories	NHS reference costs 2013/14
	Stroke	1700.01 (1253.62-1943.16)	HRG AA29, weighted average of non-elective cost across all CC score categories	NHS reference costs 2013/14
	Ventricular arrhythmia	1807.61 (1366.10-2144.33)	HRG EB07, weighted average of non-elective cost across all CC score categories	NHS reference costs 2013/14
Secondary care: emergency day case	Heart failure	525.11 (280.26-707.17)	HRG EB03, weighted average of day care cost	NHS reference costs 2013/14
	Stroke	394.91 (277.76-427.29)	HRG AA29, weighted average of day care cost	NHS reference costs 2013/14
Medications	amiodarone 200mg 1/day	6.43	See table 7.10	BNF 2014 (online)
	levothyroxine 100mcg 1/day	6.36	See table 7.10	BNF 2014 (online)
	carbimazole 10mg 1/day	116.62	See table 7.10	BNF 2015 (online)
	prednisolone 30mg	25.65	See table 7.10	BNF 2014 (online)

Table 7.10: Unit costs of medications

Drug	Formulation	Pack size	Cost per pack (£)	90-day cost (£)	Source
amiodarone 200mg 1/day	200mg	28	2	6.43	BNF 2014 (online)
levothyroxine 100mcg 1/day	100mcg	28	1.98	6.36	BNF 2014 (online)
carbimazole 10mg 1/day	5mg	100	64.79	116.62	BNF 2014 (online)
prednisolone 30mg 1/day	5mg	28	1.33	25.65	BNF 2014 (online)

Table 7.11: Composition of "no symptoms" and "untreated AIH" state cost

Service	Count	Unit costs (£)	90-day cost (£)
GP visit	1	46.00	46.00
Thyroid monitoring	0.5	4.45	2.23
Follow-up cardiology consultation	0.5	123.49	61.74
amiodarone 200mg 1/day	90	6.43	6.43
Total (recommended)			116.40
total (less frequent)			115.30
total (more frequent)			118.62

Patients in the "no symptoms" state included patients on regular amiodarone therapy. Resource use assumptions in this state were taken from a published model of amiodarone monitoring (Elliott, Putman et al. 2014) and were assumed to include 1 GP consultation every 3 months, 1 consultant cardiologist visit every 6 months and ongoing prescription of amiodarone. It was assumed that patients were maintained on 200mg daily dose of amiodarone. The cost of monitoring varied across the comparison groups: patients with recommended frequency of TFT monitoring had 1 test every 6 months, those monitored less frequently were assumed to have 1 test every 12 months, more frequent monitoring was assumed to be once in 3 months, on average.

Patients identified to have overt AIH were not assumed to have any additional contacts with healthcare services until the start of treatment on top of standard amiodarone treatment. They were assumed to be referred to treatment to re-establish control of thyroid function within the next cycle.

Table 7.12: Cost in untreated AIT state

Service	Count	Unit cost (£)	90-day cost (£)
GP visit	1	46.00	46.00
Thyroid monitoring	0.5	4.45	2.225
Cardiology consultation	1	123.49	123.49
amiodarone 200mg 1/day	90	6.43	6.43
Total (recommended)			178.14
total (less frequent)			177.03
total (more frequent)			180.36

The cost of patients diagnosed with AIT but not yet receiving treatment was assumed to be the same as the "no symptoms" state, except the cardiology consultant follow-up is more frequent (once in 3 months, rather than 6 months) due to the appearance of symptoms associated with thyrotoxicity.

Table 7.13: Cost of medically treated AIH

Service	Count	Unit cost (£)	90-day cost (£)
GP visits	3	46.00	138.00
Thyroid monitoring	3	4.45	13.35
Cardiology consultation	0.5	123.49	61.74
amiodarone 200mg 1 tablet daily	90	6.43	6.43
levothyroxine 100mcg 1 tablet daily	90	6.36	6.36
Total			225.89

Patients diagnosed with overt hypothyroidism (TSH>10 mU/L) are maintained on amiodarone and receive thyroid replacement therapy with thyroxine (Basaria and Cooper 2005). Resource use in this state assumed monthly GP visits, monthly monitoring of thyroid function and a prescription of levothyroxine. This was added to the 6-monthly cardiology consultant visit and cost of medication during ongoing amiodarone therapy.

Table 7.14: Cost of medically treated AIT

Service	Count	Unit cost (£)	90-day cost (£)
Endocrinology consultation	1	186.54	186.54
GP visits	3	46.00	138.00
Thyroid monitoring	3	4.45	13.35
Cardiology consultation	1	123.49	123.49
Carbimazole 10mg 1/d - 11% of patients	180	116.62	12.83
Prednisolone (30 mg 1/d 1 month, 20 mg 1/d 1 month, 10mg 1/d 1 month) - 89% of patients	360	25.65	22.83
Total			497.03

AIT can be split into two pathologically different categories: type I and type II AIT. The former is iodine-induced hyperthyroidism, which occurs due to

underlying physiological factors. The latter is acquired from the pharmacological effect of a medication, which may develop spontaneously. Due to complexity involved in both the diagnosis and treatment of AIT, there is no consensus on the correct treatment strategy for this condition. Thionamides are generally used as the first-line treatment for Type I AIT in the form of carbimazole, which is then converted to methimazole and blocks synthesis of thyroid hormones (Martino, Bartalena et al. 2001). The efficacy of carbimazole may be limited by large presence of iodine, in which case potassium perchlorate may be used (Reichert and de Rooy 1989). Glucocorticoids are used in the treatment of Type II AIT (Bogazzi, Tomisti et al. 2009).

The assumed resource use during medical treatment of AIT was based on a previously published model of amiodarone monitoring (Elliott, Putman et al. 2014) and included monthly GP visits, monthly monitoring of thyroid function, 3-monthly endocrinology consultation. Assumptions regarding drug therapy were based on the proportion of patients with type I and II AIT, which was 11% and 89%, respectively (Bogazzi, Bartalena et al. 2007). Patients with Type I AIT were assumed to receive carbimazole and 89% were prescribed a glucocorticoid (prednisolone). This was added to the 3-monthly cardiology consultant visit. Discontinuation of amiodarone during treatment for AIT is recommended by the BNF (British Medical Association 2014). Therefore, the model assumed that amiodarone would be stopped during the treatment phase. Once euthyroid state is achieved, amiodarone therapy was assumed to resume.

Table 7.15: Cost of surgically treated AIT

Service	Count	Unit cost (£)	90-day cost (£)
Endocrinology consultation	2	186.54	373.08
GP visits	3	46.00	138.00
Thyroid monitoring	3	4.45	13.35
Cardiology consultation	1	123.49	123.49
Carbimazole 10mg 1/d	180	116.62	116.62
Major thyroid procedures w/o CC	1	3172.76	3172.76
Total			3936.30

In case of non-response to pharmacological treatment, a thyroidectomy may be carried out to resolve AIT (Brennan, van Heerden et al. 1987). The cost of surgical management of AIT was assumed to include the cost of a planned

hospital admission for the thyroidectomy (represented by the HRG for major thyroid procedure), monthly GP visits, monthly monitoring of thyroid function, 2 consultant endocrinology outpatient visits, 3-monthly cardiologist consultant outpatient visit and prescription of carbimazole. Amiodarone therapy was assumed to be suspended for the duration of treatment.

Table 7.16: Cost of MACE state

Service	Diagnosis	Weight	Cost (£)
Inpatient	Heart failure	0.51	2771.87
	Myocardial infarction	0.25	2243.86
	Stroke	0.04	1700.01
	Ventricular arrhythmia	0.15	1807.61
Day care	Heart failure	0.03	525.11
	Stroke	0.01	394.91
	Weighted average		2333.46

MACE comprised heart failure, myocardial infarction, stroke and ventricular arrhythmia (Yiu, Jim et al. 2009). The cost of the MACE state was calculated as a weighted average of the unit costs of non-elective admissions for each of the above conditions derived from NHS reference costs 2013/14.

Table 7.17: Cost post-treatment of AIH and AIT

Service	Count	Unit cost (£)	90-day cost (£)
GP visit	1	46.00	46.00
Thyroid monitoring	0.5	4.45	2.23
Cardiology consultation	0.5	123.49	61.74
amiodarone 200mg 1/day	90	6.43	6.43
Total			116.40

The resource use following successful treatment of AIT and AIH was assumed to be the same as for euthyroid patients, with monitoring conducted at recommended frequency.

7.3.10Analysis design

Base case analysis

The economic consequences of each competing monitoring strategy were assessed using an incremental analysis of the cost and utility per patient generated in the 5 year horizon of the model. Pairwise incremental analyses were

carried out between the three strategies in the order of effectiveness using incremental cost-effectiveness ratios (ICERs), which is a ratio of the difference in cost and QALYs accumulated at 5 years calculated for each pair (e.g. if less frequent monitoring, recommended and higher frequency are in ascending order of QALYs generated, head-to-head comparisons are recommended vs. less frequent and more frequent vs. recommended). In the expected value (EV) analysis, the ICER has analytic value only if it is positive; a negative ICER does not, by itself, provide an unambiguous measure of cost-effectiveness as either the numerator or denominator of the ICER may be negative. Moreover, it does not make much sense to represent an ICER in decreasing order of effectiveness (e.g. a positive ICER with negative cost and effectiveness), unless the less effective option has been explicitly identified as the intervention in the study design. It is assumed in this case that the healthcare decision maker is pursuing health improvement as an objective (or avoiding the adverse consequence of ADE on patient health), rather than seeking cost savings, potentially at the expense of patient utility. Therefore, ICERs were represented only in situations where both the incremental cost and QALYs accumulated were positive or both were negative. In other situations, the costs and outcomes were presented and interpreted separately.

Deterministic sensitivity analysis

The base case analysis described in the previous section was used to examine the incremental cost per unit of outcome across the comparators in order to inform an optimal decision. The relative importance of different factors in the decision problem may be assessed by measuring the marginal effect of key model parameters on the ICER. It may help to identify elements of the model where changes in the values of individual parameters may influence strategy selection. If such values rely on incorrect assumptions or their true value is uncertain, incorrect conclusions may be drawn from the analysis. This assessment was done via tornado and one-way sensitivity analyses. The effect on strategy selection was measured in terms of net monetary benefit (NMB), which is calculated as follows:

$$NMB = QALYs \cdot WTP - Cost$$

NMB is the monetary value of the health benefit of a strategy net of cost. It is preferable to the ICER in sensitivity analyses involving more than two alternatives as it explicitly incorporates the decision-maker's willingness to pay and does not require head-to-head comparisons across the alternatives. The value attached to

each QALY generated was assumed to be the ceiling willingness-to-pay of £20,000 per QALY, as recommended by NICE (NICE 2004).

Model parameters were grouped into categories: treatment effect of monitoring, transition probabilities, costs and utilities. One-way sensitivity analyses were carried out to gauge the effect of varying model inputs on the results of the cost-effectiveness analysis. Defensible ranges were selected which reflect clinically valid intervals for each parameter in the deterministic sensitivity analysis.

Probabilistic sensitivity analysis

Up to this point all model parameters, including transition probabilities, costs and utilities, have been represented using expected values. These figures were used to conduct the deterministic analysis and obtain point estimates for the ICER. Analysis of the impact of variation in parameter values was limited to one-way sensitivity analyses. However, a sound modelling analysis needs to assess and report the uncertainty around cost and effectiveness estimates and its impact on decision uncertainty (Briggs, Weinstein et al. 2012). A probabilistic sensitivity analysis was undertaken to incorporate parametric uncertainty. This involved measuring the effect of varying estimated values of all parameters simultaneously by making distributional assumptions for probabilities, costs and utilities which feed into the model, constructing confidence intervals for ICER values and representing the uncertainty involved in the decision problem set out at the beginning of this chapter.

Monte Carlo simulation was used to simulate the patient pathways and generate ICER estimates. 5000 iterations of the model were run in order to eliminate first-order uncertainty in the process and produce stable estimates of incremental cost and effectiveness. Uncertainty around the mean of the ICER was assessed by estimating the 2.5^{th} and 97.5^{th} percentile of the ICER using bootstrapping with replacement. Estimates were mapped on cost-effectiveness planes in order to visually illustrate the distribution. Decision uncertainty was illustrated using a cost-effectiveness acceptability curve (CEAC) which represents the decision problem in the context of the decision-maker's willingness-to-pay ranging from £0 to £100,000. This interval was selected as £0/QALY represents a situation where only cost-saving strategies may be accepted and £100,000/QALY represents the upper bound of the amount of resources which may be devoted by a healthcare system in a developed economy. The region which is relevant in context of current policy ranges between £20,000 and £30,000/QALY, which is the reference WTP for a QALY listed by NICE (NICE 2004).

Distributional assumptions for transition probabilities

All probabilities were modelled using beta distribution. This distribution type is appropriate for probabilities as it restricts values to a range between 0 and 1 and is skewed towards the middle of the range.

Uncertainty in probability of transition to ADE states is informed by the standard errors of the λ and γ parameters of the hazard function. The standard deviations were estimated by bootstrapping the value of the transition probability in each cycle with 5000 iterations, assuming a joint normal distribution of the parameters. Collinearity was controlled by applying Cholesky decomposition. Probability coherence was ensured by harmonising probabilities in cases when their sum exceeded 1.

All transition probabilities in the treatment pathway were assigned distribution parameters. Figures from original observational data were used to calculate distributional parameters directly in cases where such information was available. In other cases, standard deviations or standard errors were converted to beta parameters α and β using the following formulas:

$$\alpha = \frac{\mu^2 \cdot (1-\mu)}{\sigma^2 - \mu}$$

$$\beta = \frac{\mu \cdot (1-\mu)^2}{\sigma^2 - \mu}$$

Distributional assumptions for costs

All unit costs were modelled using the gamma distribution. This distribution type was appropriate as it supports a heavy right tail, which reflects the typically strong positive skew in cost variables.

Unit costs sourced from the NHS reference costs include lower and upper quartile estimates. These figures were used to fit a gamma 95% confidence interval and derive the gamma distribution parameters. Certain cost parameters were not assumed to vary from case to case. These included medication costs, which use standardised charges across the healthcare system. Others, such as primary care consultation costs were reported in PSSRU without a range or confidence interval.

Distributional assumptions for utilities

Utilities, just like probabilities, are restricted to a range between 0 and 1. The beta distribution was therefore deemed to be an appropriate tool to utility estimates. Details of the beta parameters of utility variables are listed in Table 7.8.

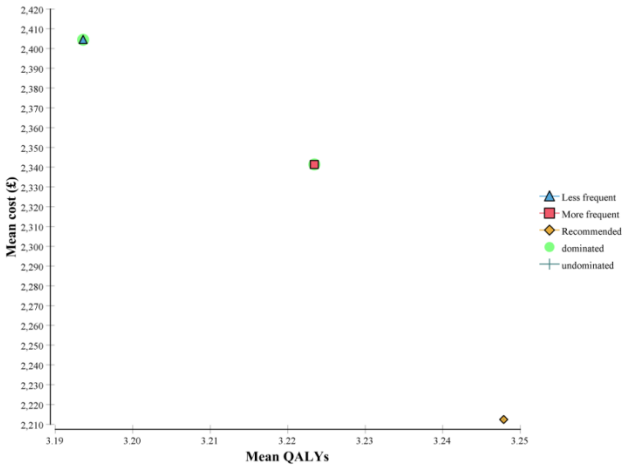
7.4 Results

7.4.1 Base case analysis

Table 7.18: Results of deterministic analysis.

Strategy	Mean discounted QALYs	Mean discounted cost (£)
Less frequent	3.194	2404.44
More frequent	3.223	2341.40
Recommended	3.248	2212.37

Fig. 7.2: Expected values of cost and utility accumulated in each strategy



The strategy involving recommended frequency of TFT monitoring generated more QALYs per patient over the 5 year time horizon at lower mean cost per patient. Therefore, based on expected values of cost and effectiveness, the strategy involving recommended frequency monitoring was dominant over the other two strategies. This is illustrated in Fig. 7.2, which maps the three strategies on the cost-effectiveness plane. Recommended monitoring is located lower and to the right of the points representing the other strategies, which represents absolute dominance.

In terms of expected values, all three strategies were similarly effective in terms of QALYs generated over the model horizon. Although recommended monitoring was the most effective strategy, it yielded only 0.0245 QALYs per patient over 5 years more than the next best strategy in terms of effectiveness (less frequent monitoring) and 0.0543 more QALYs compared to more frequent monitoring. The mean cost per patient was lowest in the recommended monitoring group, generating a mean cost saving of £129 over less frequent monitoring and £192 over more frequent monitoring.

7.4.2 Deterministic sensitivity analysis

Given the small difference in estimates of mean cost and effectiveness between treatment arms, changes in assumed values of individual input parameters in the model may have an impact on the choice of strategy. Sensitivity analyses of the effect of changing key variables in the model on net monetary benefit (NMB) of each strategy were calculated.

Uncertainty in the estimates of model parameters was investigated across four types of input parameter: effect of monitoring on risk of ADE (via hazard ratios estimated in the time-to-event analysis), transition probabilities in the treatment pathway, unit costs and utilities. The ranges used in the one-way analyses are summarised in Table 7.19

Table 7.19: Parameter ranges used in the DSA.

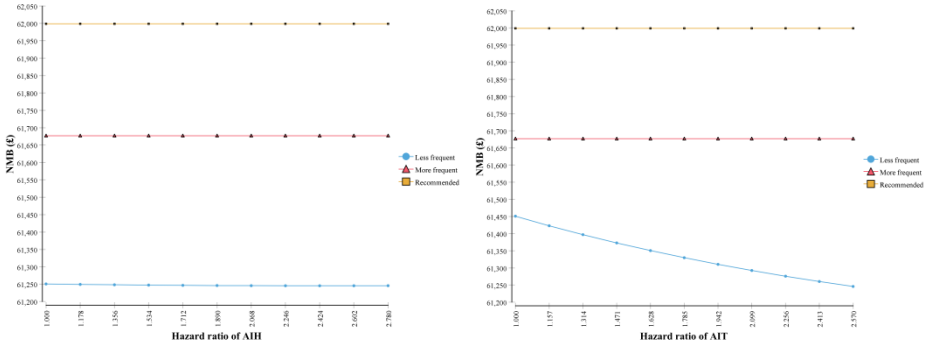
Effectiveness of monitoring			
Variable	Low	High	Rationale
HR of AIH (less frequent/recommended)	1	2.78	In order to examine the sensitivity of the model results to changes in the estimated effect of monitoring, the hazard ratio of each ADE in the less frequent group was changed to 0.
HR of AIT (less frequent/recommended)	1	2.57	
HR of MACE (less frequent/recommended)	1	2.13	
Transition probabilities			
Probability of thyroidectomy in AIT	0	0.083	The original source did not report a confidence interval. Of interest to this study was the reduction in the mean cost of treatment if a lower proportion of patients required thyroidectomy in AIT.
Probability of death from thyroidectomy	0.0086	0.11	Due to high mortality following thyroidectomy, QALYs generated may be sensitive to the estimated survival in this patient group. The probability was derived from a small study of 34 cases of thyridectomy, and was thus prone to uncertainty. The probability was reduced to the baseline probability of death in patients with thyrotoxicity to test the scenario of no excess death in surgery.

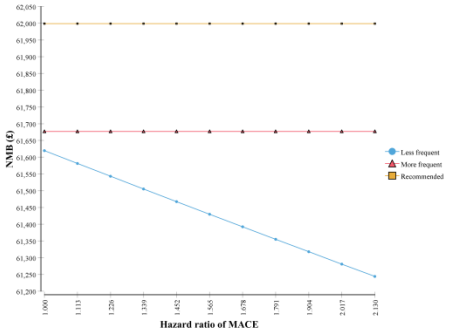
Probability of death from AIH/AIT	0.0074	0.0086	Examine the effect of no excess mortality from AIH/AIT by reducing probability to baseline mortality for patients with arrhythmia.
Probability of MACE to death	0.0086	0.1539	Death resulting from MACE episode had an impact on QALY generation. The effect of reducing this probability to the baseline probability of death in patients with thyroid dysfunction was explored in order to examine a scenario of no excess mortality from MACE.
Cost of treatment of ADEs			
MACE	1724.54	2333.46	The cost was varied to the 25th percentile of the estimated cost of each component of MACE listed in NHS Reference Costs.
AIT surgical	3234.48	3936.29	The cost was varied to 25th percentile of NHS reference cost for thyroid surgery, as the main component of the cost for this state.
AIT medical	405.73	497.03	The cost was varied to 25th percentile of the endocrinology and cardiology outpatient visit cost in NHS reference costs.
AIH medical	187.59	225.89	The cost was varied to 25th percentile of the cardiology outpatient visit cost in NHS reference costs.
Utility			
Untreated AIH	0.821	1	The scenario of no utility decrement was examined. Utility decrements for AIH and AIT states was increased to 1 which was the upper bound of the confidence interval, reflecting no loss of utility attributed to thyroid dysfunction.
Untreated AIT	0.8	1	
Treated AIH	0.8557	1	
Treated AIT	0.9755	1	

Assumptions regarding effect of monitoring on HR of ADEs

In order to assess the impact of varying assumptions regarding the hazard ratios input into the model on optimal strategy selection, one-way sensitivity analyses were performed to estimate the effect if the assumed effectiveness of recommended monitoring is reduced. Hazard ratios of AIH, AIT and MACE for less frequent and more frequent strategies relative to recommended monitoring were reduced to one to represent a scenario with no difference in the rate of ADE across groups.

Fig. 7.3: One-way sensitivity analysis of estimated effectiveness of TFT monitoring.



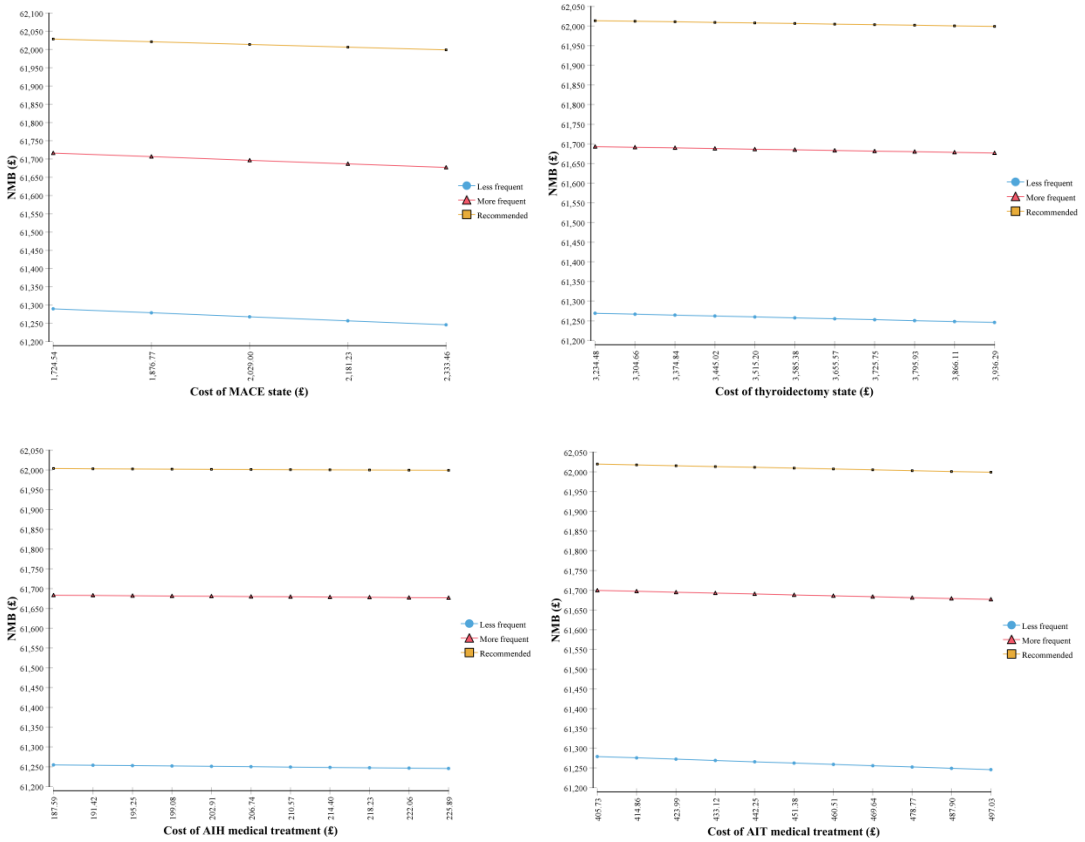


The largest effect on NMB was in the HR of MACE. Changing the HR of each ADE in turn did not result in a change of optimal strategy.

Effect of uncertainty in cost parameters

Another potential source of uncertainty in the model is the cost of treatment of AIH, AIT and MACE.

Fig. 7.4: One-way sensitivity analyses of cost parameters.

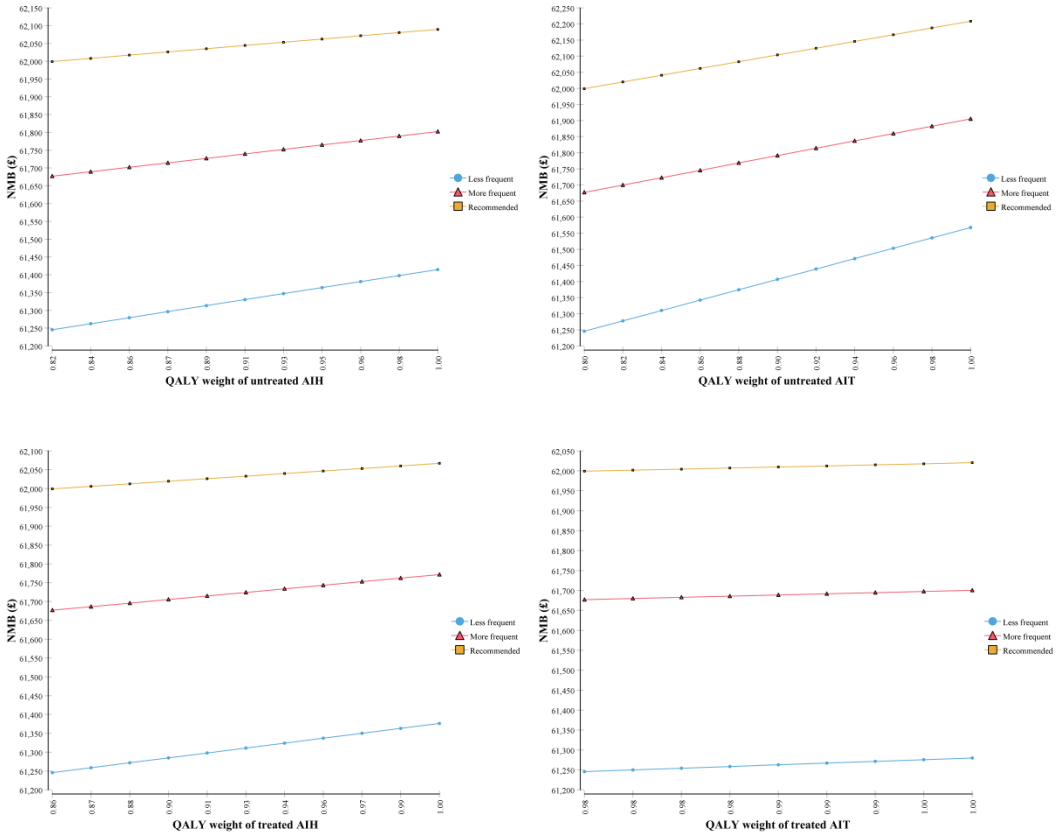


Varying the state cost for AIH, AIT and MACE did not result in a change in optimal strategy.

Effect of uncertainty in utility parameters

Another potential source of uncertainty which could impact on strategy selection is utility weights. The impact of varying utilities attached to AIH, AIT and MACE states across defensible ranges of the original mean values was illustrated in the following figures:

Fig. 7.5: One-way sensitivity analyses of utility parameters.

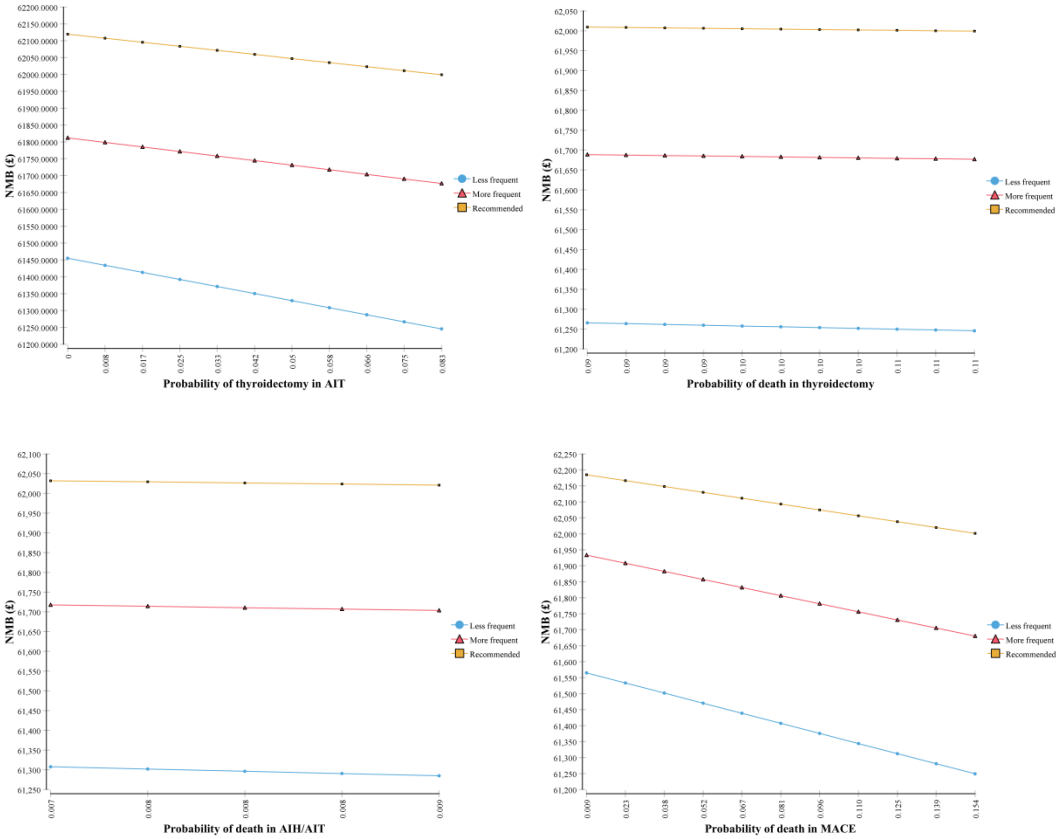


Varying the assumed QALY weights associated with treated or untreat thyroid dysfunction did not result in a change of optimal strategy.

Effect of uncertainty in transition probabilities

The effect of varying assumed probabilities of transitioning between states in the model was illustrated in a tornado analysis. The probabilities included the probability of surgical treatment of AIT and transition to the death state from AIH/AIT/MACE, during and after treatment.

Fig. 7.6: One-way sensitivity analyses of transition probability parameters



7.4.3 Probabilistic sensitivity analysis.

Table 7.20: Results of probabilistic sensitivity analysis.

Strategy	Mean (SD) QALYs generated	Mean (SD) cost generated (£)
Less frequent	2.98 (0.44)	2241.95 (314.48)
More frequent	3.01 (0.442)	2182.86 (306.82)
Recommended	3.03 (0.444)	2064.45 (281.35)

Based on expected values of cost and outcomes generated using Monte Carlo simulation, recommended monitoring was dominant against the other two strategies as it generated more QALYs per patient (0.02 (95% C.I. -0.02,0.06) and 0.05 (0.01,0.10) compared to more frequent and less frequent monitoring, respectively) at lower cost per patient at the 5 year horizon (-£118 (-268,32)) and -£178 (-341,-14) compared to more frequent and less frequent monitoring, respectively). However, due to substantial parametric uncertainty, the bootstrapped 95% confidence intervals around the ICER estimates were extremely large.

In order to illustrate the effect of parametric uncertainty in the cost-effectiveness analysis, bootstrapped estimates of incremental cost and QALY per patient were mapped on cost-effectiveness planes. Two-way comparisons were made in the order of effectiveness of each intervention: recommended vs. more frequent; more frequent vs. less frequent.

Fig. 7.7: Cost-effectiveness plane: recommended vs. more frequent monitoring

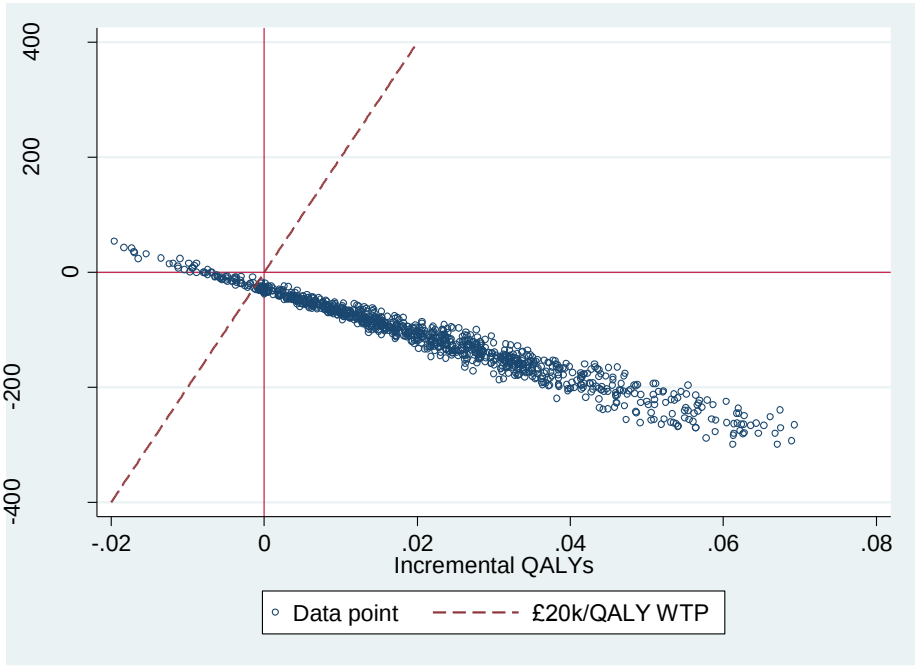
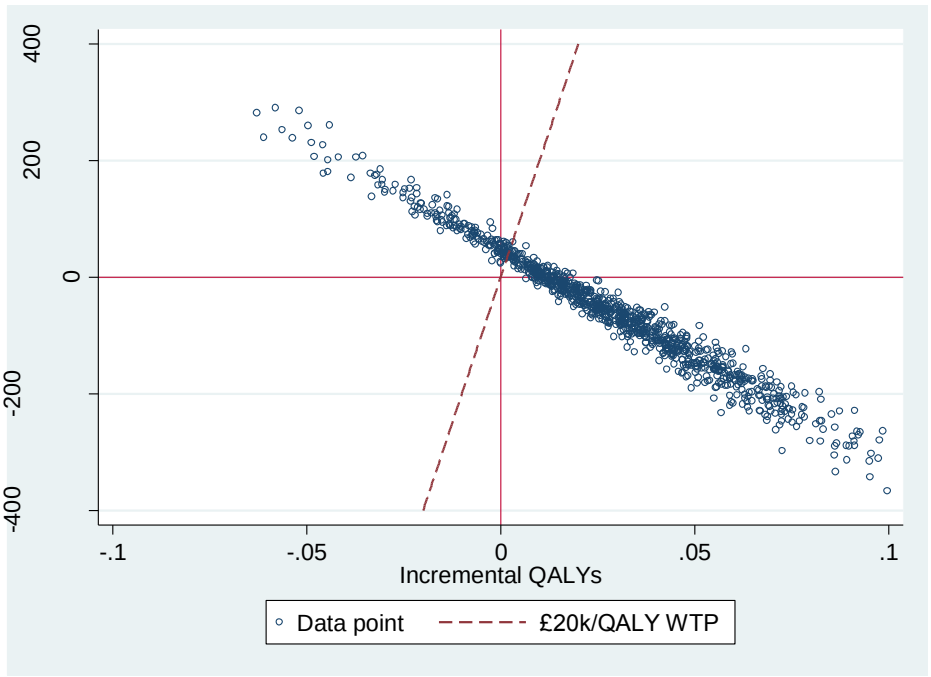
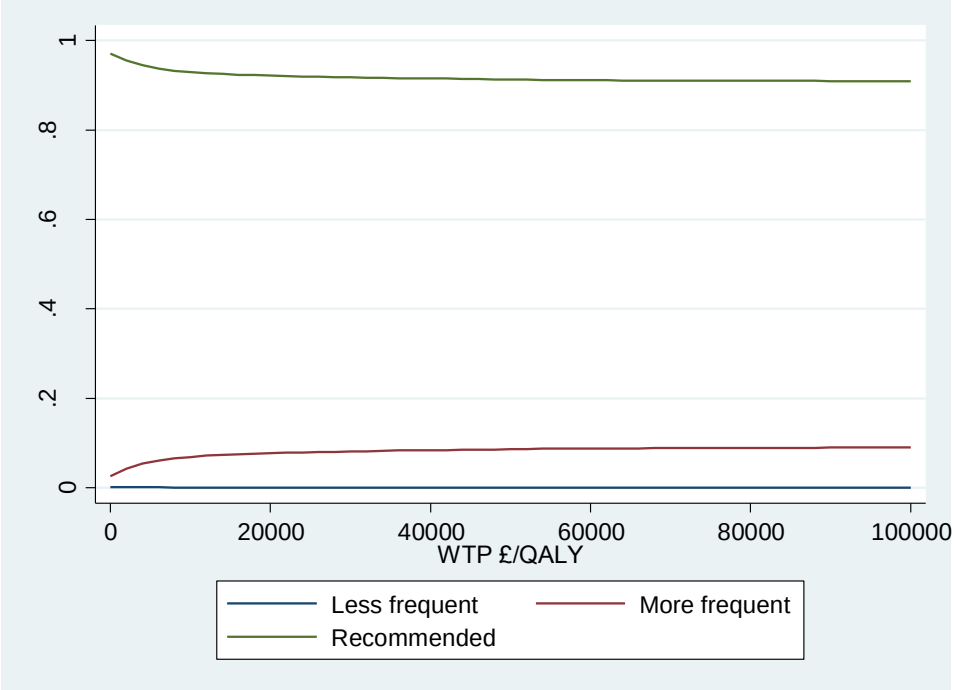


Fig. 7.8: Cost-effectiveness plane: more frequent vs. less frequent monitoring.



Decision uncertainty in the context of a comparison of all three strategies was represented using a cost-effectiveness acceptability curve.

Fig. 7.9: Cost-effectiveness acceptability curve, 5 year horizon.



Monitoring thyroid function during amiodarone therapy at recommended frequency of once in 6 months was the strategy which had the highest probability of being cost-effective compared to alternative strategies involving monitoring less or more frequently than the frequency listed in guidelines, irrespective of the ceiling willingness-to-pay per additional QALY gained. The preferred strategy had a probability of 97% of being cost saving and at least 90% cost-effective at a ceiling willingness to pay of up to £100,000 per QALY.

7.5 Discussion

This chapter described an investigation of the cost-effectiveness of different frequencies of monitoring amiodarone therapy using a decision-analytic model design. Based on expected values of incremental cost and effectiveness, a strategy involving recommended frequency of monitoring was associated with both an increase in QALYs generated and lower cost accumulated at 5 years compared to strategies involving monitoring at a less or more frequent rate, although the actual differences in incremental cost and QALYs gained were small. In the deterministic analysis, the baseline strategy resulted in 3.248 QALYs gained at 5 years, at a mean discounted cost of £2212.37. This represented a larger number of QALYs generated at a lower cost compared to the other two

strategies. This is explained by the improvements in the prevention of thyroid dysfunction resulting from regular TFT monitoring and the resulting reduced cost of treatment of AIH and AIT in primary and secondary care. The preliminary analysis demonstrated that monitoring above recommended levels did not result in lower risk of AIH or AIT (conversely, the analysis showed an increased risk after adjustment for treatment assignment, albeit not statistically significant at the 5% level). This was reflected in the cost-effectiveness analysis, which showed that the additional investment in more intensive monitoring was not justified.

Differences in effectiveness and cost between comparators were relatively small, meaning that changing the assumed value of a single variable may lead to a change in the preferred strategy. One-way sensitivity analyses were conducted to examine the effect of uncertainty in individual model parameters on the results of the EV analysis by comparing the NMB of each strategy, assuming a WTP of £20,000/QALY. First, the treatment effect of monitoring estimated in the time-to-event analysis was reduced by decreasing the hazard ratios of AIH and AIT in the model to one to gauge the effect of a reduction in the benefit of regular monitoring on strategy selection. The parameter which was associated with the largest change in net benefit of less frequent monitoring was HR of MACE, but there was no change in optimal policy even if HR was assumed to be 1 (no protective effect of regular monitoring against MACE). The effect of cost inputs was examined by reducing the key drivers of cost resulting from ADEs, including thyroidectomy, MACE and medical management of AIH/AIT. The base case analysis results were robust to changes in cost estimates, which had a marginal impact on net benefit of each strategy. Similarly, the base case results were robust to changes in utility inputs. The QALY weight of untreated AIT had the largest impact on net benefit, but the relative effect on NMB in each strategy was not large enough to change the preferred strategy. Other parameters examined were transition probabilities which determined the progression of patients though the disease pathway. Changes in the assumed probability of thyroidectomy in AIT and death in MACE had the largest effect on NMB, but did not result in changes in strategy.

A probabilistic sensitivity analysis was conducted to gauge the combined effect of uncertainty of all input parameters on the model output. It demonstrated that a strategy involving TFT monitoring at the recommended frequency of once in 6 months was highly likely to be cost-effective: the probability of cost saving was 97% and probability of cost effectiveness was above 90% across the whole range of ceiling willingness-to-pay assessed in the PSA (£0 to £100,000 per QALY).

Two previous studies have investigated the cost-effectiveness of interventions aimed at reducing errors and ADEs through improved monitoring. The PINCER study, which examined a pharmacist-led intervention to reduce clinically important errors in primary care found that the intervention was dominant compared to a comparator of simple feedback in the base case scenario and in case of targeting monitoring errors (Elliott, Putman et al. 2014). Targeting the intervention to improve amiodarone monitoring generated an ICER of £475 (standard erorr £15). Inclusion of parametric uncertainty impacted on the confidence of cost-effectiveness of the intervention: at £20k/QALY WTP, the intervention was cost-effective in 59% of cases, targeting monitoring errors was cost-effective 53% of the time and targeting amiodarone reached 100% confidence of cost-effectiveness at the standard threshold.

The results of the current analysis were consistent with the results of the PINCER study in terms of EV analysis: recommended frequency monitoring was found to be cost-saving compared to monitoring less frequently than recommended (which was considered a monitoring error according to PINCER methodology). Similarities were found in the results of the PSA; both studies demonstrated that regular monitoring of amiodarone is a cost-effective and potentially cost-saving strategy. There were important distinctions in the model designs: the PINCER study used information from published studies on the probability of progression of AIH and AIT, whilst the current study estimated transition probabilities for ADE from a time-to-event analysis populated by UK-based observational data. In addition, the current study explored the effect of frequency of monitoring on the risk of ADE, rather than just presence or absence of monitoring, and found that monitoring above recommended frequency generated a higher cost and did not result in improvement in utility.

A US-based trial-based economic evaluation of a pharmacist-led outreach programme to improve monitoring of cardiovascular medications reported an ICER of $522,000 (USD) per QALY (<1% probability of cost-effectiveness at WTP of $30k/QALY in the PSA), however when targeted at patients presenting with risk factors, such as presence of diabetes mellitus and chronic kidney disease, the intervention was dominant and highly likely to be cost-effective at the same WTP level (Smith, Raebel et al. 2011). This suggests that interventions to improve monitoring of medications beyond current practice have limited value for the money invested when rolled out to the whole population, but may be cost-effective when carefully targeted at those patients for whom improvements in medication safety may prove beneficial. In the context of this finding, further

research is needed to determine whether monitoring amiodarone therapy more frequently than recommended in guidelines may be a cost-effective option in patient subgroups with a higher risk of developing amiodarone-induced thyroid dysfunction.

The principal strength of this study lies in the reliable design of the analysis used to answer the research question. Based on the literature review described earlier in this thesis, this is the first study which has employed observational data from a large database of electronic health records to inform a modelling-based economic analysis of monitoring of a high-risk medication in UK primary care. The application of time-to-event analysis informed by a large dataset from a validated, high-quality source of data of prescription records, diagnoses and monitoring tests allowed the investigator to accurately estimate the effect of monitoring frequency on the risk of specific ADEs linked to amiodarone in a UK primary care context. The application of the propensity score in the initial analysis design stage permitted to balance the distribution of baseline covariates across comparison groups, controlling for selection bias arising from non-random treatment assignment. Accelerated-time models in the form of Weibull regression permitted to estimate the hazard function of AIH and AIT whilst taking into account non-informative censoring.

The time horizon employed in the modelling analysis is a potential limitation of the study. Although the 5-year horizon of the model was justified by design aspects, described in more detail in Section 7.3.5, a lifetime horizon may have added more credibility to the study results. Further research is needed in order to find a functional form of the hazard function which may be more applicable in decision-analytic modelling of survival data under conditions of exponentially rising hazard of ADE. The cohort simulation design was chosen for its advantages in terms of applicability of the results of a prior survival analysis to a Markov chain design and due to the simplicity it affords in terms of calculation of incremental cost and effectiveness. This design allowed little opportunity to assess the cost-effectiveness of monitoring in subgroups of interest, such as frail older people or those with risk factors for thyroid dysfunction. Alternative modelling methods such as discrete event simulation or microsimulation may shed light on risk factors and patient characteristics which may potentially affect the cost-effectiveness of monitoring frequency.

If a healthcare planner's decision on allocation of resources to routine amiodarone monitoring in primary care were based purely on EV analysis presented in this

study, the choice of recommended monitoring would be unambiguous. However, a decision based purely on expectations is only appropriate in case of an immediate decision or when opportunities to obtain additional information are limited. Assuming that such criteria do not apply in the context of decision-making in UK primary care, any decision needs to incorporate the uncertainty of costs and outcomes of alternative options. The PSA presented in Section 7.4.3 found that recommended frequency monitoring was the optimal strategy, irrespective of the ceiling WTP per QALY, even when uncertainty in model parameters had been fully taken into account.

In any situation where the third-party payer is willing to invest any amount of money per QALY gained for an intervention to improve the safety of amiodarone therapy, monitoring in accordance with guidelines is the cost-effective option, therefore the current recommended frequency of monitoring (once in 6 months) can be deemed to be within the acceptable range. Given the lack of evidence in favour of changing the current recommendations, this analysis supports the current policy of encouraging monitoring of thyroid function once in 6 months during long-term amiodarone therapy.

Comment on appropriateness of standard health-maximisation approach in economic evaluation of healthcare safety interventions.

Healthcare safety interventions may have additional non-health benefits in addition to avoiding health and utility losses caused by adverse events and cost savings resulting from avoiding treatment of preventable harm. Patients who have experienced an adverse event as a result of a medical error may feel isolated, helpless and fearful that the error may reoccur in the future. The psychological burden following a medical error may also affect clinicians due to feelings of guilt, fear of litigation and worry of the impact the error may have on their future career, their professional reputation and the opinion of colleagues (Delbanco and Bell 2007). Reduction in dreadfulness, improvements in controllability and voluntariness have been identified as important attributes of healthcare safety interventions, alongside reduction in the likelihood and health impact of errors (Steuten and Buxton 2010). The occurrence of errors in healthcare undermines the confidence of patients and the trust of the public in the system, which may produce suboptimal outcomes. Therefore, interventions which prevent medical errors have the potential to produce extra-consequentialist benefits which are hard to measure as part of an economic evaluation. This has

led some authors to suggest that standard methods for economic evaluation involving health-maximisation may undervalue the benefits of interventions in healthcare safety and that alternative methods which incorporate these non-health aspects would more accurately assess the value for money of preventive interventions (Steuten and Buxton 2010, Payne, McAllister et al. 2013).

There is evidence that medication safety is not the sole area of healthcare where patients may seek additional non-health benefits from interventions beyond improvement of their physical or mental condition. For example, qualitative research of the preferences of patients undergoing treatment for psychosis attach value to psychosocial factors, such as participation in social activity, independence and self-confidence, as well as reduction in psychotic symptoms (Byrne and Morrison 2014). Genetic testing services have been found to have both positive and negative effect on patient anxiety and depression resulting from receiving information regarding the risk of cancer (Payne, Nicholls et al. 2008). Other studies have reported that societal preferences may be influenced by intangible aspects of the intervention or condition. There is evidence that individuals place a higher value to preventing health losses caused by failures in healthcare, such as hospital-acquired infections and medication errors, compared to adverse effects which were perceived to have been caused by the patient themselves, such as conditions related to lifestyle choice or sport injuries (Singh, Lord et al. 2012). This suggests that public preferences are affected by responsibility for health losses, with a higher value attached to preventing health losses which were imposed upon patients by poor quality care. Some experts put into question the "QALY is a QALY" health-maximisation approach in economic evaluation, which assumes that units of health (such as a QALY or another measure used in cost-utility analysis) are assumed to have the same value irrespective of context, in cases where the end users of healthcare or society attach a higher priority to certain types of care or a particular healthcare context (Mooney 1994, Coast, Smith et al. 2008).

Decision problems involving multiple competing attributes may be objectively incorporated into an economic analysis using conjoint analysis (CA). This method estimates the relative value of attributes and trade-offs between different attributes by the end users of a healthcare technology in order to select an option with the highest net benefit (Ryan 1999). Cost and outcomes in monetary terms are included as attributes, allowing different technology options to be ranked by WTP.

One suggested alternative involves the use of a QALY maximisation approach into an evaluative frameworks which includes a wider set of objectives, known as multi-criteria decision analysis (MCDA) (Baltussen and Niessen 2006). This approach allows a combination of multiple criteria which may include incremental cost, QALYs or WTP generated using standard economic methods with other aspects which may be important to population welfare, such as severity of disease (or prevented ADE), impact of the intervention on under-served population groups, as well as potential non-health benefits of healthcare safety explored by Steuten et al and Delbanco et al. The issue with this approach is that whilst measures of non-health benefit have been developed, they lack in conceptualisation and consistency of definition, such as in the case of empowerment as a potential benefit of genetic screening (Herbert, Gagnon et al. 2009).

In the presence of sufficient evidence that interventions of healthcare safety generate hard-to-measure extra-consequentialist benefits to patients, their families and healthcare professionals, it may be argued that the ceiling willingness-to-pay threshold set by NICE should be raised for these interventions in order to compensate for these missing intangible effects in the denominator of the ICER. However, this would involve placing a monetary value on inherently subjective effects such as anxiety and dreadfulness, which would be practically challenging.

Amartya Sen's capability approach represents an alternative school of thought in academic health economics which posits that individual welfare is determined by what one is capable of doing (capability), rather than what they actually achieve (functioning) (Sen 1993). In his work, Sen illustrates using as an example a person who is starving compared to another who is fasting. Because the latter made a choice to forego food, their welfare may be higher, despite the same effect in terms of functioning. The proponents of this theory argue that the classic extra-welfarist health-maximisation approach to economic evaluation does not sufficiently take into account an individual's capability for well-being, which is not necessarily related to their current state of health. The idea that capability by itself may determine individual welfare was conceptualised by Martha Nussbaum in terms of ten capabilities which determine individual well-being, including the ability to live a life of normal length, being able to live in good health in the absence of physical or psychological harm and the ability to actively take part in society, amongst others (Nussbaum 2003). Certain extra-consequentialist attributes of healthcare safety fit into this conceptual framework. Preventing

medical errors reduces dreadfulness, which improves the ability to live without fear and anxiety (point 5 of Nussbaum's list), as well as allowing a patient better control over their treatment outcomes (point 10) and making them feel less isolated (point 7). This suggests that the capability approach has the potential to be an appropriate theoretical framework for economic evaluation of intervention to reduce errors in healthcare.

Another alternative option is a move towards cost-effectiveness analysis which does not involve measuring patient utility, which would avoid the issue of measuring patient and public preferences. It has been suggested that economic research in medication safety should move away from preference-based outcomes in economic evaluation towards objective measures such as mortality, readmission rates or length of stay. The advantages of using such outcomes is that they can be measured much more precisely than incidence of ADEs and effect of ADEs on patient utility and may provide a less ambiguous marker of the value of medication safety initiatives. Studies which use objective outcomes may be understood and adopted more readily by policy-makers and practicing physicians alike, easing translation of research findings unto practice. A GP may be more likely to appreciate the benefits of monitoring if the findings are expressed in terms of cost per ADE or preventable admission avoided, rather than incremental cost per QALY.

Chapter 8

Discussion

8.1 Introduction to the chapter

A number of medications routinely prescribed in the UK primary care setting are listed in guidelines for regular safety monitoring during ongoing therapy. In many cases, however, guidelines are inconsistent, lack necessary information to guide monitoring strategy and rarely rely on evidence from well-designed observational or randomised studies to justify recommendations for regular monitoring. Whilst previous studies of monitoring have demonstrated low compliance to monitoring recommendations in the UK and abroad, there is little evidence from quantitative studies regarding the impact of regular monitoring on the risk of ADE or the extent to which medication monitoring may prevent unnecessary health loss and economic cost resulting from ADEs. This thesis aimed to examine the current conduct of regular monitoring of a subset of high-risk medications which are indicated for regular laboratory monitoring in published guidelines in the UK and assess the effectiveness and cost-effectiveness of different frequencies of monitoring long-term therapy. In order to achieve these aims, a retrospective cohort study was conducted using data on prescriptions, laboratory tests and diagnoses from two large databases of UK electronic health records. Time-to-event analyses were conducted to assess the effect of assignment into comparison groups involving different frequency of monitoring during long-term therapy with the study medications. The economic effect of the frequency of thyroid function monitoring in amiodarone therapy was examined using a model-based cost-effectiveness analysis which combined the causal effect of monitoring on the hazard rate of amiodarone-induced ADEs estimated in the time-to-event analysis with evidence on the cost and utility effect of ADEs derived from published studies.

8.2 Summary of findings

Chapter 1 of this thesis highlighted the issue of adverse drug events in healthcare and described monitoring as a tool used within general practice to improve the safety profile of certain high-risk medications in general practice. When used

appropriately, monitoring may be used to identify changes related to an ADE in a timely manner in order to allow corrective action to be taken before it causes severe harm to a patient's health. Monitoring may be conducted before, at the initiation and during ongoing treatment with a medication. The efficacy of a monitoring strategy is conditional on a number of factors, which include the presence of an accurate marker which can identify the development of an ADE in a timely manner with a high signal-to-noise ratio. Monitoring errors occur, but as with any other type of medication error, very few cause actual harm.

Chapter 2 included a review of the literature on ADEs and medication errors in healthcare. Most research in medication safety has been conducted in the hospital setting, with fewer studies investigating the incidence, nature and causal factors of community-acquired ADEs. The review found that ADEs commonly occur in routine practice, with an estimated 0.5% and 25% of hospital admissions caused by ADEs acquired in the community setting.

Several studies identified in the review estimated the direct cost impact of community-acquired ADEs. The cost burden of ADEs from the perspective of the healthcare provider was substantial and ranged from £1167 to £6830 per ADE-related admission (converted to 2014 GBP); the aggregate direct medical cost of ADE-related hospitalisations in the UK was estimated at £656 million when inflated to 2015 prices (Pirmohamed, James et al. 2004). There was much less evidence on the societal costs of ADEs, such as the cost of informal care and lost productivity.

The studies identified in the review were heterogeneous in their study design, methods of identifying ADEs and included medications, which contributed to a large amount of variation in estimates across studies. There was no "gold standard" method for identifying ADEs in the medication safety literature, with studies employing chart review, spontaneous reporting, data mining using the electronic health record or a combination of the above. Another issue highlighted in the review was inconsistency in the terminology used to define medication harm, with multiple terms including "adverse drug event", "adverse drug reaction", "drug complication" used within studies to define events of interest, frequently without a clear explanation of how such events were defined in observed data.

One of the aspects of the studies of ADE identified in the review was their assessment of preventability of drug-related harm. It was found that 36% of preventable ADEs may have been averted with more careful monitoring at the

initiation or during therapy (Gurwitz, Field et al. 2003). Reviews of studies of community-acquired ADEs have found that between 21% and 45% of ADEs-related hospitalisations may be preventable through better monitoring (Howard, Avery et al. 2007, Thomsen, Winterstein et al. 2007). These findings suggest significant scope for reduction of preventable ADEs through monitoring of ongoing therapy and make a strong case for implementing interventions to improve monitoring of high-risk medications in order to improve medication safety. There was a relationship between severity of ADE and its preventability: amongst community-acquired ADEs which were serious enough to require hospital admission, the proportion caused by error was over one in three. This is a strong argument in favour of interventions to reduce medication error in primary care: whilst the proportion of errors which cause actual harm is small, the most severe and life-threatening ADEs are caused by errors and should be prevented.

The review found that most ADEs were caused by a small set of medications which have been in use for many years and whose risks are well known in practice, rather than newly introduced medications or those which have been included in guidelines for potentially inappropriate prescribing. Amongst the drugs most commonly implicated in ADEs in healthcare were cardiovascular medications which are indicated for monitoring during ongoing therapy in UK guidelines: ACE inhibitors, diuretics and amiodarone. This was the main motivation for selecting these medications as the focus of analysis of the effectiveness and cost-effectiveness of monitoring.

A review of published studies of monitoring was undertaken in Chapter 3. It examined the evidence on the compliance to medication monitoring in primary care in the context of current guidelines for the conduct of laboratory tests both at initiation of therapy and during ongoing therapy. In addition, it summarised previous studies which evaluated interventions to improve monitoring.

The evidence from selected studies found that although monitoring of ACE inhibitors, loop diuretics and amiodarone is indicated in guidelines, in practice a large proportion of patients on regular therapy with these medications fail to receive any of the required tests. The compliance with thyroid function monitoring in amiodarone was particularly poor, with the most optimistic estimate of compliance suggesting that over 1 in 3 patients on amiodarone do not have their thyroid function checked at all during established therapy, despite evidence from observational studies that hypothyroidism and thyrotoxicity occur frequently during regular intake of this medication. The reported proportions of patients with

inadequate follow-up monitoring varied greatly cross studies, which may be explained by differences in study design, length of observation, methods used to identify the presence of monitoring as well criteria for identifying continued therapy and periods of discontinuation. Only two studies have previously examined the effect of monitoring of the medications included in the analysis described in this thesis on the probability of ADE, both of which involved a retrospective cohort design (McDowell, Coleman et al. 2010, Raebel, Ross et al. 2010). The findings from these studies did not support regular monitoring of U&E in antihypertensive medications (including ACEI and diuretics) in all patients, with evidence from one study that it may be effective at reducing the risk of hyperkalaemia in a subgroup of patients with pre-existing diabetes and CKD (Raebel, Ross et al. 2010). The review found no studies which looked into the effect of thyroid function monitoring during amiodarone therapy on the risk of thyroid dysfunction.

The review identified studies which evaluated interventions to improve compliance to monitoring guidelines. Observational studies generally showed a larger effect of CDSS systems on monitoring completion, when compared to RCTs. The effect was larger in studies which combined the electronic system technology with professional support from pharmacists, suggesting a synergy arising from technology and the expert knowledge of the pharmacist to improve the practice of monitoring.

There was insufficient economic evidence to justify regular monitoring of ACEI, loop diuretic or amiodarone therapy. Cost-effectiveness analyses alongside RCTs assessing interventions to improve monitoring found that monitoring of urea & electrolytes during ACE inhibitor therapy is only effective in patients with CKD (Smith, Raebel et al. 2011). Extending the monitoring intervention to all patients on the drug was not likely to be cost-effective. Only one study examined the cost-effectiveness of an intervention to improve thyroid function monitoring during amiodarone therapy and found it to be highly likely to be cost-effective at standard ceiling WTP per QALY (Elliott, Putman et al. 2014). However, no evidence of the cost and utility effect of monitoring of these medications at different frequencies outside the context of an intervention was found. Given the poor levels of compliance to guidelines for regular monitoring of ACEI, loop diuretics and, in particular, amiodarone in UK primary care, there was a strong rationale for carrying out a study to examine the economic effect of alternative strategies for regular monitoring of these medications in order to inform guidelines for best practice.

The review of literature on medication monitoring outlined in Chapter 3 highlighted several aspects of study design which need to be considered in studies of monitoring. These include controlling for the confounding effect of non-random assignment to monitoring when using observational data, correct classification of therapy discontinuation and monitoring exposure from data retrospectively collected data in order to adequately observe prevalent users of the drug and have a sufficiently long follow-up period in order to observe frequency of testing. These methodological aspects were incorporated into the design of the cohort analysis described in this thesis and are more outlined in detail in Chapter 4.

Chapter 4 outlined the aims of the time-to-event analysis of monitoring and provided a rationale for the use of CPRD and HES to populate the analyses. The advantages of using large databases of EHRs for this programme of research included access to a large patient sample representative of the UK population, follow-up of sufficient length to observe both monitoring frequency and ADE outcomes over multiple calendar years. Previous studies have validated CPRD and HES use in the UK through comparisons of commonly recorded diagnoses using chart reviews. The depth of information contained on CPRD allowed the investigator to construct several patient covariates in order to control for known confounding factors. These included chronic disease and polypharmacy scores, in addition to demographic factors and presence of specific co-prescribed medications and medical conditions which may influence monitoring decisions. The use of CPRD and HES entailed certain limitations for the research described in this thesis. It was not possible to identify monitoring conducted in the hospital setting due to unavailability of outpatient and laboratory test records in HES. Around one-third of CPRD practices did not have a link to hospital data, which limited the sample size in the analyses which included ADE-related hospitalisation as an outcome.

Chapter 5 outlined how EHRs from CPRD and HES were used to estimate the effect of monitoring frequency on the risk of specific ADEs caused by the medications included in the study. The merits of alternative methods of estimating causal effect in epidemiological research using observational data with non-random assignment were discussed. Standard regression and propensity score methods were identified as valid tools to control confounding in observational studies. The propensity score, defined as the probability of treatment assignment conditional on observed baseline covariates, has been used previously to investigate the effect of monitoring using observational data and

was found to be a valid method of isolating the average treatment effect in cases where treatment assignment may be influenced by observed patient characteristics (Raebel, Ross et al. 2010). Propensity score weights were the preferred method in the context of this study: it allowed for more than two exposure levels (3 levels of monitoring frequency) whilst balancing patient groups when baseline patient characteristics affect both the chosen monitoring regimen and the rate of ADE. Ordered probit regression was used to estimate propensity scores for the three different exposures, given the values of baseline covariates derived from CPRD. Marginal mean weighting through stratification (MMWS) was chosen as the most appropriate method to attach analytic weights to each patient conditional on their treatment assignment and propensity score quintile. This method was considered superior to other available tools, such as matching, covariate adjustment and inverse probability of treatment weighting (IPTW). It ensured that the comparison groups were balanced in terms of their baseline characteristics, separated the design and analysis stages, which allowed the calibration of each step in the model and was less likely to produce outliers due to extreme values in the estimated propensity score distributions.

Time-to-event analysis was chosen as the appropriate method to estimate the effect of assigned monitoring frequency on the risk of specific ADEs linked to the study medications. This method was chosen due to the desire to measure the effect of monitoring in terms of the instantaneous hazard rate, which was a suitable approach when dealing with different lengths of observation and non-informative censoring across individuals in the cohort. Weibull regression was used in order to explicitly estimate the hazard function of ADE in each individual analysis, whilst applying analytic weights based on propensity scores. The shape and scale parameters of the hazard function estimated using Weibull regression may be used to estimate transition probabilities for a decision-analytic model, which were used in the cost-effectiveness analysis in Chapter 7, which was another rationale of using this type of time-to-event regression analysis.

Chapter 6 summarised the results of the descriptive analysis of monitoring and time-to-event analysis of the effect of monitoring on the risk of ADEs in the case of each of the medications investigated in this programme of research. The descriptive analysis of monitoring using CPRD records found that many patients were not monitored at the recommended frequency listed in guidelines. Compliance to monitoring guidelines was lowest in the cohort of patients on long-term amiodarone therapy, where 34% of patients had no record of any monitoring and just 37% had records of thyroid function tests at least once in 6

months. At the design stage of the time-to-event model, it was found that 99.6% of patients in the ACEI cohort and 99.9% of patients in the loop diuretics cohort were comparable in terms of their probability of being assigned into each monitoring regimen, using previously validated design parameters (Cochran and Rubin 1973). In the amiodarone cohort, there was a smaller overlap of propensity scores across comparison groups (86.6% and 71.4% of the cohort was comparable in the analyses of AIH/AIT and MACE, respectively). The observed characteristics of patients monitored below recommended frequency (less frequently than once in 6 months) and patients monitored above this frequency differed to the point where a substantial part of the cohort was ineligible for analysis due to incomparability. The exclusion of these patients was necessary in order to ensure that all three comparison groups could be adequately compared and the propensity score weighting system served its purpose in controlling for selection bias. Given that the amiodarone cohort was far smaller than the other two medication cohorts, particularly in HES-linked practices (5760 patients), a further reduction in sample size impacted on the potential for the analysis to identify the clinical effect of monitoring. However, as demonstrated in Chapter 6, the analysis of amiodarone monitoring still demonstrated a statistically significant effect of monitoring in all three test-ADE pairs.

The time-to-event analyses of the effect of monitoring frequency on the rate of ADEs (refer to Chapter 6 for details) identified statistically significant findings in several drug-test categories. The analysis found a protective effect of regular monitoring of thyroid function in amiodarone in terms of all three ADEs studied: hypothyroidism, thyrotoxicity and emergency admission for MACE. The positive effect in the MACE model suggests that it may also prevent cardiovascular morbidity linked to amiodarone-induced AIT. This is a strong argument in favour of interventions to improve compliance to TFT monitoring in amiodarone as MACE has a major impact on patient health, utility and cost of treatment. However, this result should be interpreted with caution. The causal effect of increased TFT monitoring and episodes of MACE is more difficult to establish, as other types of monitoring and treatment which accompanied TFT monitoring were likely to have a competing effect on cardiovascular health. Further evidence from observational and randomised studies are needed in order to confirm the presence of this effect.

There was a beneficial effect of monitoring of U&Es more frequently than recommended in loop diuretic therapy: it reduced the rate of electrolyte disorders treated both in primary and secondary care. However, the true effect of

monitoring frequency was non-linearly related to the risk of ADE as those monitored less frequently also had a lower rate of ADEs compared to those monitored according to guidelines. Further work in this area is needed in order to explain the relationship between monitoring frequency and electrolyte ADEs induced by loop diuretics.

Monitoring below recommended frequency was found to be beneficial in certain cases: this includes U&E monitoring in ACEI and loop diuretics (in terms of hyperkalaemia treated in primary and secondary care and diuretic-induced hypokalaemia/hyonatraemia treated in general practice). The analysis did not support the current policy of encouraging regular monitoring of ACEI and loop diuretics.

Chapter 6 additionally explored a number of methodical issues, such as the exclusion of patients with no recorded monitoring and identification of exposure and ADE outcomes in the same time interval. An additional issue was the choice of propensity score stratification to control confounding over alternative methods such as standard regression and inverse probability weighting. The implication of these decisions on the study was explored using sensitivity analyses.

The model-based cost-effectiveness analysis described in Chapter 7 hypothesised that the benefits of regular TFT monitoring in amiodarone therapy reported in Chapter 6 translated into economic benefits to the NHS. The cost-effectiveness analysis demonstrated that a strategy of monitoring in line with current guidelines which recommend TFT once in 6 month was dominant against alternative strategies which involve monitoring below or above this reference interval. One-way sensitivity analyses demonstrated that varying the estimates of key transition probability, cost and utility parameters had an effect on the net monetary benefit produced by each alternative strategy, but did not result in a change of optimal strategy. Even under conservative scenarios involving reduced protective effect of monitoring, likelihood of adverse outcome (e.g. MACE), cost of treatment and utility losses resulting from thyroid dysfunction, recommended frequency monitoring remained the strategy with the highest net monetary benefit. A probabilistic sensitivity analysis which looked at the combined effect of uncertainty in all model inputs still maintained that recommended frequency of monitoring was most likely to be the cost-effective option, with a 97% probability of producing cost savings and a minimum 90% probability of being the cost effective option, assuming WTP ceiling ranging 0-£100,000/QALY. The analysis

supported current guidelines which recommend TFT monitoring once in 6 months during long-term amiodarone therapy.

8.3 Strengths and limitations of the study

One of the major challenges in this programme of research involved handling systematic differences across comparison groups in the time-to-event analysis. Monitoring is designed to identify and mitigate risks of drug-related harm in patients and naturally GPs would wish to employ it in such a way as to maximise the gains in medication safety with respect to resource and time constraints. Patients who are deemed at higher risk of ADE are thus more likely to be monitored intensively. Therefore, it is reasonable to assume that systematic differences across groups of patients assigned to different monitoring regimens may exist in an observational study. The findings from the analyses in Chapter 6 reaffirm this assumption: the ordered probit models found a relationship between the monitoring frequency assigned to patients and several patient characteristics, including age, sex, multimorbidity, polypharmacy and the presence of certain medical conditions and co-prescribed medications. These findings, which are supported by evidence from previously conducted observational studies of monitoring practice, introduced confounding by indication into the observational analysis of monitoring. This represented a major threat to the internal validity of the study and needed to be addressed.

This study used propensity scores, which are estimated probabilities of assignment of a particular participant to a specific treatment conditional on the values of their baseline characteristics (Rosenbaum 1983). Propensity scores were estimated using a probit regression model which included covariates which were assumed to influence treatment assignment based on evidence from literature. Propensity scores were turned into adjusting weights using mean marginal weighting through stratification (MMWS). This method offered a number of advantages: it ensured that the comparison groups were comparable by selecting patients with propensity scores within a common range (Hong 2012), allowed the investigator to separate the design of the model (estimation of propensity scores, stratification) from the estimation process (fitting Weibull regression to the time-to-event data) ensured that each step could be correctly specified and calibrated. In addition, the weights computed using MMWS are representative of the stratum of the propensity score distribution, and are thus less likely to be affected by outliers in the set of estimated propensity scores and misspecification of the probit model which is used to estimate them. The

performance of the adjustment tool was tested by creating pseudo-datasets and checking covariate balance against the original dataset.

Several other statistical tools may be applied to separate the treatment effect from other factors which potentially affect both the decision to carry out a particular treatment and the final outcome of interest, such as multivariate regression and stratification. Propensity scores may be used in a number of ways to achieve a balanced sample: these include matching, covariate adjustment and inverse probability weighting. A sensitivity analysis in Chapter 6 was conducted in order to compare the performance of MMWS to IPTW and standard regression in the ACEI dataset. Whilst MMWS and IPTW produced similar estimates of the effect of monitoring on the rate of ADE, the results in the standard Weibull model diverged from the base case analysis. In particular, the effect in the less frequent monitoring group was opposite to the one found in the original analysis. A major limitation of the use of standard regression to examine the effect of follow-up monitoring is the assumption of independence of the exposure variable from the set of baseline covariates included in the model, which may produce biased estimates of the treatment effect if this assumption doesn't hold. The choice of monitoring strategy is clearly determined by the value of baseline characteristics, as demonstrated in Chapter 6 of this thesis, and supported by evidence from previous studies of monitoring. Methods which adjust for the initial probability of assignment conditional on the value of baseline characteristics are more appropriate in this situation. Nonetheless, the sensitivity analysis demonstrated that the results of the time-to-event analysis were highly sensitive to the choice of design. This is a major source of structural uncertainty in this study which may impact on its internal validity. Further methodological work is required in order to establish the appropriate method for controlling for non-random assignment in observational studies in monitoring to ensure the removal of selection bias.

In order to ensure comparability across comparison groups, all patients included in the analysis were required to have a similar probability of being assigned to each of the three regimens. This was done in order to enable each patient to be compared to individuals assigned to the other two strategies (common support). This resulted in the exclusion of patients who don't fall under this criteria and loss of sample size. The loss was minor in ACEI and loop diuretics cohorts (0.4% and 0.1%, respectively). However, differences in characteristics of patients assigned to alternative monitoring regimens was much wider in the amiodarone cohort, resulting in the loss of 28.6% of patients. Given that amiodarone cohort had the smallest starting sample size, this was a threat to external validity of the results

in this cohort. However, the potential loss of internal validity due to selection bias outweighed the methodological limitations arising from loss of sample size. Moreover, the resulting analysis identified a statistically significant effect of monitoring in all three outcomes, meaning that the study was adequately powered to identify differences in the rate of ADE across comparison groups. Full details of the exclusion in each cohort are included in Section 6.7.

A limitation of the observational studies of monitoring using records from CPRD was the inability to adequately observe ADEs in patients who had no record of monitoring during follow-up. It is a feature of the data source: absence of conducted laboratory tests precluded the presence of records of laboratory measurements and thus the condition of the patient was unobservable. Given that two-thirds of cases of hyperkalaemia and AKI were identified from abnormal test results, the inclusion of this group in the analysis was likely to underestimate the actual number of ADEs and produce misleading results. The missing laboratory values could not be easily replicated using standard methods for handling missing data, such as multiple imputation, as there is reason to presume that the values are not missing at random, but rather both the absence of monitoring and the risk of ADE were affected by other patient or treatment factors. As a result, only patients with at least one relevant monitoring test during the observed period were included.

An exploratory analysis in Chapter 6 examined this issue in the context of the ACEI cohort. Upon examination of the data, it was found that patients with no recorded monitoring were, on average, older, with more comorbidities and cases of CKD and diabetes, compared to the group who were monitored below recommended frequency. This may suggest that patients in this group may have been monitored in secondary care, as patients with this profile are usually indicated for monitoring. It was not possible to verify this as hospital laboratory test records were not available to this study. The presence of diagnoses of hyperkalaemia and AKI in CPRD records for these patients may also point to the possibility that monitoring (and identification of abnormal tests) may have been conducted in a setting other than primary care. A sensitivity analysis was conducted in order to gauge the effect of inclusion of patients with no observed monitoring on the results of the time-to-event analysis in the ACEI cohort. It found a statistically significant positive effect of monitoring on the hazard rates of both hyperkalaemia and AKI. The absence of hospital laboratory test records was likely to introduce bias into the analysis. The potential loss of generalisability and validity arising from the exclusion of this group was outweighed by the potential

threat to internal validity arising from the difficulties involved in the correct specification of exposure and outcomes in this group.

The exclusion of patients with no recorded monitoring did not impact on the ability of the study to answer the research questions set out in Chapter 1. The aim of the study was to investigate the effect of monitoring in line with guidelines, in comparison to monitoring less or more frequently than this benchmark interval. Data on patients who have had records of laboratory tests during the observed period was sufficient to achieve this aim. Nonetheless, this somewhat limited the comparability of this study to previous work in the area and potential impact of the study in the context of policy in medication monitoring, as it was not possible to include the absence of monitoring as an alternative strategy.

A major methodological issue examined in Chapter 6 was linked to identification of monitoring exposure and ADE outcomes in the same follow-up period. An increased rate of diagnoses of ADE may have been driven by intensive monitoring, which may skew the results of the time-to-event analysis in favour of less frequent monitoring strategies. Another issue was the potential re-classification into less frequent testing in patients who have not experienced an ADE during the follow-up period. Upon examination of the data in the ACEI cohort, a positive relationship was found between monitoring frequency and both hyperkalaemia and AKI. This effect in the unadjusted data was unsurprising, given that patients with pre-existing risk factors for ADE were more likely to both undergo intensive monitoring and experience ADEs. The unadjusted risk ratios were consistent with findings from previous studies which linked monitoring and the risk of ADEs linked to ACEI. Whilst the effect in the adjusted data was surprising, it is not a true reflection of the relative rates of ADE as relative risk does not take into account the length of follow-up, which is considerably longer in patients monitored less frequently (due to reduced probability of discontinuation and death). Nonetheless, it was not possible to rule out that some of the diagnosed ADEs may have been driven by increased testing alone. Despite the application of appropriate tools designed to control the confounding effect of non-random treatment assignment, this effect may have overestimated the hazard ratios in the more frequent monitoring group.

Patients who were monitored in the first period and did not have a record of ADE were less likely to be monitored regularly in subsequent periods, which suggested the presence of a "reassurance" effect. Although this issue was relevant to a

subset of the cohort, it could have potentially underestimated the protective effect of monitoring in the time-to-event analysis.

The sensitivity analysis in Section 6.13 examined an alternative method which used separate time intervals to identify exposure and outcomes in the time-to-event analysis. The alternative analysis found that the rate of both hyperkalaemia and AKI was significantly higher in the group of patients with less frequent monitoring, which wasn't observed in the original analysis. However, the generalizability of this result to the study population was low. A large proportion of the original cohort was excluded due to early discontinuation from therapy, death or censoring due to ADE. As a result, these patients exited the dataset before they reached the new follow-up period. This was therefore not an appropriate method of answering the research questions set in this programme of research.

Another issue linked to the use of routinely collected EHR data was insufficient information on the indications for conducted monitoring and the rationale behind the decisions of GPs to assign individual patients to a particular monitoring regimen. It was impossible to make accurate judgment regarding the GP's decisions based on information from EHRs alone. The researcher may observe a missing test but it is not possible to determine with certainty that it constitutes an error as a number of factors may have driven it which are unobservable, such as the attitude of the prescribing GP towards the risk of ADE in individual patients, their opinion of the merits of regular monitoring, workload and time constraints; on most occasions it was impossible to determine whether missed monitoring was due to non-compliance with monitoring guidelines by the GP or non-adherence to ordered tests by the patient. In addition, the CPRD test data did not contain information on the indication for the test and an assumption was made that the conduct of relevant tests was linked to the monitoring regimen, rather than an unrelated reason. These are issues inherent to EHR data extracted from large databases and it may be necessary to access patient charts in order to accurately investigate the nature of individual cases of monitoring error and ADE.

The reliability of an observational study which employs routinely collected data is heavily dependent on the quality of the data source from which it obtained patient-level records. This study used primary care prescriptions, laboratory test and diagnosis data from CPRD, which undergoes continuous quality assessment and validation (Jick, Kaye et al. 2003). The quality of data recorded on CPRD has previously been assessed by comparing against chart reviews and GP

questionnaires and achieved a positive predictive value between 80% and 90% for common chronic conditions, including diabetes, inflammatory bowel disease, psychosis and dementia (Khan, Harrison et al. 2010). The recording of cardiovascular events, such as atrial fibrillation and cardiac arrhythmia has been relatively poorer, meaning that some events relevant to this study may have been missed. In order to mitigate this issue, linkage to HES database was used to identify severe events which resulted in hospital admission. The HES database has been validated by cross-checking hospital charts for records of chronic conditions, including cancer and vascular disease (Aylin, Lees et al. 2007, Garout, Tilney et al. 2008). Completeness of electronic records in hospitals has seen a substantial improvement in recent years, with improvement in accurate recording of diagnoses (96%) and episodes (86%) (Burns, Bottle et al. 2011), which may be a positive effect of the introduction of Payment by Results and general improvements in EHR technology. Although this study used high-quality, validated sources of EHR to populate analysis of monitoring, misrecorded or missing episodes were still a possibility, which is an inevitable limitation of EHRs. The investigator had no access to patient charts and validating though cross-checking with other sources of data was not possible.

Improvements in the quality of recording of EHRs may have acted as a double edged sword. Whilst improved completeness of records and consistency in clinical coding helped to reduce the number of missing episodes and diagnoses in the current analysis, it may have artificially driven up estimates of monitoring completion and recorded ADEs during the study period. Improved electronic recording has previously been identified as a cause of a rise in reported incidence of ADE-related admissions (Wu, Jen et al. 2010). The current analysis identified an improvement in TFT monitoring of amiodarone from the year 2006 and onwards, which may have been partly driven by improved recording of laboratory tests on CPRD. It is not clear how much of the effect was driven by improved data quality as no study has hitherto investigated the patterns in recording of laboratory test records on CPRD.

A limitation of economic analysis was the short horizon of the model, which was limited to 5 years due to the uncertainties involved in the extrapolation of the hazard function based on original follow-up of 2.5 years. Observational studies which observe monitoring and ADEs over a lengthier time period, in addition to further advancement of methods for extrapolation of hazard rates are both required in order to build a model with a lifetime horizon. The cohort simulation design of the model did not allow for a subgroup analysis. Additional models

which will explore the cost-effectiveness of monitoring in high-risk groups are planned as part of future work to advance the research outlined in this thesis.

Whilst time-to-event models were fitted in all three medication cohorts, cost-effectiveness was only assessed in the case of TFT monitoring of amiodarone. The choice to focus the economic work on this particular cohort was motivated by several factors: the literature review identified only one previous cost-effectiveness analysis of monitoring and none based on observed data on monitoring and ADEs, which represented a major gap in the evidence base; compliance to monitoring guidelines was poor in the amiodarone cohort, a fact which has been recognised in previous studies but the economic effect of this disparity between recommended and actual practice has never been formally investigated; a statistically significant improvement in the hazard rate of ADE linked to regular monitoring was only observed in the case of amiodarone. As the hypothesised economic benefits of monitoring are linked to reduction in the cost of treatment of ADE and loss of patient utility arising from reduced risk of ADE, absence of clinical benefit of regular monitoring in the time-to-event analysis meant that monitoring ACEI and loop diuretics at recommended frequency was unlikely to be cost-effective.

One of the strengths of the time-to-event analysis of the effect of monitoring described in Chapters 5 and 6 of this thesis was the length of follow-up of exposure and outcomes, which was 48 months in the ACEI and loop diuretic cohorts and 33 months in the amiodarone cohort. This allowed the investigator to observe recorded monitoring tests, diagnoses of ADEs and abnormal test results over a longer time period than previous studies of monitoring, which typically extended for 6-15 months (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, Matheny, Sequist et al. 2008, Johnson, Canty et al. 2010, McDowell, Coleman et al. 2010, Avery, Rodgers et al. 2012, Bundy, Marsteller et al. 2012). This made it possible to classify patients according to the frequency of monitoring tests. This design was combined with strong methods to determine continuity of therapy and monitoring exposure designed to minimise misclassification bias. To address the first point, the computerised prescription history of each included patient was checked to identify discontinuation of therapy. Absence of recorded prescriptions between days 90 and 270 meant that the patient was discontinued shortly after initiation and was thus removed from the dataset. Gaps in prescription for an entire exposure period were taken as evidence of discontinuation of therapy (refer to Chapter 5 for precise definitions). Exposure was only observed during periods when the patient was on active therapy. To

identify the monitoring assignment for each patient, computerised records of laboratory tests were searched using a previously compiled and validated list of Read codes to identify relevant laboratory tests. Patients were classified into groups according to the number of tests carried out per exposure period. Exposure was identified at different time points during the observed period for each patient. There was no "gold standard" for identifying therapy persistence and monitoring exposure in previously published monitoring studies, with each author making individual assumptions regarding what can be regarded continuous therapy or appropriate length of observation of regular monitoring. The current analysis incorporated what were deemed the most relevant methods used in previous studies, whilst taking into account the available information and limitations of the chosen data source.

This study accounted for multimorbidity and polypharmacy, which are two important determining factors in monitoring, using specially constructed measures from CPRD data. Both measures needed to be created from raw data as neither measure of burden of chronic disease nor a measure of co-prescription was immediately available to the investigator. An effort was made to include all possible chronic conditions using a systematic search of diagnostic Read codes and prescription of medications associated with 40 different chronic conditions from CPRD data, based on a chronic disease score which has been validated in a UK population (Barnett, Mercer et al. 2012). A polypharmacy measure was constructed from prescription records by classifying CPRD product codes by BNF chapter. In the ACEI and loop diuretic cohorts, additional covariates were created to control for concomitant treatment with drugs which are equally indicated for regular monitoring of U&E in guidelines and diagnoses which are likely to result in more frequent monitoring, including CKD and diabetes. As with any retrospectively collected data from EHRs, there was a risk that the resulting covariate measures could be inaccurate due to missing or misrecorded diagnoses and prescriptions. Although the CPRD database has been previously validated for as a tool for pharmacoepidemiological research, there was no unerring method for ensuring the completeness of records, such as cross-checking with paper-based patient records, as the initial data extraction was the only point of access to patient-level data.

The cost-effectiveness analysis of amiodarone monitoring outlined in Chapter 7 was the first economic analysis of frequency of regular amiodarone monitoring based on data observed from actual practice. It employed a modelling method which incorporated the results of three separate time-to-event analyses in order

to populate the initial transition probabilities in the model, which reduced the reliance other sources of data, such as incidence rates derived from literature, which may misrepresent the population and setting of the current analysis. Extensive sensitivity analyses were conducted in order to gauge the robustness of the model to changes in individual inputs using one-way sensitivity analyses, as well as to combined uncertainty of all parameters in a probabilistic sensitivity analysis.

8.4 How this study fits into and contributes to the existing evidence base

This programme of research contributed to the evidence base on the practice of follow-up monitoring of three medications which pose a risk of ADE in the context of UK primary care: ACEI, loop diuretics and amiodarone. The findings from Chapter 6 of subpar compliance to published guidelines on follow-up monitoring in all three medications fits into the existing evidence base of studies which have described similar discrepancies in monitoring practice in other countries. Moreover, compliance to guidelines for follow-up monitoring was considerably poorer in the amiodarone cohort compared to ACEI or loop diuretics, with a majority of patients deemed to be inadequately monitored (62%) and over a third of patients in the cohort had no recorded TFT tests during the course of therapy. The poor conduct of amiodarone monitoring relative to other medications has also been observed in previous research (Raebel, Chester et al. 2006, Tjia, Fischer et al. 2011).

The time-to-event analysis may be compared with previous studies of monitoring in terms of particular design aspects, which were adapted to the analysis following assessment of appropriateness of design decisions described in Chapter 3. This study identified monitoring exposure in ACEI and loop diuretics cohort based on presence of either serum creatinine, potassium or sodium, with a requirement that at least two tests are recorded on the same day, which is a similar approach to Tjia et al (Tjia, Fischer et al. 2011). Other authors used stricter definitions, which may underestimate exposure as laboratory tests are not necessarily accurately recorded in EHRs. In the amiodarone cohort, TSH was used to determine presence of TFT testing as triidothyronine and free thyroxine may not, by itself, be used to diagnose thyroid dysfunction (British Thyroid Association 2006, Johnson, Canty et al. 2010, Tjia, Fischer et al. 2011). Similarly to some previous studies of monitoring, this study determined exposure to monitoring and ADE outcomes through retrospective observation of patients for a set time period

following initiation of therapy (Hurley, Roberts et al. 2005, Raebel, McClure et al. 2007, McDowell, Coleman et al. 2010).

The time-to-event analysis found that monitoring U&Es during ACE inhibitor therapy below recommended frequency did not result in reduction in the hazard rate of hyperkalaemia or AKI. Conversely, it found that the rate of ADE was lower in patients monitored less frequently. A previous study which employed a retrospective cohort design to estimate the effect of regular monitoring of U&E on the risk of ACEI-induced hyperkalaemia in patients with diabetes found that conducting monitoring reduced the risk of ADE by half (RR 0.5, C.I. 0.37,0.66) (Raebel, Ross et al. 2010). The effect of monitoring was larger in patients with CKD (RR 0.29), but not in patients without CKD and <50 years of age. However, the results of this study are not directly comparable to the findings from this programme of research as it examined patients with pre-existing diabetes. A UK-based study which examined the general population of patients on long-term antihypertensive therapy found that regular monitoring of U&E in ACEI and diuretics was associated with an increased risk of hospital admission (H.R. 1.37,1.21,1.55) (McDowell, Coleman et al. 2010). However, McDowell and colleagues have looked at all admissions to hospital, including those unrelated to the drug, making it difficult to draw comparisons to the studies described in this thesis.

The comparison of the findings from the economic analysis of amiodarone described in Chapter 7 of this thesis to existing evidence is limited by that fact that only one previous study assessed the cost-effectiveness of regular thyroid function monitoring of amiodarone. An RCT-based economic evaluation of a pharmacist-led IT intervention to improve amiodarone monitoring in English general practices found that it was highly likely to be cost-effective (ICER: £515 per QALY, 100% probability of cost-effectiveness at ceiling WTP of £20,000/QALY). This study differed from the current analysis described in this thesis in terms of its study aim, which was to investigate the cost-effectiveness of an intervention to improve monitoring, rather than alternative strategies of monitoring in standard practice. The main parameter which was used in the model was reduction in the odds of a monitoring error (adjusted O.R. 0.57) derived from a previously conducted RCT. This was used as an assumption that improved monitoring increases the chance of early treatment of hyperkalaemia, which was used as a basis for expected reduction in cost and improvement in QALY generation over the horizon of the model. Thus, the model did not estimate the direct economic effect of monitoring, but rather of improvements in early

detection of ADE through improved compliance to guidelines as a result of the intervention. Due to differences in the study aim and methods used to conduct the economic analysis, it is difficult to make direct comparisons, although both studies hypothesised that improvements in regular monitoring would lead to lower cost of treatment of ADE and higher number of QALYs generated during the model horizon.

The key difference of this study to previous investigations, and one of its major strengths, is the length of follow-up which included multiple cycles of monitoring, which made it possible to ascertain the frequency of monitoring and its effect on ADE outcomes. Previously conducted studies either examined completion of monitoring as an endpoint, or the effect or presence of monitoring over a short period, typically 12-15 months. Therefore, this study added much-needed evidence on the effect of monitoring of ACEI, loop diuretics and amiodarone beyond the first year of therapy, and investigated the effect of monitoring at different frequencies in real practice across a long period of follow-up. In addition, this study excluded the first 90 days of the follow-up period to avoid inadvertently including testing at initiation of therapy and running the risk of overestimating monitoring frequency, which was not taken into account by the studies identified in the literature review described earlier in this thesis. This ensured that the study sample included only those patients who were indicated for follow-up monitoring.

To the best knowledge of the investigator, this was the first study which examined the cost-effectiveness of different strategies of follow-up monitoring of amiodarone compared to standard practice. The findings from this analysis may be used in future studies which evaluate interventions to improve amiodarone monitoring. The evidence produced in this thesis will be useful to guideline development groups to substantiate recommendations for improvement in the practice of monitoring of this medication in primary care.

This was the first study to employ routinely collected data to populate a model-based cost-effectiveness analysis of monitoring. A time-to-event analyses in the form of Weibull models were fitted using retrospectively collected data from primary care EHRs to directly estimate the hazard rate of specific ADEs, and subsequently to estimate probabilities of occurrence of ADE in patients assigned to different frequency of monitoring, whilst controlling for the effect of confounding variables using propensity score weights. The findings from this study are therefore more robust than previous studies which used literature and expert opinion to inform the association between monitoring and ADEs.

8.5 Implications for policy

The time-to-event analysis in Chapter 6 did not find a protective effect of regular monitoring of U&E in ACEI and loop diuretic therapy. The analyses in the ACEI cohort found that patients who received U&Es tests less frequently than once in 12 months were significantly less likely to be treated for hyperkalaemia both in primary care and admitted hospital care and did not have a reduced rate of AKI. The hazard rate for AKI and hyperkalaemia was higher and remained unchanged, respectively, in patients monitored above recommended frequency. Similarly, there was no strong evidence in support of regular U&E monitoring in loop diuretic therapy. Although a protective effect of frequent monitoring was identified, patients monitored less frequently than recommended had a lower rate of hypokalaemia and hyponatraemia compared to those with recommended monitoring. The hazard rate of fall-related injuries linked to loop diuretics remained unchanged across comparison groups. There was thus no evidence from this study to support the need for regular monitoring of U&Es during long-term ACEI or loop diuretic therapy. Efforts should be geared towards symptom-based monitoring and identification of patients who are at an elevated risk of ADE, such as pre-existing or worsening renal insufficiency, diabetes or heart failure, which has been shown to be both clinically and cost-effective in previous investigations (Raebel, Ross et al. 2010).

Current NICE guidelines do not mention explicitly how often U&E should be checked when ACEI or loop diuretics are used in the treatment of hypertension. However, in the case of heart failure management, periodic monitoring of the patient's condition is recommended irrespective of the drug regimen prescribed, which would include an assessment of renal function and potassium level. Local guidelines, based on national guidelines as well as information from a range of other sources, recommend monitoring once in 12 months. Given that no tangible evidence exists in support of regular monitoring in all patients, the author encourages the development of unified, evidence based guidelines on "best practice" of monitoring in primary care. These guidelines must recognise differences in the needs of patient subgroups, especially patients with known risk factors for hyperkalaemia and AKI, such as prevalent heart failure and renal insufficiency. Specific information regarding the frequency of monitoring for each subgroup, thresholds for abnormal tests and step-by-step instructions on the optimal response to an abnormal test result are needed.

The time-to-event analysis of regular amiodarone TFT monitoring found that patients monitored below recommended frequency of once in 6 months were significantly more likely to suffer AIH, AIT and MACE. In addition, the economic analysis found that regular monitoring of TFT in amiodarone was highly likely to be cost-effective at the standard ceiling WTP level set by NICE. Given these results, the current study provides support for current guidelines for monitoring of TFT during long-term amiodarone therapy and policy efforts to improve compliance to these guidelines by GPs.

Given that the current compliance to guidelines in amiodarone TFT monitoring is poor (63% were monitored below recommended frequency and 34% had no record of TFT test during therapy), improvements in the practice of monitoring have the potential to produce substantial benefits in terms reduced incidence of thyroid dysfunction, cost savings from an NHS perspective and improved individual utility. Policy efforts in the context of amiodarone monitoring should be focused on implementing interventions which have been previously shown to be both clinically and cost-effective in a primary care setting. A number of interventions have been shown to be effective at improving monitoring based on evidence from RCTs: computer-generated feedback on patients indicated for monitoring combined with pharmacist support (Avery, Rodgers et al. 2012), automated reminders given to GPs by a computerised system linked to the EHR with pharmacist feedback (Raebel, Chester et al. 2006). These findings provide a tangible rationale for the implementation of computerised decision support with professional support by pharmacists as a model of care to improve monitoring of amiodarone and in other situations where regular monitoring is shown to be cost-effective.

Policy makers in charge of developing guidelines need to recognise the root of the problem of poor compliance to recommended monitoring, which is perceived low risk of ADE by GPs and a lack of understanding of the merits of regular monitoring of therapy of certain medications which pose an elevated risk of ADE in patients. As a result, other clinical tasks may take priority under conditions of time and resource constraints in the practice. The opinion of those who argue for implementation of monitoring in all patients prescribed a medication which poses a potential risk of ADE is that all it takes is one lapse or error of omission to cause severe or even fatal consequences to a patient. However, the case for intervening to ensure that all patients are monitored is weaker with the availability of methods of identify patients who are at higher risk of ADE and targeting this

group may yield improvements in medication safety and make the intervention more likely to be cost-effective.

The increased importance of healthcare safety as a policy issue has resulted in a culture of regulation and proliferation of guidelines for the safe use of medications. However, some authors have questioned the approach of attempting to achieve perfect safety in healthcare. Striving to reduce the risk of undesirable effects of medical care to zero may entail excessive use of resources which are better used elsewhere (Warburton 2005). The costs and benefits of increased efforts to prevent errors and ADEs do not behave in a linear fashion; assuming that interventions tackle the most common ADEs which are easily preventable, eventually large amounts of resources would need to be used in order to tackle more difficult issues, which are harder to prevent. Policy-makers need to clearly draw the line and set priorities which are achievable and provide a cost-effective solution with respect to a limited healthcare budget. In the context of medication monitoring, a cost-effectiveness analysis backed by empirical data from real-life practice, such as the one outlined in Chapter 7 of this thesis may be a useful tool to identify interventions which improve medication monitoring without taking away scarce resources from other areas of healthcare.

8.6 Implications for research

Studies of monitoring which aim to inform monitoring guidelines need to consider the effect of monitoring decisions on the risk of ADE. Investigations of rates of inadequate monitoring are useful to describe current practice, but do not explain the effect of disparities between best practice guidelines and actual practice on the rate of occurrence of ADEs. Therefore, they are of limited use for informing evidence-based guidelines

Studies which aim to explain the practice of regular monitoring during therapy and assess its value as a tool to prevent and ameliorate ADEs need to be of adequate length to observe the actual patterns of testing which occur in practice. Observing ongoing monitoring over one cycle may not be informative as it does not explain how monitoring is conducted in the treatment of chronic conditions over the course of several years. If monitoring is conducted in the first year of therapy, it does not mean that it will carry on in subsequent years, which may lead to suboptimal quality of care and ADEs in the future. This is particularly important in cases where the risk of an ADE increases over time, which was the case in amiodarone-related thyroid dysfunction shown in the time-to-event

analysis in this study. The appropriateness of monitoring in such situations needs to be assessed over the whole duration of therapy, or at least for a sufficient length of time of observation to allow appropriate extrapolation to the full length of therapy or over a lifetime. One of the key messages from the time-to-event and cost-effectiveness analysis of amiodarone was that in the absence of evidence from long-term follow up of patients on amiodarone, estimation of the clinical and economic effect of monitoring of long-term therapy is constrained by significant uncertainty in the probability of ADE occurrence.

The programme of research outlined in this thesis highlighted a number of methodological problems involved in observational research of monitoring. In particular, problems linked with identification of exposure and ADE outcomes from retrospectively collected real-life data require further investigation in order to minimise their impact on internal validity. These is significant scope for further academic work in this area in order to develop, validate and compare alternative methods to evaluate monitoring as a preventive strategy using routinely collected EHRs. The development and consistent application of robust methods to control bias will generate much-needed evidence to inform evidence-based practice in medication monitoring.

Research in the area of medication safety needs to give priority to studies which investigate the effect of monitoring on the risk of ADEs. Studies assessing interventions to improve monitoring which use monitoring completion as surrogate outcome are not as informative in the absence of tangible evidence of the benefits of such improvements in terms of final outcomes, including the risk of medication-related harm, mortality, cost of treatment and patient utility. Ultimately, interventions which aim to improve medication safety need to be assessed in terms of their ability to prevent ADEs in order to demonstrate their advantage over current practice.

The cost-effectiveness analysis of amiodarone monitoring outlined in Chapter 7 of this thesis evaluated outcomes in terms of incremental QALYs gained over the horizon of the model, which represents a QALY maximisation approach which is the conventional method of valuing outcomes in economic evaluation of healthcare technology. However, there is sufficient evidence to suggest that this approach may not adequately take into account certain non-health benefits of improvement in medication safety, such as reduction in dreadfulness and isolation, as well as improvements in control over one's treatment outcomes (Delbanco and Bell 2007, Steuten and Buxton 2010). Incorporating these extra-

consequentialist outcomes into an economic evaluation is not straightforward and would require a substantial modification of design. However, ignoring these aspects risks underestimating the true economic value of interventions in this area, making it less likely that new effective innovations to improve monitoring are implemented. Research on the cost-effectiveness of monitoring (and medication safety in general) should be focused in several directions: firstly, the potential for the application of a wider set of outcomes in addition to incremental cost and QALYs in a framework such as multi-criteria decision analysis (MCDA) (Baltussen and Niessen 2006). Secondly, more research is needed to investigate the potential of the capability approach as an alternative framework for economic evaluation of interventions in healthcare safety (Sen 1993). Thirdly, it must be established whether using cost-effectiveness analysis based on process outcomes such as number of deaths, hospitalisations or cost per ADE/error avoided may hold potential in this area due to the difficulty in correctly estimating utility losses linked to ADEs. This approach may be better understood by physicians and policy-makers, benefiting the translation of research findings into practice.

8.7 Concluding remarks

An effective medication monitoring strategy identifies the early onset of ADEs during the course of pharmacotherapy, thus helping to avoid harm to the patient and reduce unnecessary use of healthcare resources. There is a clear disparity between what is perceived to be acceptable practice of regular monitoring of ACEI, loop diuretics and amiodarone, as described in published guidelines, and actual conduct of laboratory testing in UK general practice. Up until now, there have been few studies which have attempted to quantify the effect of missed monitoring tests on the risk of ADE. There is little evidence that the monitoring frequencies listed in guidelines represent a cost-effective use of NHS resources.

The programme of research outlined in this thesis investigated the effect of different frequencies of regular monitoring on the risk of ADEs linked to ACEI, loop diuretics and amiodarone using EHRs from CPRD and HES and the economic effect of monitoring amiodarone. The time-to-event analysis demonstrated that monitoring at recommended frequency may reduce the risk of amiodarone-induced thyroid dysfunction. The economic analysis found that complying with recommendations on amiodarone monitoring is highly likely to be both cost-saving and effective at improving patient utility.

This analysis supports the current policy of encouraging monitoring of thyroid function in long-term amiodarone users once in 6 months. Based on findings of the observational studies of ACEI and loop diuretic monitoring, it is unlikely that monitoring of all patients on long-term prescription of ACEI and loop diuretics represents an effective strategy. Therefore, the analysis did not support the need for regular monitoring of U&E in all patients on long-term ACEI or loop diuretic therapy.

This PhD project has demonstrated that observational research using large databases of EHRs holds significant potential for research in medication monitoring. This required identifying and overcoming a number of complex methodological challenges, including selection bias, issues linked to identifying and interpreting ADEs from routinely collected data, retrospectively dealing with discontinuities in therapy and monitoring exposure, fitting transition probabilities in a decision-analytic model based on extrapolated data. This programme of research contributed to the development of methodology for future investigations of the effectiveness and cost-effectiveness of medication monitoring.

Further investigations of the kind described in this thesis are needed in order to build a body of evidence on the value of regular monitoring in the context of the medications included in this work, as well as other medications listed in guidelines for monitoring. In parallel, more work is required in order to shed light on the true reasons for poor compliance with monitoring guidelines and the attitudes of physicians towards monitoring as a tool to prevent medication-related harm.

References

Ahuja, T. S., D. Freeman, Jr., J. D. Mahnken, M. Agraharkar, M. Siddiqui and A. Memon (2000). "Predictors of the development of hyperkalemia in patients using angiotensin-converting enzyme inhibitors." *Americal Journal of Nephrology*, **20**(4): 268-272.

Akbarov, A., E. Kontopantelis, M. Sperrin, S. J. Stocks, R. Williams, S. Rodgers, A. Avery, I. Buchan and D. M. Ashcroft (2015). "Primary Care Medication Safety Surveillance with Integrated Primary and Secondary Care Electronic Health Records: A Cross-Sectional Study." *Drug Safety*, **38**(7): 671-682.

Albert, S. G., L. E. Alves and E. P. Rose (1987). "Thyroid dysfunction during chronic amiodarone therapy." *Journal of the American College* of *Cardiology*, **9**(1): 175-183.

Alldred, D. P., C. Standage, A. G. Zermansky, B. Jesson, I. Savage, B. D. Franklin, N. Barber and D. K. Raynor (2008). "Development and validation of criteria to identify medication-monitoring errors in care home residents." *International Journal of Pharmacy Practice*, **16**(5): 317-323.

Andrews, L. B., C. Stocking, T. Krizek, L. Gottlieb, C. Krizek, T. Vargish and M. Siegler (1997). "An alternative strategy for studying adverse events in medical care." *Lancet*, **349**(9048): 309-313.

Aronson, J. K. and R. E. Ferner (2005). "Clarification of terminology in drug safety." *Drug Safety*, **28**(10): 851-870.

Audit Commission (2010). "Improving data quality in the NHS. Annual report on the PbR assurance programme." London.

Austin, P. C. (2011). "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behavioral Research,* **46**(3): 399-424.

Avery, A. J., S. Rodgers, J. A. Cantrill, S. Armstrong, K. Cresswell, M. Eden, R. A. Elliott, R. Howard, D. Kendrick, C. J. Morris, R. J. Prescott, G. Swanwick, M. Franklin, K. Putman, M. Boyd and A. Sheikh (2012). "A pharmacist-led information technology intervention for medication errors (PINCER): a multicentre, cluster randomised, controlled trial and cost-effectiveness analysis." *Lancet,* **379**(9823): 1310-1319.

Aylin, P., A. Bottle and A. Majeed (2007). "Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models." *BMJ,* **334**(7602): 1044.

Aylin, P., T. Lees, S. Baker, D. Prytherch and S. Ashley (2007). "Descriptive study comparing routine hospital administrative data with the Vascular Society of Great Britain and Ireland's National Vascular Database." *European Journal of Vascular and Endovascular Surgery*, **33**(4): 461-465; discussion 466.

Baltussen, R. and L. Niessen (2006). "Priority setting of health interventions: the need for multi-criteria decision analysis." *Cost Effectiveness and Resoure Allocation,* **4**: 14.

Barnett, K., S. W. Mercer, M. Norbury, G. Watt, S. Wyke and B. Guthrie (2012). "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study." *Lancet,* **380**(9836): 37-43.

Basaria, S. and D. S. Cooper (2005). "Amiodarone and the thyroid." *American Journal of Medicine,* **118**(7): 706-714.

Bates, D. W. (1996). "Medication errors. How common are they and what can be done to prevent them?" *Drug Safety,* **15**(5): 303-310.

Bates, D. W., D. L. Boyle, M. B. Vander Vliet, J. Schneider and L. Leape (1995). "Relationship between medication errors and adverse drug events." *Journal of General Internal Medicine,* **10**(4): 199-205.

Bates, D. W., D. J. Cullen, N. Laird, L. A. Petersen, S. D. Small, D. Servi, G. Laffel, B. J. Sweitzer, B. F. Shea, R. Hallisey and et al. (1995). "Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group." *JAMA,* **274**(1): 29-34.

Bates, D. W., A. C. O'Neil, D. Boyle, J. Teich, G. M. Chertow, A. L. Komaroff and T. A. Brennan (1994). "Potential identifiability and preventability of adverse events using information systems." *Journal of American Medical Informatics Association,* **1**(5): 404-411.

Bates, D. W., N. Spell, D. J. Cullen, E. Burdick, N. Laird, L. A. Petersen, S. D. Small, B. J. Sweitzer and L. L. Leape (1997). "The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group." *JAMA,* **277**(4): 307-311.

Bates, D. W., J. M. Teich, J. Lee, D. Seger, G. J. Kuperman, N. Ma'Luf, D. Boyle and L. Leape (1999). "The impact of computerized physician order entry on medication error prevention." *Journal of American Medical Informatics Association,* **6**(4): 313-321.

Beach, J. E., G. A. Faich, F. G. Bormel and F. J. Sasinowski (1998). "Black box warnings in prescription drug labeling: results of a survey of 206 drugs." *Food and Drug Law Journal,* **53**(3): 403-411.

Beijer, H. J. and C. J. de Blaey (2002). "Hospitalisations caused by adverse drug reactions (ADR): a meta-analysis of observational studies." *Pharmacy World and Science,* **24**(2): 46-54.

Bell, K. J. L., L. Irwig, J. C. Craig and P. Macaskill (2008). "Use of randomised trials to decide when to monitor response to new treatment." *BMJ,* **336**(7640): 361-365.

Black, A. D., J. Car, C. Pagliari, C. Anandan, K. Cresswell, T. Bokun, B. McKinstry, R. Procter, A. Majeed and A. Sheikh (2011). "The impact of eHealth on the quality and safety of health care: a systematic overview." *PLoS Medicine,* **8**(1): e1000387.

Black, N. (1996). "Why we need observational studies to evaluate the effectiveness of health care." *BMJ*, **312**(7040): 1215-1218.

Boersma, E., D. Poldermans, J. J. Bax, E. W. Steyerberg, I. R. Thomson, J. D. Banga, L. L. van De Ven, H. van Urk, J. R. Roelandt and D. S. Group (2001). "Predictors of cardiac events after major vascular surgery: Role of clinical characteristics, dobutamine echocardiography, and beta-blocker therapy." *JAMA,* **285**(14): 1865-1873.

Bogazzi, F., L. Bartalena, E. Dell'Unto, L. Tomisti, G. Rossi, P. Pepe, M. L. Tanda, L. Grasso, E. Macchia, F. Aghini-Lombardi, A. Pinchera and E. Martino (2007).

"Proportion of type 1 and type 2 amiodarone-induced thyrotoxicosis has changed over a 27-year period in Italy." *Clinical Endocrinology*, **67**(4): 533-537.

Bogazzi, F., L. Tomisti, G. Rossi, E. Dell'Unto, P. Pepe, L. Bartalena and E. Martino (2009). "Glucocorticoids are preferable to thionamides as first-line treatment for amiodarone-induced thyrotoxicosis due to destructive thyroiditis: a matched retrospective cohort study." *Journal of Clinical Endocrinology and Metabolism*, **94**(10): 3757-3762.

Brennan, M. D., J. A. van Heerden and J. A. Carney (1987). "Amiodarone-associated thyrotoxicosis (AAT): experience with surgical management." *Surgery*, **102**(6): 1062-1067.

Brennan, T. A., L. L. Leape, N. M. Laird, L. Hebert, A. R. Localio, A. G. Lawthers, J. P. Newhouse, P. C. Weiler and H. H. Hiatt (1991). "Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I." *New England Journal of Medicine*, **324**(6): 370-376.

Briggs, A. H., M. C. Weinstein, E. A. Fenwick, J. Karnon, M. J. Sculpher, A. D. Paltiel and I.-S. M. G. R. P. T. Force (2012). "Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6." *Medical Decision Making*, **32**(5): 722-732.

Briggs, C. (2004). "Parametric survival and decision models: relating continuous hazards to discrete-time transition probabilities." Paper presented to the Health Economists' Study Group Conference, Glasgow.

British Medical Association & Royal Pharmaceutical Society (2014). "British National Formulary." London: BMJ Group and Pharmaceutical Press.

British National Formulary (online). (2015). "Amiodarone." Retrieved 14/08/2015, from https://www.medicinescomplete.com/mc/bnf/current/PHP906-supraventricular-and-ventricular-arrhythmias.htm#PHP907.

British National Formulary (online). (2015). "Angiotensin-converting enzyme inhibitors." Retrieved 14/08/2015, from https://www.medicinescomplete.com/mc/bnf/current/PHP1124-angiotensin-converting-enzyme-inhibitors.htm?q=ace%20inhibitor&t=search&ss=text&p=1#_hit.

British National Formulary (online). (2015). "Loop diuretics." Retrieved 14/08/2015, from https://www.medicinescomplete.com/mc/bnf/current/PHP805-loop-diuretics.htm?q=loop%20diuretic&t=search&ss=text&p=1#_hit.

British Thyroid Association (2006). "UK guidelines for the use of thyroid function tests."

Budnitz, D. S. and P. M. Layde (2007). "Outpatient drug safety: new steps in an old direction." *Pharmacoepidemiology and Drug Safety*, **16**(2): 160-165.

Budnitz, D. S., M. C. Lovegrove, N. Shehab and C. L. Richards (2011). "Emergency hospitalizations for adverse drug events in older Americans." *New England Journal of Medicine*, **365**(21): 2002-2012.

Budnitz, D. S., D. A. Pollock, K. N. Weidenbach, A. B. Mendelsohn, T. J. Schroeder and J. L. Annest (2006). "National surveillance of emergency department visits for outpatient adverse drug events." *JAMA*, **296**(15): 1858-1866.

Bundy, D. G., J. A. Marsteller, A. W. Wu, L. D. Engineer, S. M. Berenholtz, A. H. Caughey, D. Silver, J. Tian, R. E. Thompson, M. R. Miller and C. U. Lehmann (2012). "Electronic health record-based monitoring of primary care patients at risk of medication-related toxicity." *The Joint Commission Journal on Quality and Patient Safety*, **38**(5): 216-223.

Burns, E. M., A. Bottle, P. Aylin, A. Darzi, R. J. Nicholls and O. Faiz (2011). "Variation in reoperation after colorectal surgery in England as an indicator of surgical performance: retrospective analysis of Hospital Episode Statistics." *BMJ*, **343**: d4836.

Buxton, M., N. Caine, D. Chase, D. Connelly, A. Grace, C. Jackson, J. Parkes and L. Sharples (2006). "A review of the evidence on the effects and costs of implantable cardioverter defibrillator therapy in different patient groups, and modelling of cost-effectiveness and cost-utility for these groups in a UK context." *Health Technology Assessment*, **10**(27): iii-iv, ix-xi, 1-164.

Byrne, R. and A. P. Morrison (2014). "Service users' priorities and preferences for treatment of psychosis: a user-led Delphi study." *Psychiatric Services*, **65**(9): 1167-1169.

Canas, F., M. J. Tanasijevic, N. Ma'luf and D. W. Bates (1999). "Evaluating the appropriateness of digoxin level monitoring." *Archives of Internal Medicine*, **159**(4): 363-368.

Carroll, K. J. (2003). "On the use and utility of the Weibull model in the analysis of survival data." *Controlled Clinical Trials*, **24**(6): 682-701.

Chan, T. Y. K. and J. A. J. H. Critchley (1994). "Reporting of adverse drug reactions in relation to general medical admissions to a teaching hospital in Hong Kong." *Pharmacoepidemiology and Drug Safety*, **3**(2): 85-89.

Charlson, M. E., P. Pompei, K. L. Ales and C. R. MacKenzie (1987). "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." *Journal of Chronic Diseases*, **40**(5): 373-383.

Chiquette, E., M. G. Amato and H. I. Bussey (1998). "Comparison of an anticoagulation clinic with usual medical care: anticoagulation control, patient outcomes, and health care costs." *Archives of Internal Medicine*, **158**(15): 1641-1647.

Chobanian, A. V., G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo, Jr., D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright, Jr., E. J. Roccella (2003). "The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report." *JAMA,* **289**(19): 2560-2572.

Clague, H. W., Y. Twum-Barima and S. G. Carruthers (1983). "An audit of requests for therapeutic drug monitoring of digoxin: problems and pitfalls." *Therapeutic Drug Monitoring*, **5**(3): 249-254.

Classen, D. C., S. L. Pestotnik, R. S. Evans, J. F. Lloyd and J. P. Burke (1997). "Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality." *JAMA*, **277**(4): 301-306.

Classen, S., W. M. MStat, K. Walsh and W. Mann (2006). "The relationship of classes of commonly prescribed medications to functional status and quality of life for frail home-based older adults." *Physical & Occupational Therapy in Geriatrics*, **24**(1): 25-44.

Coast, J., R. Smith and P. Lorgelly (2008). "Should the capability approach be applied in health economics?" *Health Economics*, **17**(6): 667-670.

Cochran, W. G. (1968). "The effectiveness of adjustment by subclassification in removing bias in observational studies." *Biometrics*, **24**(2): 295-313.

Cochran, W. G. and D. B. Rubin (1973). "Controlling bias in observational studies: A review." *The Indian Journal of Statistics*, Series A: 417-446.

Coleman, J. J., R. E. Ferner and S. J. Evans (2006). "Monitoring for adverse drug reactions." *British Journal of Clinical Pharmacology*, **61**(4): 371-378.

Coleman, J. J., S. E. McDowell, S. J. Evans, P. S. Gill and R. E. Ferner (2010). "Oversight: a retrospective study of biochemical monitoring in patients beginning antihypertensive drug treatment in primary care." *British Journal of Clinical Pharmacology*, **70**(1): 109-117.

Cordarone X (2004). Summary of product characteristics, Sanofi Aventis.

Cox, A. R., C. Anton, C. H. Goh, M. Easter, N. J. Langford and R. E. Ferner (2001). "Adverse drug reactions in patients admitted to hospital identified by discharge ICD-10 codes and by spontaneous reports." *British Journal of Clinical Pharmacology*, **52**(3): 337-339.

Craig, S. (2010). "Hyponatraemia in emergency medicine." Retrieved 09/03/2015, from http://emedicine.medscape.com/article/767624-overview.

D'Agostino, R. B., Jr. (1998). "Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group." *Statistics in Medicine*, **17**(19): 2265-2281.

D'Agostino, R. B. and H. Kwan (1995). "Measuring effectiveness. What to expect without a randomized control group." *Medical Care*, **33**(4 Suppl): AS95-105.

Dale, J., S. Connor and K. Tolley (2003). "An evaluation of the west Surrey telemedicine monitoring project." *Journal of Telemedicine and Telecare*, **9 Suppl 1**: S39-41.

Darchy, B., E. Le Mière, B. Figuérédo, E. Bavoux and Y. Domart (1999). "Iatrogenic diseases as a reason for admission to the intensive care unit: Incidence, causes, and consequences." *Archives of Internal Medicine*, **159**(1): 71-78.

Dave, S. and I. Petersen (2009). "Creating medical and drug code lists to identify cases in primary care databases." *Pharmacoepidemiology and Drug Safety*, **18**(8): 704-707.

Davies, C., A. Briggs, P. Lorgelly, G. Garellick and H. Malchau (2013). "The "hazards" of extrapolating survival curves." *Medical Decision Making*, **33**(3): 369-380.

Davies, E. C., C. F. Green, S. Taylor, P. R. Williamson, D. R. Mottram and M. Pirmohamed (2009). "Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes." *PLoS One*, **4**(2): e4439.

Delbanco, T. and S. K. Bell (2007). "Guilty, afraid, and alone--struggling with medical error." *New England Journal of Medicine*, **357**(17): 1682-1683.

Department of Health (2000). "An organisation with a memory." London.

Department of Health (2001). "Building a safer NHS for patients: implementing an organisation with a memory." London.

Department of Health (2003). "Making Amends: a consultation paper setting out proposals for reforming the approach to clinical negligence in the NHS." London.

Department of Health (2004). "The quality and outcomes framework." London.

Department of Health (2008). "Pharmacy in England. Building on strengths-delivering the future." London.

Department of Health (2014). "NHS Reference Costs 2013 to 2014." London.

Dixon, J., C. Sanderson, P. Elliott, P. Walls, J. Jones and M. Petticrew (1998). "Assessment of the reproducibility of clinical coding in routinely collected hospital activity data: a study in two hospitals." *Journal of Public Health Medicine*, **20**(1): 63-69.

Dormuth, C. R., B. R. Hemmelgarn, J. M. Paterson, M. T. James, G. F. Teare, C. B. Raymond, J. P. Lafrance, A. Levy, A. X. Garg, P. Ernst and S. Canadian Network for Observational Drug Effect (2013). "Use of high potency statins and rates of admission for acute kidney injury: multicenter, retrospective observational analysis of administrative databases." *BMJ*, **346**: f880.

Dosiou, C., J. Barnes, A. Schwartz, R. Negro, L. Crapo and A. Stagnaro-Green (2012). "Cost-effectiveness of universal and risk-based screening for autoimmune thyroid disease in pregnant women." *Journal of Clinical Endocrinology and Metabolism*, **97**(5): 1536-1546.

Dosiou, C., G. D. Sanders, S. S. Araki and L. M. Crapo (2008). "Screening pregnant women for autoimmune thyroid disease: a cost-effectiveness analysis." *European Journal of Endocrinology*, **158**(6): 841-851.

Drummond, M. (1997). "Methods for the economic evaluation of health care programmes." Oxford, Oxford University Press.

Eccles, M., N. Freemantle and J. Mason (1998). "North of England evidence based development project: guideline for angiotensin converting enzyme inhibitors in primary care management of adults with symptomatic heart failure." *BMJ*, **316**(7141): 1369-1375.

Edwards, I. R. and J. K. Aronson (2000). "Adverse drug reactions: definitions, diagnosis, and management." *Lancet*, **356**(9237): 1255-1259.

Elliott, R. A., K. D. Putman, M. Franklin, L. Annemans, N. Verhaeghe, M. Eden, J. Hayre, S. Rodgers, A. Sheikh and A. J. Avery (2014). "Cost effectiveness of a pharmacist-led information technology intervention for reducing rates of clinically important errors in medicines management in general practices (PINCER)." *Pharmacoeconomics*, **32**(6): 573-590.

Erewash, Central & Greater Derby, South & Dales, Amber Valley Primary Care Trusts (2006). "Drug Monitoring Guidelines."

Esnaola, N. F., S. B. Cantor, S. I. Sherman, J. E. Lee and D. B. Evans (2001). "Optimal treatment strategy in patients with papillary thyroid cancer: a decision analysis." *Surgery*, **130**(6): 921-930.

Fang, M. C., A. S. Go, Y. Chang, L. H. Borowsky, N. K. Pomernacki, N. Udaltsova and D. E. Singer (2014). "Long-term survival after ischemic stroke in patients with atrial fibrillation." *Neurology*, **82**(12): 1033-1037.

Feder, H. M., Jr., C. Osier and E. G. Maderazo (1981). "Chloramphenicol: A review of its use in clinical practice." *Reviews of Infectious Diseases*, **3**(3): 479-491.

Feldstein, A. C., D. H. Smith, N. Perrin, X. Yang, M. Rix, M. A. Raebel, D. J. Magid, S. R. Simon and S. B. Soumerai (2006). "Improved therapeutic monitoring with several interventions: a randomized trial." *Archives of Internal Medicine*, **166**(17): 1848-1854.

Ferner, R.E. (1999). "Errors in prescribing, preparing, and giving medicines: definition, classification, and prevention." In: Aronson, J.K. (1999). "Side effects of drugs" 22nd Edition. Amsterdam: Elsevier Science, 1999.

Fialova, D., E. Topinkova, G. Gambassi, H. Finne-Soveri, P. V. Jonsson, I. Carpenter, M. Schroll, G. Onder, L. W. Sorbye, C. Wagner, J. Reissigova, R. Bernabei and H. O. C. P. R. G. Ad (2005). "Potentially inappropriate medication use among elderly home care patients in Europe." *JAMA*, **293**(11): 1348-1358.

Fichman, M. P., H. Vorherr, C. R. Kleeman and N. Telfer (1971). "Diuretic-induced hyponatremia." *Annals of Internal Medicine*, **75**(6): 853-863.

Fick, D. M., J. L. Waller, J. R. Maclean, R. V. Heuvel, J. Tadlock, M. Gottlieb and C. B. Cangialose (2001). "Potentially inappropriate medication use in a Medicare managed care population: association with higher costs and utilization." *Journal of Managed Care Pharmacy*, **7**(5): 407-413.

Field, T. S., B. H. Gilman, S. Subramanian, J. C. Fuller, D. W. Bates and J. H. Gurwitz (2005). "The costs associated with adverse drug events among older adults in the ambulatory setting." *Medical Care*, **43**(12): 1171-1176.

Franklin, M., V. Berdunov, J. Edmans, S. Conroy, J. Gladman, L. Tanajewski, G. Gkountouras and R. A. Elliott (2014). "Identifying patient-level health and social care costs for older adults discharged from acute medical units in England." *Age and Ageing*, **43**(5): 703-707.

Ftouh, S., M. Thomas and G. Acute Kidney Injury Guideline Development (2013). "Acute kidney injury: summary of NICE guidance." *BMJ*, **347**: f4930.

Gandhi, T. K., H. R. Burstin, E. F. Cook, A. L. Puopolo, J. S. Haas, T. A. Brennan and D. W. Bates (2000). "Drug complications in outpatients." *Journal of General Internal Medicine*, **15**(3): 149-154.

Gandhi, T. K., S. N. Weingart, J. Borus, A. C. Seger, J. Peterson, E. Burdick, D. L. Seger, K. Shu, F. Federico, L. L. Leape and D. W. Bates (2003). "Adverse drug events in ambulatory care." *New England Journal of Medicine*, **348**(16): 1556-1564.

Garout, M., H. S. Tilney, P. P. Tekkis and P. Aylin (2008). "Comparison of administrative data with the Association of Coloproctology of Great Britain and Ireland (ACPGBI) colorectal cancer database." *International Journal of Colorectal Disease*, **23**(2): 155-163.

Gelfand, J. M. (2005). "The UK General Practice Research Database." In: Strom, B. L. (2005) "Pharmacoepidemiology." Wiley-Blackwell.

Gibson, P. G., H. Powell, J. Coughlan, A. J. Wilson, M. Abramson, P. Haywood, A. Bauman, M. J. Hensley and E. H. Walters (2003). "Self-management education and regular practitioner review for adults with asthma." *Cochrane Database of Systematic Reviews*: CD001117.

Glasziou, P., L. Irwig and D. Mant (2005). "Monitoring in chronic disease: a rational approach." *BMJ*, **330**(7492): 644-648.

Glasziou, P. P. (2008). "Evidence-Based Medical Monitoring: From Principles to Practice." BMJ Books.

Goettler, M., S. Schneeweiss and J. Hasford (1997). "Adverse drug reaction monitoring--cost and benefit considerations. Part II: cost and preventability of adverse drug reactions leading to hospital admission." *Pharmacoepidemiology and Drug Safety*, **6 Suppl 3**: S79-90.

Goldman, R. E., C. S. Soran, G. L. Hayward and S. R. Simon (2010). "Doctors' perceptions of laboratory monitoring in office practice." *Journal of Evaluation of Clinical Practice*, **16**(6): 1136-1141.

GP Online. (2015). "NHS England unveils £15m fund to employ pharmacists in GP practices." Retrieved 20/10/2015, from http://www.gponline.com/nhs-england-unveils-15m-fund-employ-pharmacists-gp-practices/article/1355033.

Gray, A. (2003). "Adverse events in the National Health Service: an economic perspective." University of Oxford.

Gurwitz, J. H., T. S. Field, L. R. Harrold, J. Rothschild, K. Debellis, A. C. Seger, C. Cadoret, L. S. Fish, L. Garber, M. Kelleher and D. W. Bates (2003). "Incidence and preventability of adverse drug events among older persons in the ambulatory setting." *JAMA*, **289**(9): 1107-1116.

Gurwitz, J. H., T. S. Field, J. Judge, P. Rochon, L. R. Harrold, C. Cadoret, M. Lee, K. White, J. LaPrino, J. Erramuspe-Mainard, M. DeFlorio, L. Gavendo, J. Auger and D. W. Bates (2005). "The incidence of adverse drug events in two large academic long-term care facilities." *American Journal of Medicine*, **118**(3): 251-258.

Hallas, J., B. Harvald, L. F. Gram, E. Grodum, K. Brosen, T. Haghfelt and N. Damsbo (1990). "Drug related hospital admissions: the role of definitions and intensity of data collection, and the possibility of prevention." *Journal of Internal Medicine*, **228**(2): 83-90.

Havranek, E. P., F. Abrams, E. Stevens and K. Parker (1998). "Determinants of mortality in elderly patients with heart failure: the role of angiotensin-converting enzyme inhibitors." *Archives of Internal Medicine*, **158**(18): 2024-2028.

Hazell, L. and S. A. Shakir (2006). "Under-reporting of adverse drug reactions : a systematic review." *Drug Safety*, **29**(5): 385-396.

Health & Social Care Information Centre (2015). "OPCS-4 Classification." Retrieved 29/06/2015, from http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/opcs4/index_html

Health & Social Care Information Centre. (2015). "Hospital Episode Statistics." Retrieved 09/03/2015, from http://www.hscic.gov.uk/hes.

Herbert, R. J., A. J. Gagnon, J. E. Rennick and J. L. O'Loughlin (2009). "A systematic review of questionnaires measuring health-related empowerment." *Research and Theory ofr Nursing Practice*, **23**(2): 107-132.

Hippisley-Cox, J. and C. Coupland (2010). "Unintended effects of statins in men and women in England and Wales: population based cohort study using the QResearch database." *BMJ*, **340**: c2197.

Hippisley-Cox, J., D. Stables and M. Pringle (2004). "QRESEARCH: a new general practice database for research." *Informatics in Primary Care*, **12**(1): 49-50.

Hirsh, J. (1995). "Optimal intensity and monitoring warfarin." *American Journal of Cardiology*, **75**(6): 39B-42B.

HM Treasury (2011). "Appraisal and Evaluation in Central Government." London.

Hobbs, F. D., A. K. Roalfe, R. C. Davis, M. K. Davies, R. Hare and C. Midlands Research Practices (2007). "Prognosis of all-cause heart failure and borderline left ventricular systolic dysfunction: 5 year mortality follow-up of the Echocardiographic Heart of England Screening Study (ECHOES)." *European Heart Journal*, **28**(9): 1128-1134.

Hoch, I., A. D. Heymann, I. Kurman, L. J. Valinsky, G. Chodick and V. Shalev (2003). "Countrywide computer alerts to community physicians improve potassium testing in patients receiving diuretics." *Journal of American Medical Informatics Association*, **10**(6): 541-546.

Holland, R., I. Brooksby, E. Lenaghan, K. Ashton, L. Hay, R. Smith, L. Shepstone, A. Lipp, C. Daly, A. Howe, R. Hall and I. Harvey (2007). "Effectiveness of visits from community pharmacists for patients with heart failure: HeartMed randomised controlled trial." *BMJ*, **334**(7603): 1098.

Hong, G. (2012). "Marginal mean weighting through stratification: a generalized method for evaluating multivalued and multiple treatments with nonexperimental data." *Psychological Methods*, **17**(1): 44-60.

Honigfeld, G. (1996). "Effects of the clozapine national registry system on incidence of deaths related to agranulocytosis." *Psychiatric Services*, **47**(1): 52-56.

Horler, A. R. (1981). "Blood disorders." In: Davies, D. M. (1981) "Textbook of Adverse Drug Reactions." Oxford University Press.

Hougardy, D. M., G. M. Peterson, M. D. Bleasel and C. T. Randall (2000). "Is enough attention being given to the adverse effects of corticosteroid therapy?" *Journal of Clinical Pharmacy and Therapeutics*, **25**(3): 227-234.

Houghton, S. G., D. R. Farley, M. D. Brennan, J. A. van Heerden, G. B. Thompson and C. S. Grant (2004). "Surgical management of amiodarone-associated thyrotoxicosis: Mayo Clinic experience." *World Journal of Surgery*, **28**(11): 1083-1087.

Howard, R. L., A. J. Avery, P. D. Howard and M. Partridge (2003). "Investigation into the reasons for preventable drug related admissions to a medical admissions unit: observational study." *Quality and Safety in Health Care*, **12**(4): 280-285.

Howard, R. L., A. J. Avery, S. Slavenburg, S. Royal, G. Pipe, P. Lucassen and M. Pirmohamed (2007). "Which drugs cause preventable admissions to hospital? A systematic review." *British Journal of Clinical Pharmacology*, **63**(2): 136-147.

Hurley, J. S., M. Roberts, L. I. Solberg, M. J. Gunter, W. W. Nelson, L. Young and F. J. Frost (2005). "Laboratory safety monitoring of chronic medications in ambulatory care settings." *Journal of General Internal Medicine*, **20**(4): 331-333.

Jafari-Fesharaki, M. and M. M. Scheinman (1998). "Adverse Effects of Amiodarone." *Pacing and Clinical Electrophysiology*, **21**(1): 108-120.

Jick, S. S., J. A. Kaye, C. Vasilakis-Scaramozza, L. A. Garcia Rodriguez, A. Ruigomez, C. R. Meier, R. G. Schlienger, C. Black and H. Jick (2003). "Validity of the general practice research database." *Pharmacotherapy*, **23**(5): 686-689.

Johnson, J. A. and J. L. Bootman (1995). "Drug-related morbidity and mortality. A cost-of-illness model." *Archives of Internal Medicine*, **155**(18): 1949-1956.

Johnson, S. G., K. Canty, S. Billups and J. Schimmer (2010). "Adherence to amiodarone monitoring recommendations before and after implementation of a centralized pharmacy service: a cohort study." *Journal of Pharmacy Practice*, **23**(6): 536-539.

Johnson, W. G., T. A. Brennan, J. P. Newhouse, L. L. Leape, A. G. Lawthers, H. H. Hiatt and P. C. Weiler (1992). "The economic consequences of medical injuries. Implications for a no-fault insurance plan." *JAMA*, **267**(18): 2487-2492.

Joint Formulary Committee (2014) "British National formulary." London, BMJ Group and Pharmaceutical Press.

Joint Specialty Committee on Renal Medicine of the Royal College of Physicians of London and the Renal Association (2006). "Chronic kidney disease in adults: UK guidelines for identification, management and referral."

Kalra, P. A., M. Kumwenda, P. MacDowall and M. O. Roland (1999). "Questionnaire study and audit of use of angiotensin converting enzyme inhibitor and monitoring in general practice: the need for guidelines to prevent renal failure." *BMJ*, **318**(7178): 234-237.

Khan, N. F., S. E. Harrison and P. W. Rose (2010). "Validity of diagnostic coding within the General Practice Research Database: a systematic review." *British Journal of General Practice*, **60**(572): e128-136.

Kinsella, S., S. Moran, M. O. Sullivan, M. G. Molloy and J. A. Eustace (2010). "Hyponatremia independent of osteoporosis is associated with fracture occurrence." *Clinical Journal of the American Society of Nephrology*, **5**(2): 275-280.

Klarin, I., A. Wimo and J. Fastbom (2005). "The association of inappropriate drug use with hospitalisation and mortality: a population-based study of the very old." *Drugs & Aging,* **22**(1): 69-82.

Kohn, K. T., Corrigan, J. M., & Donaldson, M. S. (1999) "Committee on quality of health care in America. To err is human: building a safer health system." *National Academy Press*, *199*, 9.

Koplan, K. E., A. D. Brush, M. S. Packer, F. Zhang, M. D. Senese and S. R. Simon (2012). ""Stealth" alerts to improve warfarin monitoring when initiating interacting medications." *Journal of General Internal Medicine*, **27**(12): 1666-1673.

Krahenbuhl-Melcher, A., R. Schlienger, M. Lampert, M. Haschke, J. Drewe and S. Krahenbuhl (2007). "Drug-related problems in hospitals: a review of the recent literature." *Drug Safety*, **30**(5): 379-407.

Krska, J., J. A. Cromarty, F. Arris, D. Jamieson, D. Hansford, P. R. Duffus, G. Downie and D. G. Seymour (2001). "Pharmacist-led medication review in patients over 65: a randomized, controlled trial in primary care." *Age and Ageing*, **30**(3): 205-211.

Kuperman, G. J., D. Boyle, A. Jha, E. Rittenberg, N. Ma'Luf, M. J. Tanasijevic, J. M. Teich, J. Winkelman and D. W. Bates (1998). "How promptly are inpatients treated for critical laboratory results?" *Journal of American Medical Informatics Association*, **5**(1): 112-119.

Kurth, T., A. M. Walker, R. J. Glynn, K. A. Chan, J. M. Gaziano, K. Berger, J. M. Robins (2005). "Results of multivariate logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. " American Journal of Epidemiology, **163**(3): 262-270.

Lazarou, J., B. H. Pomeranz and P. N. Corey (1998). "Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies." *JAMA*, **279**(15): 1200-1205.

Leape, L. L., D. W. Bates, D. J. Cullen, J. Cooper, H. J. Demonaco, T. Gallivan, R. Hallisey, J. Ives, N. Laird, G. Laffel and et al. (1995). "Systems analysis of adverse drug events. ADE Prevention Study Group." *JAMA* **274**(1): 35-43.

Leape, L. L., T. A. Brennan, N. Laird, A. G. Lawthers, A. R. Localio, B. A. Barnes, L. Hebert, J. P. Newhouse, P. C. Weiler and H. Hiatt (1991). "The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II." *New England Journal of Medicine*, **324**(6): 377-384.

Leendertse, A. J., A. G. Egberts, L. J. Stoker, P. A. van den Bemt and H. S. Group (2008). "Frequency of and risk factors for preventable medication-related hospital admissions in the netherlands." *Archives of Internal Medicine*, **168**(17): 1890-1896.

Lessing, C., A. Schmitz, B. Albers and M. Schrappe (2010). "Impact of sample size on variation of adverse events and preventable adverse events: systematic review on epidemiology and contributing factors." *Quality and Safety in Health Care*, **19**(6): e24.

Lewis, J. D., R. Schinnar, W. B. Bilker, X. Wang and B. L. Strom (2007). "Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research." *Pharmacoepidemiology and Drug Safety*, **16**(4): 393-401.

Lindley, C. M., M. P. Tully, V. Paramsothy and R. C. Tallis (1992). "Inappropriate medication is a major cause of adverse drug reactions in elderly patients." *Age and Ageing*, **21**(4): 294-300.

Lipsitz, L. A. (2012). "Understanding health care as a complex system: the foundation for unintended consequences." *JAMA*, **308**(3): 243-244.

Lo, H. G., M. E. Matheny, D. L. Seger, D. W. Bates and T. K. Gandhi (2009). "Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care." *Journal of the American Medical Informatics Association*, **16**(1): 66-71.

Loh, K. C. (2000). "Amiodarone-induced thyroid disorders: a clinical review." *Postgraduate Medical Journal*, **76**(893): 133-140.

London and South East Medicines Information Service, South West Medicines Information Service and Croydon Primary Care Trust. (2008) "Suggestions for Drug Monitoring in Adults in Primary Care." NHS.

Maddirala, S., A. Khan, A. Vincent and K. Lau (2008). "Effect of angiotensin converting enzyme inhibitors and angiotensin receptor blockers on serum

potassium levels and renal function in ambulatory outpatients: risk factors analysis." *American Journal of the Medical Sciences*, **336**(4): 330-335.

Madsen, J. J. (1993). "Comparison of concurrent and retrospective methods of detecting adverse drug reactions." *American Journal of Hospital Pharmacy*, **50**(12): 2556-2557.

Martino, E., F. Aghini-Lombardi, S. Mariotti, L. Bartalena, M. Lenziardi, C. Ceccarelli, G. Bambini, M. Safran, L. E. Braverman and A. Pinchera (1987). "Amiodarone iodine-induced hypothyroidism: risk factors and follow-up in 28 cases." *Clinical Endocrinology*, **26**(2): 227-237.

Martino, E., L. Bartalena, F. Bogazzi and L. E. Braverman (2001). "The effects of amiodarone on the thyroid." *Endocrine Reviews*, **22**(2): 240-254.

Martino, E., M. Safran, F. Aghini-Lombardi, R. Rajatanavin, M. Lenziardi, M. Fay, A. Pacchiarotti, N. Aronin, E. Macchia, C. Haffajee and et al. (1984). "Environmental iodine intake and thyroid dysfunction during chronic amiodarone therapy." *Annals of Internal Medicine*, **101**(1): 28-34.

Matheny, M. E., T. D. Sequist, A. C. Seger, J. M. Fiskio, M. Sperling, D. Bugbee, D. W. Bates and T. K. Gandhi (2008). "A randomized trial of electronic clinical reminders to improve medication laboratory monitoring." *Journal of the American Medical Informatics Association*, **15**(4): 424-429.

Matzuk, M. M., M. Shlomchik and L. M. Shaw (1991). "Making digoxin therapeutic drug monitoring more effective." *Therapeutic Drug Monitoring*, **13**(3): 215-219.

McDowell, S. E., J. J. Coleman, S. J. Evans, P. S. Gill and R. E. Ferner (2010). "Laboratory monitoring and adverse patient outcomes with antihypertensive therapy in primary care." *Pharmacoepidemiology and Drug Safety*, **19**(5): 482-489.

McDowell, S. E. and R. E. Ferner (2011). "Biochemical monitoring of patients treated with antihypertensive therapy for adverse drug reactions: a systematic review." *Drug Safety*, **34**(11): 1049-1059.

McKee, M., A. Britton, N. Black, K. McPherson, C. Sanderson and C. Bain (1999). "Interpreting the evidence: choosing between randomised and non-randomised studies." *BMJ*, **319**(7205): 312-315.

Medicines and Healthcare Products Regulatory Agency (2013). "Trends in UK spontaneous Adverse Drug Reaction (ADR) reporting between 2008-2012." Accessed on 21/10/2015 at: http://www.yccscotland.scot.nhs.uk/news/2014/Pages/Trends-in-UK-spontaneous-ADR-reporting-2008-%E2%80%93-2012-.aspx

Mehta, P. A., S. W. Dubrey, H. F. McIntyre, D. M. Walker, S. M. Hardman, G. C. Sutton, T. A. McDonagh and M. R. Cowie (2009). "Improving survival in the 6 months after diagnosis of heart failure in the past decade: population-based data from the UK." *Heart*, **95**(22): 1851-1856.

Menzin, J., L. Boulanger, O. Hauch, M. Friedman, C. B. Marple, G. Wygant, J. S. Hurley, S. Pezzella and S. Kaatz (2005). "Quality of anticoagulation control and costs of monitoring warfarin therapy among patients with atrial fibrillation in clinic settings: a multi-site managed-care study." *Annals of Pharmacotherapy*, **39**(3): 446-451.

Mooney, G. (1994). "What else do we want from our health services?" *Social Science and Medicine*, **39**(2): 151-154.

Morris, C. J., S. Rodgers, V. S. Hammersley, A. J. Avery and J. A. Cantrill (2004). "Indicators for preventable drug related morbidity: application in primary care." *Quality and Safety in Health Care*, **13**(3): 181-185.

Naranjo, C. A., U. Busto, E. M. Sellers, P. Sandor, I. Ruiz, E. A. Roberts, E. Janecek, C. Domecq and D. J. Greenblatt (1981). "A method for estimating the probability of adverse drug reactions." *Clinical Pharmacology & Therapeutics*, **30**(2): 239-245.

National Health Service (2007). "Changes to the GMS contract 2006/07."

National Institute for Health and Care Excellence (2004). "Guide to the Methods of Technology Appraisal." NICE: London.

National Institute for Health and Care Excellence (2006). "Methods for development of NICE public health guidance." NICE: London.

National Institute for Health and Care Excellence (2010). "Chronic heart failure: Management of chronic heart failure in adults in primary and secondary care (Clinical guideline 108)." NICE: London.

National Institute for Health and Care Excellence (2011). "Chronic heart failure quality standard (Quality standard 9)." NICE: London.

National Institute for Health and Care Excellence (2011). "Hypertension: Clinical management of primary hypertension in adults (Clinical Guideline 127)." NICE: London.

National Institute for Health and Care Excellence (2014). "Atrial fibrillation: the management of atrial fibrillation." NICE: London.

National Institute for Health and Care Excellence (2014). "Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care (Clinical guideline 182)." NICE: London

National Institute for Health and Care Excellence (2014). "Lipid Modification - Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease (Clinical Guideline 181)". NICE: London.

NHS Scotland (2014). "Stroke statistics." Accessed on 21/10/2015 at: https://isdscotland.scot.nhs.uk/Health-Topics/Stroke/Publications/2014-01-28/2014-01-28-Stroke-Report.pdf?80306643248

Nobili, A., S. Garattini and P. M. Mannucci (2011). "Multiple diseases and polypharmacy in the elderly: challenges for the internist of the third millennium." *Journal of Comorbidity*, **1**(1): 17.

Nolan, J. P., N. J. Tarsa and G. DiBenedetto (1985). "Case-finding for unsuspected thyroid disease: costs and health benefits." *American Journal of Clinical Pathology*, **83**(3): 346-355.

North West (2002). "Drug Information Letter - Monitoring Drug Therapy."
Nussbaum, M. (2003). "Capabilities as fundamental entitlements: Sen and social justice." *Feminist Economics*, **9**(2-3): 33-59.

Nussbaum, M. (2003). "Capabilities As Fundamental Entitlements: Sen And Social Justice." Feminist Economics **9**(2-3): 33-59.

Office of National Statistics (2014). "Mortality in the United Kingdom, 1983-2013." Accessed on 21/10/2015 at: http://www.ons.gov.uk/ons/rel/mortality-ageing/mortality-in-the-united-kingdom/1983-2013/index.html

Ogilvie, R. I. and J. Ruedy (1967). "Adverse drug reactions during hospitalization." *Canadian Medical Association Journal*, **97**(24): 1450-1457.

Oram, R. A., T. J. McDonald and B. Vaidya (2013). "Investigating hypokalaemia." *BMJ*, **347**: f5137.

Orsini, M. J., P. A. Orsini, D. B. Thorn and J. N. Gallina (1995). "An ADR surveillance program: increasing quality, number of incidence reports." *Formulary*, **30**(8): 454-461.

Osman, F., M. D. Gammage, M. C. Sheppard and J. A. Franklyn (2002). "Clinical review 142: cardiac dysrhythmias and thyroid dysfunction: the hidden menace?" *Journal of Clinical Endocrinology and Metabolism*, **87**(3): 963-967.

Padilla, M. C., M. J. Armas-Hernandez, R. H. Hernandez, Z. H. Israili and M. Valasco (2007). "Update of diuretics in the treatment of hypertension." *American Journal of Therapeutics*, **14**(2): 154-160.

Palen, T. E., M. Raebel, E. Lyons and D. M. Magid (2006). "Evaluation of laboratory monitoring alerts within a computerized physician order entry system for medication orders." *American Journal of Managed Care*, **12**(7): 389-395.

Palmer, B. F. (2004). "Managing hyperkalemia caused by inhibitors of the renin-angiotensin-aldosterone system." *New England Journal of Medicine*, **351**(6): 585-592.

Passarelli, M. C., W. Jacob-Filho and A. Figueras (2005). "Adverse drug reactions in an elderly hospitalised population: inappropriate prescription is a leading cause." *Drugs and Aging*, **22**(9): 767-777.

Patanwala, A. E., A. B. Sanders, M. C. Thomas, N. M. Acquisto, K. A. Weant, S. N. Baker, E. M. Merritt and B. L. Erstad (2012). "A prospective, multicenter study of pharmacist activities resulting in medication error interception in the emergency department." *Annals of Emergency Medicine*, **59**(5): 369-373.

Patel, H., D. Bell, M. Molokhia, J. Srishanmuganathan, M. Patel, J. Car and A. Majeed (2007). "Trends in hospital admissions for adverse drug reactions in England: analysis of national hospital episode statistics 1998-2005." *BMC Clinical Pharmacology*, **7**: 9.

Payne, K., M. McAllister and L. M. Davies (2013). "Valuing the economic benefits of complex interventions: when maximising health is not sufficient." *Health Economics*, **22**(3): 258-271.

Payne, K., S. Nicholls, M. McAllister, R. Macleod, D. Donnai and L. M. Davies (2008). "Outcome measurement in clinical genetics services: a systematic review of validated measures." *Value in Health*, **11**(3): 497-508.

Payne, T. H., W. P. Nichol, P. Hoey and J. Savarino (2002). "Characteristics and override rates of order checks in a practitioner order entry system." *Proceedings of the AMIA Symposium*: 602-606.

Personal Social Services Research Unit (2014). "Unit Costs of Health & Social Care 2014."

Phoenix. (2014). "Delivering the vision: hardware break/fix in healthcare." Retrieved 06/12/2014, from http://www.phoenix.co.uk/media/37887/inpscasestudyfinal.pdf.

Pharmaceutical Journal. (2014). "Community pharmacists to have access to patients' medical records." Retrieved 20/10/2015, from http://www.pharmaceutical-journal.com/news-and-analysis/news/community-pharmacists-to-have-access-to-patients-medical-records/20066358.article.

Pickard, A. S., J. A. Johnson and D. H. Feeny (2005). "Responsiveness of generic health-related quality of life measures in stroke." *Quality of Life Research*, **14**(1): 207-219.

Pinilla, J., C. Murillo, G. Carrasco and C. Humet (2006). "Case-control analysis of the financial cost of medication errors in hospitalized patients." *European Journal of Health Economics*, **7**(1): 66-71.

Pirmohamed, M. and R. E. Ferner (2003). "Monitoring drug treatment." *BMJ*, **327**(7425): 1179-1181.

Pirmohamed, M., S. James, S. Meakin, C. Green, A. K. Scott, T. J. Walley, K. Farrar, B. K. Park and A. M. Breckenridge (2004). "Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients." *BMJ*, **329**(7456): 15-19.

Podrid, P. J. (1995). "Amiodarone: reevaluation of an old drug." *Annals of Internal Medicine*, **122**(9): 689-700.

Primary Care Quality & Information Service (2008). "Drug Monitoring: A Risk Management System."

Raebel, M. A. (2012). "Hyperkalemia associated with use of angiotensin-converting enzyme inhibitors and angiotensin receptor blockers." *Cardiovascular Therapy*, **30**(3): e156-166.

Raebel, M. A., E. A. Chester, E. E. Newsom, E. E. Lyons, J. A. Kelleher, C. Long, C. Miller and D. J. Magid (2006). "Randomized trial to improve laboratory safety monitoring of ongoing drug therapy in ambulatory patients." *Pharmacotherapy*, **26**(5): 619-626.

Raebel, M. A., E. E. Lyons, S. E. Andrade, K. A. Chan, E. A. Chester, R. L. Davis, J. L. Ellis, A. Feldstein, M. J. Gunter, J. E. Lafata, C. L. Long, D. J. Magid, J. V. Selby, S. R. Simon and R. Platt (2005). "Laboratory monitoring of drugs at initiation of therapy in ambulatory care." *Journal of General Internal Medicine*, **20**(12): 1120-1126.

Raebel, M. A., E. E. Lyons, E. A. Chester, M. A. Bodily, J. A. Kelleher, C. L. Long, C. Miller and D. J. Magid (2005). "Improving laboratory monitoring at initiation of drug therapy in ambulatory care: a randomized trial." *Archives of Internal Medicine*, **165**(20): 2395-2401.

Raebel, M. A., D. L. McClure, S. R. Simon, K. A. Chan, A. Feldstein, S. E. Andrade, J. E. Lafata, D. Roblin, R. L. Davis, M. J. Gunter and R. Platt (2007). "Laboratory monitoring of potassium and creatinine in ambulatory patients receiving angiotensin converting enzyme inhibitors and angiotensin receptor blockers." *Pharmacoepidemiology and Drug Safety*, **16**(1): 55-64.

Raebel, M. A., C. Ross, S. Xu, D. W. Roblin, C. Cheetham, C. M. Blanchette, G. Saylor and D. H. Smith (2010). "Diabetes and drug-associated hyperkalemia: effect of potassium monitoring." *Journal of General Internal Medicine*, **25**(4): 326-333.

Raeder, E. A., P. J. Podrid and B. Lown (1985). "Side effects and complications of amiodarone therapy." *American Heart Journal*, **109**(5 Pt 1): 975-983.

Reason, J. (1990). "Human Error." Cambridge University Press.

Reason, J. (2000). "Human error: models and management." *BMJ*, **320**(7237): 768-770.

Reichert, L. J. and H. A. de Rooy (1989). "Treatment of amiodarone induced hyperthyroidism with potassium perchlorate and methimazole during amiodarone treatment." *BMJ*, **298**(6687): 1547-1548.

Renneboog, B., W. Musch, X. Vandemergel, M. U. Manto and G. Decaux (2006). "Mild chronic hyponatremia is associated with falls, unsteadiness, and attention deficits." *American Journal of Medicine*, **119**(1): 71 e71-78.

Reynolds, D. J. and J. K. Aronson (1993). "ABC of monitoring drug therapy. Making the most of plasma drug concentration measurements." *BMJ*, **306**(6869): 48-51.

Roberts, M. S., J. A. Stokes, M. A. King, T. A. Lynne, D. M. Purdie, P. P. Glasziou, D. A. Wilson, S. T. McCarthy, G. E. Brooks, F. J. de Looze and C. B. Del Mar (2001). "Outcomes of a randomized controlled trial of a clinical pharmacy intervention in 52 nursing homes." *British Journal of Clinical Pharmacology*, **51**(3): 257-265.

Robins, J. M., M. A. Hernan and B. Brumback (2000). "Marginal structural models and causal inference in epidemiology." *Epidemiology*, **11**(5): 550-560.

Rodriguez-Monguio, R., M. J. Otero and J. Rovira (2003). "Assessing the economic impact of adverse drug effects." *Pharmacoeconomics*, **21**(9): 623-650.

Rosenbaum, P. R. and D. B. Rubin (1983). "The central role of the propensity score in observational studies for causal effects." *Biometrika*, **70**(1): 41-55.

Rosenbaum, P. R. and D. B. Rubin (1984). "Difficulties with regression analyses of age-adjusted rates." *Biometrics*, **40**(2): 437-443.

Rubin, D. B. (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology*, **Vol.86**(No. 5): 688-701.

Ryan, M. (1999). "A role for conjoint analysis in technology assessment in health care?" *International Jounral of Technology Assessment in Health Care*, **15**(3): 443-457.

Samoy, L. J., P. J. Zed, K. Wilbur, R. M. Balen, R. B. Abu-Laban and M. Roberts (2006). "Drug-related hospitalizations in a tertiary care internal medicine service of a Canadian hospital: a prospective study." *Pharmacotherapy*, **26**(11): 1578-1586.

Sarles, H., R. Camatte and J. Sahel (1977). "A Study of the Variations in the Response Regarding Duodenal Ulcer when Treated with Placebo by Different Investigators." *Digestion*, **16**(4): 289-292.

Schneeweiss, S., J. Hasford, M. Gottler, A. Hoffmann, A. K. Riethling and J. Avorn (2002). "Admissions caused by adverse drug events to internal medicine and emergency departments in hospitals: a longitudinal population-based study." *European Journal of Clinical Pharmacology*, **58**(4): 285-291.

Schoenenberger, R. A., M. J. Tanasijevic, A. Jha and D. W. Bates (1995). "Appropriateness of antiepileptic drug level monitoring." *JAMA*, **274**(20): 1622-1626.

Schuhlen, H., A. Kastrati, J. Dirschinger, J. Hausleiter, S. Elezi, A. Wehinger, J. Pache, M. Hadamitzky and A. Schomig (1998). "Intracoronary stenting and risk for major adverse cardiac events during the first month." *Circulation*, **98**(2): 104-111.

Scottish Intercollegiate Guidelines Network (2001). "Hypertension in Older People." Edinburgh.

Sen, A. (1993). "Capability and well-being." (pp. 30-53).

Sever, P., G. Beevers, C. Bulpitt, A. Lever, L. Ramsay, J. Reid and J. Swales (1993). "Management guidelines in essential hypertension: report of the second working party of the British Hypertension Society." *BMJ*, **306**(6883): 983-987.

Shah, N. R., A. C. Seger, D. L. Seger, J. M. Fiskio, G. J. Kuperman, B. Blumenfeld, E. G. Recklet, D. W. Bates and T. K. Gandhi (2006). "Improving acceptance of computerized prescribing alerts in ambulatory care." *Journal of the American Medical Informatics Association*, **13**(1): 5-11.

Shukla, R., N. I. Jowett, D. R. Thompson and J. E. Pohl (1994). "Side effects with amiodarone therapy." *Postgraduate Medical Journal*, **70**(825): 492-498.

Simon, S. R., S. E. Andrade, J. L. Ellis, W. W. Nelson, J. H. Gurwitz, J. E. Lafata, R. L. Davis, A. Feldstein and M. A. Raebel (2005). "Baseline laboratory monitoring of cardiovascular medications in elderly health maintenance organization enrollees." *Journal of the American Geriatric Society*, **53**(12): 2165-2169.

Singer, S., S. Lin, A. Falwell, D. Gaba and L. Baker (2009). "Relationship of safety climate and safety performance in hospitals." *Health Services Research*, **44**(2 Pt 1): 399-421.

Singh, J., J. Lord, L. Longworth, S. Orr, T. McGarry, R. Sheldon and M. Buxton (2012). "Does responsibility affect the public's valuation of health care interventions? A relative valuation approach to health care safety." *Value in Health*, **15**(5): 690-698.

Smellie, W. S., J. Forth, J. J. Coleman, W. Irvine, P. C. Dore, G. Handley, D. G. Williams, P. J. Galloway, K. G. Kerr, R. Herriot, G. P. Spickett and T. M. Reynolds (2007). "Best practice in primary care pathology: review 6." *Journal of Clinical Pathology*, **60**(3): 225-234.

Smith, C. C., P. M. Bennett, H. M. Pearce, P. I. Harrison, D. J. Reynolds, J. K. Aronson and D. G. Grahame-Smith (1996). "Adverse drug reactions in a hospital general medical unit meriting notification to the Committee on Safety of Medicines." *British Journal of Clinical Pharmacology*, **42**(4): 423-429.

Smith, D. H., M. A. Raebel, K. A. Chan, E. S. Johnson, A. F. Petrik, J. R. Weiss, X. Yang and A. Feldstein (2011). "An economic evaluation of a laboratory monitoring program for renin-angiotensin system agents." *Medical Decision Making*, **31**(2): 315-324.

Smolina, K., F. L. Wright, M. Rayner and M. J. Goldacre (2012). "Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study." *BMJ*, **344**: d8059.

Smolina, K., F. L. Wright, M. Rayner and M. J. Goldacre (2012). "Long-term survival and recurrence after acute myocardial infarction in England, 2004 to 2010." *Circulation Cardiovascular Quality and Outcomes,* **5**(4): 532-540.

Soumerai, S. B., T. J. McLaughlin, D. Spiegelman, E. Hertzmark, G. Thibault and L. Goldman (1997). "Adverse outcomes of underuse of beta-blockers in elderly survivors of acute myocardial infarction." *JAMA*, **277**(2): 115-121.

Steele, A. W., S. Eisert, J. Witter, P. Lyons, M. A. Jones, P. Gabow and E. Ortiz (2005). "The effect of automated alerts on provider ordering behavior in an outpatient setting." *PLoS Medicine*, **2**(9): e255.

Steinman, M. A., S. M. Handler, J. H. Gurwitz, G. D. Schiff and K. E. Covinsky (2011). "Beyond the prescription: medication monitoring and adverse drug events in older adults." *Journal of the American Geriatric Society*, **59**(8): 1513-1520.

Stelfox, H. T., S. B. Ahmed, J. Fiskio and D. W. Bates (2004). "Monitoring amiodarone's toxicities: recommendations, evidence, and clinical practice." *Clinical Pharmacology & Therapeutics*, **75**(1): 110-122.

Steuten, L. and M. Buxton (2010). "Economic evaluation of healthcare safety: which attributes of safety do healthcare professionals consider most important in resource allocation decisions?" *Quality and Safety in Health Care*, **19**(5): e6.

Stewart, K., S. Loftus and S. DeLisle (2003). "Prescription of amiodarone through a computerized template that includes both decision support and executive functions improves the monitoring for toxicities." *AMIA Annual Symposium Proceedings*: 1020.

Strom, B. L. (2006). "Pharmacoepidemiology." John Wiley & Sons.

Sullivan, P. W. and V. Ghushchyan (2006). "Preference-Based EQ-5D index scores for chronic conditions in the United States." *Medical Decision Making*, **26**(4): 410-420.

Tabachnick, B. G. (2007). "Using Multivariate Statistics." Pearson.

Tanda, M. L., F. Bogazzi, E. Martino and L. Bartalena (2008). "Amiodarone-induced thyrotoxicosis: something new to refine the initial diagnosis?" *European Journal of Endocrinology*, **159**(4): 359-361.

Taylor, M. F. (2010). "British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices." Colchester: University of Essex.

The Renal Association (2015). "CKD stages." Retrieved 20/10/2015, from http://www.renal.org/information-resources/the-uk-eckd-guide/ckd-stages#sthash.0GMRIvZQ.dpbs.

Thomas, E. J. and T. A. Brennan (2000). "Incidence and types of preventable adverse events in elderly patients: population based review of medical records." *BMJ*, **320**(7237): 741-744.

Thomsen, L. A., A. G. Winterstein, B. Sondergaard, L. S. Haugbolle and A. Melander (2007). "Systematic review of the incidence and characteristics of

preventable adverse drug events in ambulatory care." *Annals of Pharmacotherapy*, **41**(9): 1411-1426.

Tjia, J., S. H. Fischer, M. A. Raebel, D. Peterson, Y. Zhao, S. J. Gagne, J. H. Gurwitz and T. S. Field (2011). "Baseline and follow-up laboratory monitoring of cardiovascular medications." *Annals of Pharmacotherapy*, **45**(9): 1077-1084.

Toft, A. D. and N. A. Boon (2000). "Thyroid disease and the heart." *Heart*, **84**(4): 455-460.

Tsadok, M. A., C. A. Jackevicius, E. Rahme, V. Essebag, M. J. Eisenberg, K. H. Humphries, J. V. Tu, H. Behlouli, J. Joo and L. Pilote (2011). "Amiodarone-induced thyroid dysfunction: brand-name versus generic formulations." *Canadian Medical Association Journal*, **183**(12): E817-823.

Vincent, C., S. Taylor-Adams and N. Stanhope (1998). "Framework for analysing risk and safety in clinical medicine." *BMJ*, **316**(7138): 1154-1157.

Waller, P., M. Shaw, D. Ho, S. Shakir and S. Ebrahim (2005). "Hospital admissions for 'drug-induced' disorders in England: a study using the Hospital Episodes Statistics (HES) database." *British Journal of Clinical Pharmacology*, **59**(2): 213-219.

Warburton, R. N. (2005). "Patient safety--how much is enough?" *Health Policy*, **71**(2): 223-232.

Williams, J. G. and R. Y. Mann (2002). "Hospital episode statistics: time for clinicians to get involved?" *Clinical Medicine*, **2**(1): 34-37.

World Health Organisation (1972). "International drug monitoring: the role of national centres. Report of a WHO meeting." *World Health Organisation Technical Report Series*, **498**: 1-25.

World Health Organisation (1992). "The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines." Geneva: World Health Organization.

Wu, T. Y., M. H. Jen, A. Bottle, M. Molokhia, P. Aylin, D. Bell and A. Majeed (2010). "Ten-year trends in hospital admissions for adverse drug reactions in England 1999-2009." *Journal of the Royal Society of Medicine*, **103**(6): 239-250.

Yancy, C. W., M. Jessup, B. Bozkurt, J. Butler, D. E. Casey, M. H. Drazner, G. C. Fonarow, S. A. Geraci, T. Horwich and J. L. Januzzi (2013). "2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines." *Journal of the American College of Cardiology*, **62**(16): e147-e239.

Yiu, K. H., M. H. Jim, C. W. Siu, C. H. Lee, M. Yuen, M. Mok, Y. F. Shea, K. Fan, H. F. Tse and W. H. Chow (2009). "Amiodarone-induced thyrotoxicosis is a predictor of adverse cardiovascular outcome." *Journal of Clinical Endocrinology and Metabolism*, **94**(1): 109-114.

Yoos, H. L., H. Kitzman, A. McMullen, C. Henderson and K. Sidora (2002). "Symptom monitoring in childhood asthma: a randomized clinical trial comparing peak expiratory flow rate with symptom monitoring." *Annals of Allergy, Asthma & Immunology*, **88**(3): 283-291.

Appendix

Table A1. Summary of patient numbers per medication-specific cohort

Stage of cohort selection	ACEI	Loop diuretics	Amiodarone
Original	484370	237642	21381
On long-term therapy (proportion of original)	306203 (0.63)	82178 (0.35)	8349 (0.39)
With at least some monitoring (proportion of original)	261189 (0.54)	67080 (0.28)	5545 (0.26)
From HES-linked practices(proportion of original)	166157 (0.34)	41197 (0.17)	3327 (0.16)

Table A2. Breakdown of variables in HES raw data

Data sheet	Variable	Format	Description	Defined in NHS Dictionary	Application
Patient	patid	integer	unique identifier	0	Linking to CPRD ID
	pracid	integer	unique practice id	0	Linking to CPRD ID
	ethnos	integer	patient ethnicity	0	patient characteristics
	gen_HESid	integer	unique HES identifier	0	Linking to HES ID
	n_patid_hes	integer	no. of patients on CPRD assigned same HES identifier	0	Validation
Episodes	patid	integer	unique identifier	0	Linking to CPRD ID
	spno	integer	unique spell number of hospitalisation	0	Identying spell
	epikey	integer	unique episode number	0	Identifying episode
	admidate	date	date of admission	0	start of spell
	epistart	date	date of start of episode	0	start of episode
	epiend	date	date of end of episode	0	end of episode
	discharged	date	date of discharge	0	end of spell
	eorder	integer	order of spell within episode	0	identifying order to episode of care within a hospital spell
	epidur	integer	duration of episode in days	0	identifying LOS
	admimeth	integer	method of admission	1	elective/non-elective
	admisorc	integer	source of admission	1	home/other facility
	disdest	integer	discharge destination	1	destination following spell
	dismeth	integer	method of discharge	1	discharged/died
	mainspef	integer	specialty of consultant	1	type of consultant/TFC
	tretspef	integer	specialty of treatment	1	type of treatment/TFC
	pconsult	string	consultant code	0	Consultant providing care
	intmanig	integer	intended management	1	management of secondary care

					treatment
	classpat	integer	patient classification: 1=ordinary admission; 2=day case; 3=regular day attendance; 4=regular night attendance; 5=mothers/babies using only delivery facilities; 8=not applicable (other maternity event	0	inpatient/day case
	firstreg	string	first regular day or night admission?	0	regular admission
Diagnoses: by spell	patid	integer	unique identifier	0	Linking to CPRD ID
	spno	integer	unique spell number of hospitalisation	0	Identying spell
	admidate	date	date of admission	0	start of spell
	discharged	date	date of discharge	0	end of spell
	ICD	string	diagnosis code	0	diagnosis
	ICDx	string	5/6th character of ICD-10 code	0	diagnosis
Diagnoses: by episode	patid	integer	unique identifier	0	Linking to CPRD ID
	spno	integer	unique spell number of hospitalisation	0	Identying spell
	epikey	integer	unique episode number	0	Identifying episode
	epistart	date	date of start of episode	0	start of episode
	epiend	date	date of end of episode	0	end of episode
	ICD	string	diagnosis code	0	diagnosis
	ICDx	string	5/6th character of ICD-10 code	0	diagnosis
	d_order	integer	order of diagnosis codes	0	identification of primary and secondary diagnoses
Primary diagnoses	patid	integer	unique identifier	0	Linking to CPRD ID
	spno	integer	unique spell number of hospitalisation	0	Identying spell
	admidate	date	date of admission	0	start of spell
	discharged	date	date of discharge	0	end of spell
	ICD_primary	string	primary ICD-10 code	0	primary diagnosis
	ICDx	string	5/6th character of ICD-10 code	0	diagnosis
Procedures	patid	integer	unique identifier	0	Linking to CPRD ID
	spno	integer	unique spell number of hospitalisation	0	Identying spell
	epikey	integer	unique episode number	0	Identifying episode
	admidate	date	date of admission	0	start of spell
	epistart	date	date of start of episode	0	start of episode
	epiend	date	date of end of episode	0	end of episode
	discharged	date	date of discharge	0	end of spell
	OPCS	string	OPCS-4 procedure code	0	procedure
	evdate	date	date of procedure	0	date of procedure
	p_order	integer	order of OPCS-4 codes	0	identifying primary and secondary procedures

Table A3. List of study medications and corresponding CPRD product codes

Medication class	Medication	CPRD product codes
ACE inhibitor	captopril	1121, 1143, 1144, 1807, 3069, 3310, 3839, 15958, 17624, 17633, 18269, 18325, 20849, 21943, 23478, 24482, 25998, 26995, 28486, 28820, 30039, 32048, 32514, 33336, 33646, 34544, 34562, 34719, 34936, 34937, 35302, 36742, 37655, 39227, 39512, 41617, 41633, 41743, 43432, 43507, 43649, 44527, 45228, 46851, 46951, 46957, 52293, 52499

	captopril/hydrochlorothiazide	1520, 3203, 10902, 11133, 11351, 11561, 11641, 15135, 18263, 32166, 39227
	cilazapril	12411, 12412, 12574, 13026, 15605, 16196, 16197, 16212, 21053, 23642
	enalapril	196, 448, 1299, 1904, 8105, 8106, 8800, 8830, 11197, 13755, 15085, 16708, 19208, 20188, 22439, 22708, 23252, 24041, 27871, 28127, 29530, 31587, 31716, 32241, 33057, 33078, 34400, 34453, 34712, 34768, 34798, 34952, 34953, 35794, 36753, 37080, 37087, 41417, 41694, 41746, 42723, 42894, 42901, 42902, 42908, 43411, 43563, 44657, 45217, 46974, 50334, 50780, 50863, 52010, 52882, 53719, 53915
	enalapril/hydrochlorothiazide	1021, 5189
	fosinopril	633, 4571, 5861, 13589
	imidapril	6408, 12815, 12858, 16924, 18219, 32560
	lisinopril	65, 69, 78, 277, 3720, 6806, 6807, 8268, 10882, 11987, 12313, 14387, 16701, 19198, 19204, 19223, 20975, 30921, 32597, 33977, 34471, 34696, 34799, 37778, 41522, 41532, 41538, 41573, 43412, 43413, 43416, 43418, 43566, 45300, 45324, 45337, 45816, 46975, 46979, 47159
	lisinopril/hydrochlorothiazide	2982, 6359, 6786, 9764, 17655, 21231, 33353, 37710, 38995, 39137, 39147, 39242
	moexipril	15121, 17120, 28724, 28725
	perindopril	97, 593, 5612, 5800, 6078, 11983, 14960, 33095, 35731, 37930, 37964, 37965, 37971, 38026, 38034, 38285, 38510, 43012, 43813, 45319, 45938, 51807
	perindopril/indapamide	6794, 14228, 37908, 37978
	quinapril	3929, 5159, 6765, 7314, 9731, 14477, 14478, 15096, 38854, 38899, 40355, 46365
	quinapril/hydrochlorothiazide	15031, 15108
	ramipril	80, 82, 147, 654, 709, 756, 761, 5275, 5735, 6261, 6288, 6314, 6362, 6364, 9646, 9693, 9915, 11937, 28586, 29627, 32857, 32934, 33811, 33894, 34357, 34382, 34390, 34412, 34429, 34431, 34432, 34490, 34505, 34528, 34539, 34540, 34567, 34583, 34589, 34651, 34652, 34657, 34698, 34710, 34732, 34877, 34893, 34943, 35007, 38308, 39355, 39421, 40384, 42081, 45264, 45340, 45554, 46890, 47021, 50509
	ramipril/felodipine	11567, 11965, 17006, 17474, 21162, 28438
	trandolapril	4103, 5047, 7419, 8025, 8026, 9948, 16710, 28902, 29130, 31307, 31810
	trandolapril/verapamil hydrochloride	18223, 19690, 20579
Loop diuretics	bumetanide	814, 1776, 2788, 5218, 6160, 7806, 12226, 12294, 15341, 19300, 30913, 31932, 32091, 34613, 34934, 36767, 39602, 45305
	bumetanide & amiloride hydrochloride	2493, 2495, 14587

	furosemide	6, 55, 562, 3248, 3297, 4182, 4258, 4705, 5249, 5728, 5868, 6118, 7582, 7606, 7734, 7799, 8102, 9680, 10392, 10422, 10781, 12318, 14761, 14837, 16206, 17960, 18716, 19056, 19192, 19194, 19258, 20538, 21849, 24835, 25334, 25717, 26292, 27447, 27690, 27696, 27926, 29780, 30625, 30875, 31548, 32277, 32896, 32918, 34006, 34374, 34557, 35162, 36190, 40247, 41292, 41405, 41828, 42388, 42488, 46116, 46699, 46948, 47815, 52887, 52900
	furosemide/amiloride hydrochloride	56, 193, 211, 1301, 1369, 2772, 3793, 4211, 4873, 5220, 9431, 9456, 13435, 15874, 18332, 18497, 21938, 25965, 28129, 30773, 31773, 33527, 33658, 34280, 34622, 38901, 39807, 41533, 41719, 43508, 47647
	furosemide/triamterene	3050, 2961, 11265
	furosemide/penbutolol	24832, 26529
	furosemide/spironolactone	4661, 7441
	torasemide	8052, 10066, 11268, 11487, 18096, 22658, 40190, 40738, 40898, 46525
amiodarone	amiodarone	188, 763, 7445, 8832, 12322, 17119, 27161, 27193, 29456, 31209, 32164, 32553, 33119, 34515, 34751, 34851, 35517, 41530, 41551, 46938, 49376, 50241, 53658, 53666

Table A4. Laboratory tests used to define exposure to monitoring, corresponding CPRD medical and Read codes

Drug	Term	Description	CPRD Medical code	Read code
ACE inhibitors/loop diuretics (sodium, potassium, creatinine and other U&E tests)	Serum creatinine	creatinine	5	44J3.00
	Serum potassium	potassium	8	44I4.00
	Serum sodium	sodium	9	44I5.00
	Urea and electrolytes	U&E	49	44JB
	Serum creatinine raised	creatinine	3927	44J3300
	Electrolytes normal	U&E	4351	44I2
	Plasma creatinine level	creatinine	13736	44JF
	Blood urea/renal function	U&E	13812	44J..00
	Serum electrolytes	U&E	13909	44I..00
	Plasma sodium level	sodium	14137	44h6
	Plasma potassium level	potassium	14138	44h8
	Electrolytes abnormal	U&E	15134	44I3
	Blood potassium level	potassium	17497	44hO
	Blood sodium level	sodium	17883	44h1
	Urea and electrolytes normal	U&E	18356	44120
	Urea and electrolytes abnormal	U&E	18667	44121
	Low serum potassium level	potassium	26279	44I4200
	Serum creatinine normal	creatinine	26903	44J3200

	Normal serum potassium level	potassium	26928	44I4000
	Raised serum potassium level	potassium	26929	44I4100
	Serum sodium level abnormal	sodium	26934	44I5100
	Serum sodium level normal	sodium	26935	44I5000
	Creatinine level	creatinine	27095	4Q40
	Serum creatinine abnormal	creatinine	31277	44J3000
	Serum creatinine low	creatinine	35545	44J3100
	Serum creatinine NOS	creatinine	42345	44J3z00
	Serum electrolytes NOS	U&E	43037	44IZ
	Potassium level	potassium	44370	4Q42
	Corrected serum creatinine level	creatinine	45096	44JD
	Corrected plasma creatinine level	creatinine	62062	44JC
	Sodium level	sodium	67485	4Q43
Amiodarone (thyroid function tests)	Serum TSH level	thyroid	63	442W.00
	Serum T4 level	thyroid	941	4426.00
	Thyroid hormone tests	thyroid	2698	442..00
	Serum TSH level abnormal	thyroid	3900	442A100
	Euthyroid	thyroid	4877	4421.11
	Thyroid function tests	thyroid	11367	442..13
	TSH level	thyroid	13598	442..12
	Free T3 level	thyroid	13602	4425.00
	Thyroid hormone tests normal	thyroid	13609	4421.00
	Serum T3 level	thyroid	13791	4424.00
	TSH - thyroid stim. hormone	thyroid	13811	442A.00
	Thyroid function test	thyroid	13906	442J.00
	Thyroid function tests normal	thyroid	13907	442H.00
	Free T4 level	thyroid	14459	4427.00
	Serum free T4 level	thyroid	14460	442V.00
	Thyroid hormone tests NOS	thyroid	16122	442Z.00
	Serum thyroglobulin level	thyroid	18178	442d.00
	Thyroid function tests abnormal	thyroid	18598	442I.00
	Serum T4 level normal	thyroid	19767	4426000
	Serum TSH level normal	thyroid	19769	442A000
	Plasma TSH level	thyroid	19770	442X.00
	Serum T3 level normal	thyroid	19787	4424000
	Serum total T4 level	thyroid	19788	442b.00
	Plasma total T4 level	thyroid	19789	442a.00
	Serum free triiodothyronine level	thyroid	19866	442U.00
	Thyroid hormone tests abnormal	thyroid	20970	442G.00
	Serum T4 level abnormal	thyroid	20983	4426100

	Serum T3 level high	thyroid	26936	4424100
	30 minute plasma TSH level	thyroid	26942	442K.00
	Thyroxine binding globulin	thyroid	26991	442..14
	PBI - protein bound iodine estimation	thyroid	26992	442E.00
	T3 uptake test	thyroid	26993	442D.00
	Free thyroxine index	thyroid	26994	4428.00
	TRH level	thyroid	26995	442..11
	Plasma free T4 level	thyroid	27228	442c.00
	Thyroid hormone tests high	thyroid	27278	4422.00
	Thyroid horm tests borderline	thyroid	27807	442C.00
	Thyroid hormone tests low	thyroid	33292	4423.00
	TRH - thyroid release. hormone	thyroid	33293	442B.00
	Plasma free triiodothyronine level	thyroid	34597	442Y.00
	Serum T3 level low	thyroid	35547	4424200
	Serum total T3 level	thyroid	35554	442f.00
	Blood spot TSH level	thyroid	37210	442e.00
	T3 suppression test	thyroid	42347	442F.00
	Thyroxine binding globulin-TBG	thyroid	48565	4429.00
	150 minute serum TSH level	thyroid	53286	442T.00
	T3 uptake test normal	thyroid	55155	442D000
	T3 uptake test abnormal	thyroid	59072	442D100
	60 minute plasma TSH level	thyroid	63031	442L.00
	60 minute serum TSH level	thyroid	65083	442Q.00
	90 minute plasma TSH level	thyroid	66215	442M.00
	120 minute plasma TSH level	thyroid	67108	442N.00
	Serum thyroglobulin recovery level	thyroid	68378	442g.00
	120 minute serum TSH level	thyroid	68780	442S.00
	30 minute serum TSH level	thyroid	93724	442P.00
	Thyrotropin releasing hormone (TRH) stimulation test	thyroid	94927	442h.00
	150 minute plasma TSH level	thyroid	100259	442O.00

Table A5. Description of variables used in the study

Description	Type	Estimation method	Source
Unique patient identifier	String	N/A	Patient file
Practice number	String	N/A	Patient file
Date of initiation of therapy	String	No. of days since 1960	Therapy file
No. of days of observed follow-up	Interval	Calculated as censor out date-initiation date	Therapy, clinical files
No. of observed monitoring intervals	Interval	Based on length of continuous therapy	Therapy, clinical files
Age at initiation of therapy	Interval	Calculated as initiation date-DOB	Therapy, patient files
Sex of patient	Dichotomous	0=Male; 1=Female	Patient file
No. of dates with records of monitoring test	Interval	Based on number of days with valid monitoring test	Test file
No. of monitoring tests per observed interval	Interval	Number of test days/number of periods	See source variables
Monitoring category	Ordinal	Based no. of tests per interval	See source variables
Presence of specific ADE during observed period	Dichotomous	0=no ADE; 1=ADE	Clinical file, HES diagnoses
Comorbidity score	Interval	No. of co-morbidities at baseline	Clinical and therapy files
No. of co-prescribed medications	Interval	No. of medications perscribed from different BNF chapters during the observed period	Therapy file
Specific co-prescribed medications alongside the index medication	Dichotomus	0=not co-prescribed; 1=co-prescribed	Therapy file
Chronic kidney disease stage at baseline	Categorical	1-5	Clinical file
Presence of diabetes mellitus at baseline	Dichotomous	0=no diabetes; 1=diabetes	Clinical file
Year of initiation of therapy	Categorical	0=2003; 1=2004; 2=2005; 3=2006; 4=2007; 5=2008; 6=2009 or later	Therapy file

Table A6. List of outcomes assessed in the study

Drug	Outcome	Description	CPRD medical code	Read code	Corresponding ICD-10 code(s)[ab]
ACE inhibitor	Hyperkalaemia	Diagnosis	5748	C367.00	E875
		Serum potassium>6mmol/L	8	44I4.00	
			13909	44I..00	
			14138	44h8	
			15134	44I3	
			17497	44hO	
			18667	44121	
			18356	44120	
			26928	44I4000	
			26929	44I4100	
			43037	44IZ	
			44370	4Q42	
	Acute kidney	Diagnosis	2266	K04..00	N171

	injury		25582	K04z.00	N172
					N178
					N179
		Serum creatinine rise of ≥26 µmol/L or ≥1.5 fold in a 7-day period	9	44I5.00	
			3927	44J3300	
			13736	44JF	
			26903	44J3200	
			27095	4Q40	
			31277	44J3000	
			42345	44J3z00	
			45096	44JD	
			62062	44JC	
Loop diuretics	Hypokalaemia	Diagnosis	3168	C368.00	E876
		Serum potassium <3.5mmol/L	8	44I4.00	
			49	44JB	
			4351	44I2	
			13909	44I..00	
			14138	44h8	
			15134	44I3	
			17497	44hO	
			18356	44120	
			18667	44121	
			26279	44I4200	
			26928	44I4000	
			43037	44IZ	
			44370	4Q42	
	Hyponatraemia	Serum sodium<135mmol/L	9	44I5.00	
			49	44JB	
			4351	44I2	
			13909	44I..00	
			14137	44h6	
			15134	44I3	
			17883	44h1	
			18356	44120	
			18667	44121	
			26934	44I5100	
			26935	44I5000	
			43037	44IZ	
			67485	4Q43	
	Accidental fall	Diagnosis	384	TC...11	
			6815	TC...00	
			6835	TCz..00	
			7948	TC52.00	

			8730	TCy..00	
			11308	TCyz.00	
			15112	TC5..00	
			16769	TC7..00	
			21081	TC01100	
			33887	TC4..00	
	Injury/fracture	Diagnosis			Any diagnoses from ICD-10 chapters S or T
Amiodarone	Hypothyroidism	Diagnosis	273	C04..13	E02X
					E030
					E031
					E032
					E033
					E038
					E039
			1619	C04..11	E035
			1882	C0...00	E034
					E018
					E055
					E078
					E079
			2518	C000.00	E010
					E011
					E012
					E040
					E042
					E048
					E049
					E071
			1346	C05..00	
			2572	C000.13	
			3290	C04..00	
			3436	C052.12	
			3655	C000.11	
			3941	C04z.00	
			10760	C02..12	
			14704	C04..12	
			24748	C04y.00	
			25913	C043.00	
			38976	C043z00	
			95830	C047.00	
		Serum TSH>10mU/L	63	442W.00	
			2698	442..00	

			3900	442A100	
			11367	442..13	
			13598	442..12	
			13609	4421.00	
			13811	442A.00	
			13906	442J.00	
			13907	442H.00	
			16122	442Z.00	
			18598	442I.00	
			19769	442A000	
			19770	442X.00	
			20970	442G.00	
			26942	442K.00	
			27278	4422.00	
			27807	442C.00	
			33293	442B.00	
			37210	442e.00	
			53286	442T.00	
			63031	442L.00	
			65083	442Q.00	
			66215	442M.00	
			67108	442N.00	
			68780	442S.00	
			93724	442P.00	
			100259	442O.00	
	Hyperthyroidism	Diagnosis	677	C02..00	E060
					E050
					E051
					E052
					E053
					E054
					E055
					E058
					E059
					E061
					E062
					E063
					E064
					E065
					E069
			1472	C02..11	
			1881	C00z.11	
			1882	C0...00	E034

					E018
					E055
					E078
					E079
			5257	C020.12	
		Serum TSH<0.2mU/L	63	442W.00	
			2698	442..00	
			3900	442A100	
			11367	442..13	
			13598	442..12	
			13609	4421.00	
			13811	442A.00	
			13906	442J.00	
			13907	442H.00	
			16122	442Z.00	
			18598	442I.00	
			19769	442A000	
			19770	442X.00	
			20970	442G.00	
			26942	442K.00	
			27807	442C.00	
			33292	4423.00	
			37210	442e.00	
			53286	442T.00	
			63031	442L.00	
			65083	442Q.00	
			66215	442M.00	
			67108	442N.00	
			68780	442S.00	
			93724	442P.00	
			100259	442O.00	
	MACE	Stroke			G458
					G459
					I248
					I249
					I64X
					M622
		Heart failure			I110
					I130
					I132
					I500
					I501
					I509

		Myocardial Infarction			I210
					I211
					I212
					I213
					I214
					I219
					I220
					I221
					I228
					I229
					I252
		Aneurism			I253
		Vascular arrhythmia			I470
					I471
					I472
					I479
					I490

Table A7. Criteria for the calculation of the multimorbidity score

Condition	Full identification	Diagnosis-only identification	Check[a]	Source
	Physical health			
Hypertension	Read code	Read code		Clinical
Painful condition	≥4 prescriptions of analgesic OR ≥4 anti-epileptic prescriptions in the absence of an epilepsy Read code	N/A		Therapy
Asthma	Read code AND prescription of asthma	Read code		Clinical
Heart disease	Read code	Read code		Clinical
Dyspepsia	≥4 prescriptions of BNF 0103% excluding antacids AND NOT (≥4 NSAIDs OR ≥4 aspirin/clopidogrel)	N/A		Therapy
Diabetes	Read code	Read code		Clinical
Thyroid disorder	Read code	Read code		Clinical
Rheumatoid arthritis	Read code	Read code		Clinical
Hearing loss	Read code	Read code		Clinical
COPD	Read code	Read code		Clinical
Irritable bowel syndrome	Read code OR ≥4 prescriptions of antispasmodic	Read code		Clinical
Cancer	Read code	Read code		Clinical
Constipation	≥4 prescriptions of laxative	N/A		Therapy
Stroke & transient ischaemic attack	Read code	Read code		Clinical
CKD	Read code	Read code		Clinical
Diverticular disease of intestine	Read code	Read code		Clinical
Atrial fibrillation	Read code	Read code		Clinical

Peripheral vascular disease	Read code	Read code		Clinical
Heart failure	Read code	Read code		Clinical
Prostate disorder	Read code	Read code		Clinical
Glaucoma	Read code	Read code		Clinical
Epilepsy	Read code and antiepileptic prescription	Read code		Therapy and clinical
Psoriasis or eczema	Read code AND ≥4 prescriptions of medication indicated for psoriasis or eczema (excluding simple emollients)	Read code		Clinical
Inflammatory bowel disease	Read code	Read code		Clinical
Migraine	≥4 prescriptions anti-migraine	N/A		
Blindness and low vision	Read code	Read code		Clinical
Chronic sinusitis	Read code	Read code		Clinical
Bronchiectasis	Read code	Read code		Clinical
Parkinson's disease	Read code	Read code		Clinical
MS	Read code	Read code		Clinical
Viral Hepatitis	Read code	Read code		Clinical
Chronic liver disease	Read code	Read code		Clinical
Mental health				
Depression	Read code or ≥4 prescriptions of antidepressant	Read code		Therapy and clinical
Anxiety and other neurotic condition	Read code OR ≥ 4 anxiolytic/hypnotic prescriptions OR ≥ 4 10/25mg amitriptyline & do not meet the criteria for 'Pain'	Read code		Therapy and clinical
Alcohol problems	Read code	Read code		Clinical
Psychoactive substance misuse	Read code	Read code		Clinical
Dementia	Read code	Read code		Clinical
Schizophrenia	Read code OR lithium prescription	Read code		Clinical
Learning disability	Read code	Read code		Clinical
Anorexia or bulimia	Read code	Read code		Clinical

[a] This column describes the required rule to identify presence of condition: green=diagnosis Read code only; yellow=either diagnosis or prescription; red=both diagnosis and prescription required.