The effectiveness and cost-effectiveness of public health interventions to prevent falls in children under 5 years

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Stephanie Jane Hubbard BSc MSc

Department of Health Sciences

University of Leicester

2018

Abstract

The effectiveness and cost-effectiveness of public health interventions to prevent falls in children under 5 years

Stephanie Hubbard

This work represents analyses undertaken as part of a National Institute for Health Research five-year multi-centre collaborative programme. The aim is to increase the evidence-base for interventions to prevent unintentional accidents in the home in children under-five. The focus is on the effectiveness and cost-effectiveness of interventions to increase the possession of fitted safety equipment or promote good safety practices to prevent falls.

Pairwise meta-analyses, comparing the effectiveness of an enhanced intervention to usual care, informed a Cochrane Review update. Interventions are heterogeneous containing multiple components so network meta-analysis (NMA) was used to identify the most effective of seven interventions. The most intensive intervention was most effective with households more likely to possess a fitted safety gate than in the usual care group (OR=7.73(95%CrI: 4.14 to 14.4)). Individual participant data was incorporated to explore the effect of covariates, including child age and gender, and socioeconomic status but there was little evidence on any effect.

The NMA results informed a cost-utility model to estimate the mean costs and quality adjusted life years (QALYS) associated with the interventions for increasing possession of a fitted stair safety gate. A simulated cohort of 100,000 UK households with a new-born were followed for the first three years, when there is highest risk of a fall, and then long-term. At a threshold value of £30,000 per QALY gained, none of the interventions were found to be cost-effective compared to usual care.

Appraisals of public health interventions are rarely informed by analyses beyond a narrative review and/or pairwise meta-analysis, often because of the perceived lack of high quality evidence, heterogeneity in study designs, including interventions, outcome measures and scope, and a lack of expertise. This work has illustrated that more complex evidence synthesis can be used to provide more explicit, transparent and appropriate results to inform decision making.

Acknowledgements

Contents

List of Tables

List of Figures

List of Abbreviations

A&E	Accident and Emergency
CBA	Controlled Before and After
CEA	Cost-effectiveness Analysis
CI	Confidence Interval
CPHE	Centre for Public Health Excellence
CrI	Credible Interval
D	Deviance
D_{res}	Residual deviance
DIC	Deviance information criteria
ED	Emergency Department
FE	Fixed Effects
HASS	Home and Accident Surveillance System
HES	Hospital Episode Statistics
HRQoL	Health-related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
IPB	Injury Prevention Briefing
IPD	Individual participant data
KCS	Keeping Children Safe
log	Natural logarithm
MA	Meta-analysis
MCMC	Markov Chain Monte-Carlo
ML	Maximum likelihood
MTC	Mixed treatment comparison
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
ONS	Office for National Statistics
OR	Odds ratio
PHE	Public Health England
PHI	Public Health Intervention
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALYs	Quality of Life Years
RCT	Randomised Controlled Trial
RE	Random Effects
REML	Restricted maximum likelihood
sd	Standard deviation
se	Standard error
TSD	Technical Support Document
VOI	Value of Information

1   Introduction

1.1   Aims of the thesis

An intervention is a set of actions with a coherent objective to bring about change or produce identifiable outcomes which may include policy, regulatory initiatives, single strategy projects or multi-component programmes. Public health interventions (PHI) are intended to promote or protect health or prevent ill health in communities or populations and are distinguished from clinical interventions, which are intended to prevent or treat illness in individuals (Rychetnik, Frommer et al. 2002) . Decisions on the funding of PHIs are dependent on evidence-based recommendations but a lack of evidence on the effectiveness and cost-effectiveness of the competing interventions hampers decision making. One example of this is in the area of interventions to prevent home injury in childhood (Kendrick, Ablewhite et al. 2017). To make the best use of the often complex evidence requires the use of increasingly sophisticated statistical methods. These methods have become more commonplace in the evaluation of the effectiveness and cost-effectiveness of clinical interventions, such as drugs, devices and medical procedures. By comparison (except in the evaluation of screening and immunisation which are fairly well developed), economic evaluations of other, broader, public health interventions are scarce and the methods less established. Only recently have there been efforts to consider applying decision modelling techniques to public health interventions (Shiell, Hawe et al. 2008, Drummond, Weatherly et al. 2007, Weatherly, Drummond et al. 2009).

The aim of this thesis is to examine what evaluation methods are currently used in practice, what other more advanced methods are available and to apply them to a motivating case study, and to consider whether these methods could be used more widely in the field of public health. To evaluate and illustrate their potential use, a case study provides evidence on interventions to prevent falls accidents in the home in children under 5 years of age. Advanced evidence synthesis methods are used to evaluate the effectiveness of the interventions using as much of the available evidence

as possible and then a decision model is developed to investigate the cost-effectiveness of the interventions. Due to the complexity of the interventions, and many issues arising in getting informative data, a number of challenges have to be addressed. The thesis starts with an introduction to the issues around evaluating and making decisions on public health interventions, then looks at the child falls prevention motivating example, followed by a description and application of evidence synthesis methods and decision modelling methods for the evaluation of effectiveness and cost-effectiveness respectively.

1.2   Effectiveness and cost-effectiveness analyses of interventions

This section will give a brief introduction to the evaluation of the effectiveness and cost-effectiveness of interventions. Further details are provided in Chapters 4 and 6.

1.2.1   Effectiveness analysis of interventions (evidence synthesis)

To assess the effectiveness of competing interventions on a specific outcome a comprehensive systematic review of the evidence base for the interventions needs to be undertaken. If the intervention effect estimates are available for more than one study then meta-analysis may be used to synthesise the evidence to give an overall estimate of effect across the evidence base (Higgins, Green 2011, Dias, Welton et al. 2013b). The best evidence on effectiveness comes from randomised controlled trials (RCTs); the design and conduct of the RCT should be such that as many sources of bias, systematic error or deviation from the truth in results or inferences, as possible are removed, using for example randomisation of interventions, allocation concealment and blinding of outcome assessment (Higgins, Green 2011). RCTs can have problems with external validity, they are not generalisable to the target population in terms of subjects, protocols or settings, and internal validity, problems arising from the design and conduct. Any studies considered for inclusion in a meta-analysis (e.g. RCTs, observational, and non-randomised studies) should be examined for risk of bias using

a tool such as the Cochrane Collaboration's tool (Higgins, Green 2011) as meta-analyses of studies that are at risk of bias may give seriously misleading results.

If interest is in comparing two interventions and all studies compare these two interventions head-to-head then a pairwise meta-analysis can be used (section 4.3). If there are multiple interventions of interest, evidenced from the systematic review, then a network meta- analysis (NMA) or mixed treatment comparison (MTC) analysis (section 4.11) may be appropriate to evaluate the overall effect estimates for each comparison of pairs of interventions including those not observed directly but estimated indirectly from the evidence. The probability that an intervention is "best", i.e. most effective, can be estimated and hence interventions can be ranked in the order of effectiveness. Analyses are often split by subject or study characteristics using sub-group analyses or meta-regression to investigate the effect of any heterogeneity in the studies (section 0). The data extracted from the original studies in a systematic review are often summary/ aggregated estimates of the characteristics, such as the mean age, percentage female, but meta-analysis methods can be extended to use individual participant data (IPD) from the original studies to get better estimates particularly in meta-regression where the interest is in the effect of participant characteristics (Lambert, Sutton et al. 2002) (section 4.9). The more complex meta-analysis methods are conducted using a Markov Chain Monte Carlo (MCMC) simulation approach (section 4.1).

The evaluation of effectiveness methods are described in detail in Chapter 3.6 and applied to the case-study in Chapter 5. The results from the effectiveness analysis can be used to inform the cost-effectiveness analyses of interventions and is described in 1.2.2.

1.2.2   Cost-effectiveness analysis

Limited funding for healthcare means that governments need to make choices of which interventions to fund. Economic evaluation compares alternative competing interventions in terms of the costs (intervention, resource use) and consequences

(health outcomes, effects, benefits) (Drummond 2005). Formally assessing the cost effectiveness of an intervention can help decision-makers ensure that maximum health gain, in a specified population over a specific timeframe, is achieved from the finite available resources. If resources are used for interventions that are not cost effective, the population as a whole can gain fewer benefits (NICE 2012).

A cost-utility analysis (CUA) is used in this thesis to investigate if any interventions provide greater benefit at the same or lower cost than current practice/usual care. The CUA uses monetary units for cost and determines the health outcome or benefit in terms of quality-adjusted life years (QALYs) that capture both length of life and health-related quality of life (HRQoL).

To assess the cost-effectiveness of interventions to prevent childhood falls, a comprehensive decision analytic model is defined to evaluate all possible interventions and consequences in terms of expected costs and expected health outcomes. A probabilistic decision model is estimated using a Bayesian MCMC simulation approach to allow for variability, individual subjects do not all respond the same way to the interventions, and uncertainty in the model parameters.

These methods are described in detail in Chapter 6 and applied to the case-study in Chapter 7.

1.3   Effectiveness and cost-effectiveness of public health interventions

This section will introduce public health interventions, some of the issues in evaluating the effectiveness and cost-effectiveness of these interventions and the motivating case study.

1.3.1   Public health interventions

Public health interventions, such as an intervention to prevent childhood injuries in the home, tend to be complicated and complex, programmatic and context dependent and the evidence for their effectiveness must be sufficiently comprehensive to encompass that complexity (Rychetnik, Frommer et al. 2002). Intervention

effectiveness describes the impact of an intervention programme on changing the behaviour of the population to protect health or prevent illness. The outcome of interest is often difficult to assess and other intermediate outcomes are used as surrogates, e.g. the uptake in the possession of a piece of safety equipment to prevent an accident rather than the occurrence of an accident.

In the economic evaluation of a public health intervention the intervention can generate very broad costs and benefits that are often directed at populations or communities rather than specific individuals. Standard approaches to valuing health gain (QALYs) may be inadequate, the gains may go beyond health and the individual subject. For example the effect of a home safety intervention to increase the possession of a piece of safety equipment may go beyond the effect on the child under five identified for the purpose of the intervention and may affect other children and parents, and often there is a concern about health inequalities; standard approaches tend to focus on efficiency rather than on equity (Weatherly, Drummond et al. 2009).

Some of the challenges of evaluating the effectiveness and cost-effectiveness using a decision modelling approach include:

- Scoping – defining the unit of analyses (child or family), perspective (e.g. NHS, all public services, societal), timescale (e.g. 1 year, 10 years, lifetime)
- Identifying the optimal package of interventions
- Few RCTs, complicated and complex interventions, broader population groups, little QoL and cost data available

The National Institute for Health and Care Excellence (NICE) is an independent public body that provides national guidance and advice to improve health and social care in England. NICE guidance offers evidence-based recommendations made by independent committees on a broad range of topics including public health (NICE 2014). In 2006 the Centre for Public Health Excellence (CPHE) in NICE developed the first of their public health guidance documents based on four factors, population, environment, society and organisations, that are linked to human behaviour and explain patterns of potentially preventable diseases and conditions including accidents and injuries (NICE 2012). Further details about NICE Public Health Guidance are

described in Chapter 2. The NICE guidelines are based on the core principles including using the best available evidence of what works and what it costs. This is the underlying principle of the methods applied to the case study introduced in 1.3.2 and 1.3.3. The guidelines were fairly vague when these analyses were conducted about the use of evidence synthesis and so the thesis also considers whether more advanced methods should be recommended (Chapters 2 and 8).

1.3.2   Home injuries and injury prevention

Unintentional (accidental) injuries are the leading cause of childhood death in industrialised countries, accounting for 40% of all child deaths between the ages of 1 to 14 years (Unicef 2001) and in most developed countries falls are the most commonly medically attended childhood injury (Peden, Oyegbite et al. 2008). In England, for children under the age of five, the majority of fatal and non-fatal injuries occur in the home with home injuries accounting for around half (48%) of childhood injuries presenting to A&E departments (Morrison, Stone et al. 1999, Audit Commission 2007, Office for National Statistics 2013b). Evidence indicates that children's risk of injury varies by a range of factors including age, gender, socioeconomic disadvantage, family type and size, maternal age, maternal educational level, ethnic group and neighbourhood of residence (Dowswell, Towner 2002). Falls are the 12th leading cause of disability-adjusted life years lost in 0-4 year olds (Peden, McGee et al. 2002) and can lead to long-term health, educational, social and occupational consequences in both the child and the carer (van der Sluis, Stewart et al. 2005).

It has been reported that the lack of evidence makes it difficult for policy makers and those designing and delivering interventions to know how best to design and deliver home safety interventions to increase home safety, reduce childhood injuries and address inequalities in child injury rates (Dowswell, Towner 2002). The Audit Commission Report in Child Safety, Better Safe than Sorry, 2007 was raised in many government reports but there was little evidence of systematic strategic approaches to develop, implement and monitor programmes to prevent unintentional injury in children (Audit Commission 2007). In 2007 a Cochrane review of home safety

education and provision of safety equipment for injury prevention was published with the main finding that home safety interventions provided in the home may reduce injury rates but more research was needed to confirm the findings. The results varied between studies but families receiving an intervention were more likely to possess safety equipment (Kendrick, Coupland et al. 2007a). NICE was commissioned to develop guidance on the prevention of unintentional injuries among children under 15 years of age (PH30) (NICE 2010b). The recommendations from this guidance included targeting home safety checks, providing information and advice to families, along with the provision and installation of safety equipment to the most disadvantaged families in the community.

Much of the evidence described above grouped together injuries, covered a wide range of ages of children, ignored risk factors such as gender, socioeconomic environment and ethnicity and grouped different interventions in one intervention group and one control group. The Keeping Children Safe at Home programme was proposed to develop a better understanding of how to prevent accidental injuries in pre-school children and is introduced in section 1.3.3.

1.3.3 Keeping Children Safe at Home Programme

The Keeping Children Safe at Home Programme ran from 2009-2014 and was funded by the National Institute for Health Research (NIHR). It was a collaboration between University and NHS researchers from Nottingham, Bristol, Norwich, Newcastle and Leicester. The overall aim of the programme was to increase evidence-based injury prevention by assessing the effectiveness and cost-effectiveness of interventions to prevent falls, poisonings and scalds, developing Injury Prevention Briefings (IPB) for cost-effective interventions and evaluating one of the IPB in Children's Centres. The Programme will be described in more detail in Chapter 3 and the focus of the thesis is on the effectiveness and cost-effectiveness analysis of interventions to prevent falls in children under 5 in the home, the package that I led, with the findings presented in Chapter 6 and Chapter 7 respectively.

1.4 Overview of the thesis

This chapter has provided an overview of the thesis, describing the aims, introducing the methods and the motivating example.

Chapter 2 describes the methods currently used in evaluating public health interventions and issues arising and includes the findings from a review of NICE Public Health Appraisals. Chapter 3 outlines the case study that motivated the thesis, accident prevention in the home in children under 5, and the evidence base focussed on in the thesis, i.e. evaluating the effectiveness and cost-effectiveness of interventions to increase the use of safety equipment to prevent falls in the Keeping Children Safe at Home Programme. Chapter 4 presents the methods and Chapter 5 the results of the effectiveness analysis applied to the case study evidence. Chapter 6 presents the methods and Chapter 7 the results for the cost-effectiveness analysis applied to the case study evidence. The final chapter, Chapter 8, will give an overview of the main findings and discuss any issues and opportunities for future research.

2 Evaluating the Effectiveness and Cost-effectiveness of Public Health Interventions

Public health interventions affect large population groups, can generate significant health benefits at individual and population levels but can also have harmful effects, and consume both financial and human resources (Rehfuess, Akl 2013). The evaluation of these interventions should be based on evidence that is informed, explicit, transparent and relevant to the population of interest. Evaluating public health interventions is far from straightforward and there is much discussion as to how evidence should be gathered, synthesised and used in decision making (Higgins, Green 2011). Developing recommendations relies on complex judgements on factors including magnitude of the health problem, benefits and harms, use of personnel and financial resources, transferability, as well as intervention acceptability and feasibility (Rehfuess, Akl 2013).

Systematic reviews, with and without meta-analyses, provide a transparent and consistent way of obtaining evidence of the effectiveness of interventions that minimises bias (Higgins, Green 2011). A decision modelling framework, that uses this effectiveness evidence, enables policy-relevant questions, such as which interventions represent the best use of scarce resources, to be answered (Drummond 2005).

This chapter will review the methods currently used in public health evaluations, the methodological challenges and introduce more sophisticated methods that can be used. It is based on a jointly authored paper "An exploration of synthesis methods in public health evaluations of interventions concludes that the use of modern statistical methods would be beneficial" (Achana, Hubbard et al. 2014) in which Achana conducted a systematic review of NICE public health appraisals and I discussed how more sophisticated methods of evidence synthesis may be applied to reviews of public health interventions to make the reviews more informative to decision makers.

2.1 Evaluations of public health interventions: guidelines

Guidelines specific to conducting reviews of public health and health promotion interventions were developed by the Cochrane Health Promotion and Public Health (HPPH) Field (now transitioned to the Cochrane Public Health Review Group (The Cochrane Collaboration 2018)) in 2005 and updated in 2007 (Armstrong, Waters 2007). Following a 2004 Department of Health report on improving health and reducing health inequalities in England which called for economic evaluations of public health interventions (Wanless 2004), the remit of NICE was expanded in 2006 to include the development of guidance for PH interventions based on sound appraisals of intervention effectiveness and cost-effectiveness (NICE 2012).

Both of the above sets of guidelines recommend a systematic review, with a narrative review and/or meta-analysis of primary research and previous reviews. The reporting of the systematic review should follow a set of guidelines such as the PRISMA checklist and flow chart (Moher, Liberati et al. 2009). Study quality and risk of bias should be assessed; for example RCTs should be assessed on allocation concealment, blinding of outcome assessment and follow-up in each arm and also for the balance of the distribution of confounders for non-randomised trials and controlled before and after studies. The NICE guidelines include a description of how health economic evidence should be collated and analysed if there is sufficient evidence to assess the cost-effectiveness of interventions using QALYs (NICE 2012).

2.2 Methodological challenges in public health intervention evaluation

2.2.1 Systematic reviews of evidence

A systematic review should identify all evidence on the effectiveness of an intervention. The Cochrane Handbook lists some key points in identifying evidence

from a systematic review of a public health intervention (Armstrong, Waters et al. 2011) and these along with some additional issues are discussed below.

Question of interest

There is often limited evidence on the actual question of interest, e.g. is there a reduction in the number of home accidents if a home safety intervention is offered? Some outcomes are fairly rare and make take many years to observe an effect of the intervention. Alternative, intermediate/surrogate, outcomes have to be identified, e.g. is there an increase in the possession of fitted safety equipment if a home safety intervention is offered? Assumptions are made that any intervention effect, an increase in the possession of fitted safety equipment, represents a reduction in the outcome of interest, a reduction in the number of home accidents.

Complex interventions

Public health interventions tend to be complex with multiple components. Studies may assess one, all or different combinations of these components. This makes it difficult to determine what specific intervention component or combination of components has had an effect.

The Medical Research Council (MRC) have produced guidance on "developing and evaluating complex interventions" (Craig, Dieppe et al. 2013, Craig, Dieppe et al. 2006) and their key dimensions of complexity are: the skill requirements of those delivering an intervention; the number of groups/organisational levels targeted by the intervention; the number and variability of outcomes; the degree of flexibility or tailoring of the intervention permitted.

Study designs used to evaluate interventions

Randomised controlled trials are the most useful form of evidence but are not always available for PH interventions due to issues such as feasibility and ethics. Cluster-randomised trials are increasingly used within the field of public health; where often interventions require their application at the cluster level (Donner 2004), e.g. an education intervention provided in a Health Centre reception will use the Health

Centre as the cluster. These cluster trials can contribute valuable evidence if a sufficient number of units are randomized to ensure even distribution of potential confounders among groups. For some interventions, the best available evidence may be from non-randomised studies and although they may be assessed as poor quality and have a high propensity for bias for a meta-analysis they can provide useful information in a narrative review in, for example, providing information for the development of future randomised trials (Armstrong, Waters et al. 2011). The use of different types of study design and control over bias gives rise to methodological heterogeneity.

Clinical Heterogeneity

Clinical heterogeneity might result from differences between the populations studied, the exact implementation of the interventions and control being compared or in the definition and assessment of the outcomes collected (Higgins, Thompson 2002).

Statistical Heterogeneity

Statistical heterogeneity is a consequence of clinical and methodological heterogeneity. There may be variation between studies in the underlying intervention effects being evaluated. It may be detected if variation in the results of the studies is above that compatible with chance alone. Statistical heterogeneity and how it is accounted for in meta-analyses is discussed throughout Chapter 4.

Evidence base

It is often more difficult to find the evidence for a systematic review of public health interventions as the literature can be widely scattered across multi-disciplinary areas.

Identifying health inequalities

It is usually of interest to investigate differential outcomes for different socioeconomic groups but there is often limited available information and also there is often lower participation of disadvantaged groups in research. It can be difficult to define to whom and to what degree the intervention was applied. On the other hand, this heterogeneity may increase applicability, as the populations and settings in which the

interventions will be used may be quite diverse, so this increases the likelihood that the evidence can be applied broadly.

Program by context interactions

It is often difficult to disentangle intervention effects from the influence of the context, for example social or economic environment, in which the intervention is implemented, e.g. providing free home safety equipment may have different effects depending on the ability of the householder to correctly install the safety equipment. Cluster-randomized designs can be useful in evening out important aspects of context as the intervention is allocated to a group/cluster of participants who may share a similar context, provided that the sample size is sufficient.

Sustainability

The long term viability of interventions is important to policy makers and funders, who are interested to know how health benefits are sustained beyond the intervention life. Long term impact is not usually assessed as funding usually limits studies to short term.

2.2.2 Meta-analysis

For reasons of study heterogeneity, many systematic reviews within the area of public health may not have a meta-analysis. A narrative review is used instead to describe the studies and is a useful insight into the available evidence from all types of studies and not just the highest quality evidence from RCTs. A narrative review combined with a meta-analysis, if applicable, ensures that all evidence on the effectiveness of an intervention is considered. NICE guidelines recommend that the characteristics and limitations of the data should be fully reported including the populations, intervention used, setting, sample size and any risk of bias. Reasons should be presented for why studies are not included in the meta-analysis. For studies included in a meta-analysis the level of heterogeneity between studies should be explored and considered in the analysis. Meta-regression and sub-group analyses should be used to explore the effect of varying populations and interventions, for example if studies have been conducted in areas with different levels of deprivation. Sensitivity analyses should be used to

explore the effect of any methodological heterogeneity, such as the length of follow-up or percentage responding. Any possible publication bias should be investigated (NICE 2012). Methods and issues in using meta-analysis are described in Chapter 4 in detail.

2.2.3   Economic modelling

Economic modelling with a decision modelling framework uses the effectiveness analysis described above combined with other relevant evidence and information on resource utilisation to derive comparative estimates of cost-effectiveness. They have the same issues and problems as described above in sections 2.2.1 and 2.2.2 and also have cost, consequence and equity issues. The methods and issues are described in Chapter 6.

2.3   NICE public health guidance and appraisals

NICE public health guidance makes recommendations for England on what is known from research and practice about the effectiveness and cost effectiveness of interventions and broader programmes, including the systems in which they are delivered, and the methods used to deliver them (NICE 2012). It can help the NHS and local authorities to meet standards for public health, and work towards the requirements of national planning and commissioning frameworks, enable national and local public sector organisations and partnerships to improve health and reduce health inequities, and support local authorities and schools in fulfilling their duty to promote the wellbeing of communities. Deliverers of public health improvement benefit from identified cost saving and the opportunity to re-direct resources. The first NICE Public Health Appraisals were published in 2006.

In terms of synthesising the evidence base NICE guidance 2012 states "Meta-analysis data may be used to produce a graph if the data (usually from RCTs) is sufficiently homogenous and if there is enough relevant and valid data from comparable (or the same) outcome measures. Where such data are not available, the synthesis may have

to be restricted to a narrative overview of individual studies looking at the same question".

In terms of economic evaluations NICE guidance 2012 states "Public health recommendations should be based on the balance between the estimated cost of each intervention and the expected health benefits (that is, recommendations should be cost effective). Recommendations should not be made on the basis of the total cost or the resource impact of implementing them. So, if the evidence suggests that an intervention provides health benefits and the cost per person of doing so is acceptable, it should be recommended, even if it would be expensive to implement across the whole population."

A review of all NICE PH appraisals published between March 1, 2006 and September 25, 2012 (Achana, Hubbard et al. 2014) is presented below. I was a joint author on this paper (presented in Appendix U). The aim was to identify what methods were being used for the synthesis of public health intervention effectiveness evidence. Thirty-nine completed PH appraisals were identified that contained 155 articles included in the review, with a median of 4 articles per appraisal.

The findings showed that effectiveness evidence was mostly synthesised using narrative reviews and only 9 (23%) of the 39 appraisals were informed by at least one systematic review with a meta-analysis. The other appraisals refrained from a meta-analysis citing a lack of RCTs or heterogeneity in the study designs. Those appraisals that did conduct a meta-analysis used the simplest methods (Table 2-1); a fixed or random effects pairwise meta-analysis (section 4.3) thus restricting the scope of the analysis and how the findings can be used to inform policy decisions; in some of these cases the use of network meta-analysis could have been explored as the interventions contained multiple components. Decision models were often not informed by the meta-analysis results and were based on the findings from a single study. The paper concluded that more advanced techniques in evidence synthesis methodology can be used to address some of the challenges and opportunities in the appraisal of PH interventions, including the use of sub-group analyses, meta-regression incorporating individual participant data (IPD) and network meta-analysis to compare more than two

interventions and rank these interventions in order of effectiveness (Achana, Hubbard et al. 2014). The aim of these methods would be to identify which intervention is most effective and to whom. These methods are all described in detail in Chapter 4 and the example presented in the paper (Appendix U) is described in detail in Chapter 5.

2.3.1 NICE evaluation of interventions to prevent unintentional injuries in the home in children

The only other evaluation of the effectiveness and cost effectiveness of interventions identified in the area of preventing accidents, including falls, in the home in children was reported in NICE PH29 and PH30 (Pitt, Anderson et al. 2009). PH29 presents strategies to prevent unintentional injuries among children and young people aged under 15 and was published in Nov 2010 then reviewed but not updated Feb 2013. PH30 presents guidance on preventing unintentional injuries among under-15s in the home. In these reviews the authors conducted systematic reviews but no meta-analysis due to heterogeneity of interventions and methods. They conducted a cost-effectiveness analysis of generic home safety interventions versus no intervention for all home safety accidents irrespective of the mechanism or cause of injury. Falls, scalds, poisonings, etc. were not separated. Their recommendations include:

- Incorporate unintentional injury prevention within local and national policy and strategies for children and young people's health and wellbeing.

- Installation and maintenance of permanent safety equipment in social and rented dwellings and home safety assessments.

Table 2-1 Review of quantitative methods used to synthesise public health evidence for NICE public health appraisal

Appraisal title	Systematic review report title	Included RCTs only	Main outcome	Description of main outcome	Outcome measure: statistic	Type of synthesis	Model type	Lumping[1] of interventions	Presentation of results	Assessed publication bias	Software	Used result of M-A in decision model
Prevention of sexually transmitted infections and under 18 conceptions (PH3)	Review 2 - Review of evidence for the effectiveness of screening for genital chlamydial infection in sexually active young women and men	No	Intermediate	Uptake of proactive chlamydia screening using home-collected specimens	Screening response rate (%)	M-A	Random effects	No	FP/Txt	No	RevMan, Stata	No
School-based interventions on alcohol (PH7)	Alcohol and schools: effectiveness and cost-effectiveness review	No	Final	Alcohol use	Weighted mean difference	M-A	Random effects	Yes	FP/Txt	No	Not stated	No
Smoking cessation services (PH10)	Cut down to quit' with nicotine replacement therapies	Yes	Final	6 or more months' sustained abstinence	Relative risk & Cohen's d	M-A	Random effects	Yes	FP/T/ Txt	No	RevMan	Yes
Smoking cessation services (PH10)	Final report	No	Final	6 or months' sustained abstinence	Cohen's d	M-A	Fixed & random effects	Yes	FP/T/ Txt	No	RevMan	No
Social and emotional wellbeing in primary education (PH12)	Teesside review	Yes	Intermediate	Social problem solving	Standardised mean difference	M-A	Random effects	Yes	FP/T	No	RevMan	No
Management of long-term sickness and incapacity for work (PH19)	PH19 Management of long-term sickness and incapacity for work: Economic analysis report	No	Yes	Number returning to work following sickness	Relative risk	M-A	Random	Yes	FP/T/Txt	No	Revman	Yes
School-based interventions to prevent smoking (PH23)	School-based interventions to prevent smoking: quantitative effectiveness review	Yes	Final	smoking uptake	Odds ratio	M-A	Random effects	Yes	FP/Txt	Yes	Stata	Yes
Weight management before, during and after pregnancy (PH27)	Weight management before, during and after pregnancy: evidence review	No	Intermediate	Number exceeding IoM[2] guidelines for healthy weight gain	Relative risk	M-A	Random effects	Yes	FP/T/ Txt	No	RevMan	No

Table 2-1 (continued) Review of quantitative methods used to synthesise public health evidence for NICE public health appraisal

Appraisal title	Systematic review report title	Included RCTs only	Main outcome	Description of main outcome	Outcome measure: statistic	Type of synthesis	Model type	Lumping[1] of interventions	Presentation of results	Assessed publication bias	Software	Used result of M-A in decision model
Preventing type 2 diabetes - population and community interventions (PH35)	PH35 Preventing type 2 diabetes - population and community interventions: report on cost-effectiveness evidence and methods for economic modelling	No	Intermediate	Body mass index	Weighted mean difference	M-A	Not reported	Yes	T/Txt	No	Not reported	Yes
Preventing type 2 diabetes - risk identification and interventions for individuals at high risk (PH38)	Prevention of type 2 diabetes: systematic review & meta-analysis of lifestyle, pharmacological and surgical interventions	Yes	Final	Reduce progress to diabetes for people with IGT	Hazard ratio	M-A & NMA	Random effects	No	FP/TxT	No	RevMan (M-A) WinBUGS (NMA)	Yes

Presentation of results (FP = Forest plot, T=Table, Txt=Text), M-A = pairwise meta-analysis, NMA = network meta-analysis;

1 = lumping is a term used in the literature[19, 20] to described the tendency to aggregate or treat seemingly similar but disparate /different interventions as one intervention group in order for example to facilitate inclusion of many studies in a meta-analysis. A classic example is treating different doses of a drug as if they were the same treatment

2 = American Institute of Medicine (IOM) Guidelines on Weight Management in Pregnancy

2.4   Summary

This chapter reviews some of the guidance on conducting evaluations of public health interventions, including NICE guidance, Cochrane Public Health Review guidance and MRC Process Evaluation of Complex Interventions. Some methodological challenges of public health intervention evaluation are identified. A review of methods used in NICE public health appraisals concludes that more advanced methods of evidence synthesis can be used to address some of the challenges. The next chapter and the ones following will describe the Keeping Children Safe at Home Programme and the methods of effectiveness and cost-effectiveness used to address the challenges described in this chapter.

3 Case Study: Accident and Injury Prevention in Children under Five in the Home

As stated in the introduction unintentional (accidental) injuries are the leading cause of childhood death in industrialised countries, accounting for 40% of all child deaths between the ages of 1 to 14 years (Unicef 2001) and in most developed countries falls are the most commonly medically attended childhood injury (Peden, Oyegbite et al. 2008), hence they are a major public health challenge. In this chapter unintentional home injuries will be discussed further along with some of the issues in providing and evaluating interventions to prevent the injuries. There will be a summary of the Keeping Children Safe at Home NIHR Programme and the Cochrane Review on this topic and the NICE guidance. The thesis focusses in particular on interventions to prevent falls as this is the data that I analysed and my results are presented in both the Cochrane Review (Kendrick, Young et al. 2012) and Programme report (Kendrick, Ablewhite et al. 2017).

3.1 Unintentional (accidental) injuries in the home in children under five

In England, death and admissions to hospital for children are higher in the under 5's than any other age group. This age group is unique in terms of rapid growth and developmental changes, which influence risk for a number of specific causes of injury (Spady, Saunders et al. 2004). Each year in England, an average of 60 children die from injuries in and around the home and there are approximately 40,000 emergency hospital admissions and 450,000 visits to Accident and Emergency (A&E) departments (Public Health England February 2017, Office for National Statistics 2013b, Health and Social Care Information Centre 2013).

The most common unintentional injuries in the home identified by Public Health England (PHE) include falls; choking/suffocation/strangulation; burns and scalds; poisoning; drowning; fire. The common types of home injuries in children under 5 reporting to hospital include: cuts/other open wounds; other soft tissue injuries;

bruises/contusions; concussion; grazes/splinters; burns/scalds; dislocations/sprains; poisonings (Walker 2010). The Home Accident Surveillance System (HASS) database from 2002 detailing home accidents that caused a serious enough injury to warrant a visit to hospital (Helen Shaw 2014) figures were used to assess that 80% of these injuries are slight with 20% being serious. Some of the serious injuries require only a short recovery period (92% recovering in less than a year) with little associated cost (estimated cost £2,494 at June 2009 prices), while others require long term medical support resulting in high costs (less than 1%) (Walker 2010, Department for Trade and Industry. 2003). Serious long-term childhood injuries place burdens on the NHS, other care agencies as well as the child and their carers. Estimated costs (including lost output, value of avoidance of injury, medical and support) were calculated as £33,200 for a serious injury and £10,600 for a non-fatal hospital treated injury at June 2009 prices (Walker 2010). Little is known about the minor injuries sustained and not reported to hospital.

Falls are the main cause of injury-related admissions for under-fives (around 20,000 per year) with most hospital admissions resulting from falls from furniture, stairs and steps (around 10,500 per year) (Health and Social Care Information Centre 2013). Deaths from falls, or any other unintentional accidents in the home, are rare with about five deaths from falls a year. Children under one mostly fall from beds or high chairs or while being carried. As the child gets older the risks change due to the child becoming more mobile and independent (Public Health England February 2017). Falls are the 12th leading cause of disability-adjusted life years lost in 0-4 year olds (Peden, McGee et al. 2002) and can lead to long-term health, educational, social and occupational consequences in both the child and the carer (van der Sluis, Stewart et al. 2005).

Overall rates of death from injury and poisoning in children have fallen in England and Wales, except for children in families living in socioeconomic disadvantage. This has been shown particularly for children living in households in which no adult is in paid employment (Edwards, Roberts et al. 2006). Evidence indicates that children's risk of injury varies by a range of factors including age, gender, socioeconomic disadvantage, family type and size, maternal age, maternal educational level, ethnic group and

neighbourhood of residence (Dowswell, Towner 2002). These factors should be considered in any interventions aimed at preventing unintentional accidents in the home.

3.2   Unintentional injury prevention interventions for the home

Home safety injury prevention interventions can be provided by health or social care professionals, schools, voluntary organisations and other organisations to individual children or families or groups of children or families. The aim of an intervention is to increase the use of prevention practices and equipment and hence reduce injury rates. The interventions most commonly used are home safety education and the provision of safety equipment (section 3.5.1). Home safety education can take the form of generic information or personalised information, either paper based or online, or one-to-one face-to-face with a healthcare or social care professional. Home safety equipment can consist of single items to prevent specific accidents, such as a safety gate to prevent falls down stairs, window locks to prevent falls from windows or cupboard locks to prevent poisonings, or a package of equipment to prevent multiple accident types. The use of prevention practices and equipment can be recommended through education or a home safety inspection, and equipment can be offered free of charge or discounted using a voucher scheme.

There has been concern that the uptake of interventions, both educational and equipment, varies between socioeconomic groups but there is little evidence available on this (Towner, Dowswell et al. 2005). Interventions have been aimed at specific groups, for example Watson et al 2005 offered free equipment to families with children under 5 living in deprived areas (Watson, Kendrick et al. 2005). Associations between social deprivation and increased risks of childhood injury may be linked to several underlying factors, including: overcrowded housing conditions; hazardous environments; single-parenthood; unemployment; a relatively young maternal age; a relatively low level of maternal education; stress and mental health problems on the part of caregivers; lack of access to health care .

Different types of educational information is offered and there is lots of advice online. Figure 3-1 shows the NHS Choices website on baby and toddler safety giving guidance to parents on preventing falls down stairs in babies and toddlers. The NHS recommendation is for two safety gates on the stairs, top and bottom, but much of the evidence on intervention effects reports only the possession of a single fitted safety gate on the stairs, not two, and not if the gate is used appropriately and kept closed at all times (Young, Wynn et al. 2013).

Interventions to prevent home injury are very varied, can be aimed at preventing multiple injuries and be offered by a range of providers. This makes the evaluation of their effectiveness and cost-effectiveness difficult. Safety gates can also be used in doorways to prevent a child from entering a room, such as the kitchen, where they may be a risk of poisoning or scald injuries. Some falls prevention measures, such as recommending that baby walkers are not used, do not require any safety equipment although a safety gate can prevent falls down stairs or steps in a baby walker. In 2005 Kendrick et al reported that baby walkers are used in the UK by 50% of children aged between 3 and 12 months and parents report that between 8% and 12.5% of children using walkers suffer an injury in their walker, with around 3,000 attending A&E departments reporting head injuries, lacerations, burns and scalds from stairway falls, tip overs and burns (Kendrick, Illingworth et al. 2005).

Other falls prevention interventions were aimed at preventing falls: in the bath by promoting the use of bath mats or decals in the bath, promoting not leaving children unattended on high surfaces; using window locks to prevent falls from windows; using safe rugs with non-slip linings to prevent trip falls.

NHS Choices on baby and toddler safety

https://www.nhs.uk/conditions/pregnancy-and-baby/baby-safety-tips/?#falls-in-toddlers

Here are some injury prevention tips for parents of crawling babies:

- Fit safety gates at the top and bottom of the stairs to stop a baby from climbing stairs or falling down them. Close the gates properly after you go through them.

- Continue to use safety gates at the top and bottom of the stairs until your infant is at least two years old.

- Teach your child how to climb stairs but never let them go up and down on their own

Figure 3-1 NHS Choices on baby and toddler safety – Safety gates

3.3   Child injury prevention: policy and measures

It is difficult to obtain evidence that injury prevention interventions actually prevent injuries, sections 1.3.1 and 3.5.1; the evidence focusses on interventions to increase the use of safety equipment that should prevent the unintentional accidents and injuries. A lack of evidence makes it difficult for policy makers and those designing and delivering interventions to know how best to design and deliver home safety interventions to increase home safety, reduce childhood injuries and address inequalities in child injury rates (Dowswell, Towner 2002). The Audit Commission Report in Child Safety, *Better Safe Than Sorry*, in 2007 reviewed the activities to prevent unintentional injuries in the home especially in children under five. The report highlighted that there was little evidence of systematic strategic approaches to develop, implement and monitor programmes to prevent unintentional injury in children and made recommendations to the government and local organisations to follow evidence-based guidance and to commission NICE to develop guidance (Audit Commission 2007). In 2010 NICE developed a series of guidance documents, PH29 and PH30 (described in section 2.3.1), on the prevention of unintentional injuries in children aged under 15 (NICE 2010b).

An EU report in 2012 concluded that there was scope for improvement in implementing child injury prevention measures in England and that unintentional

injury was the leading cause of health inequality in deaths in children (European Child Safety Alliance and EuroSafe 2012). In 2013 Public Health England (PHE) was established with one of the targets being to reduce unintentional and deliberate injuries for the 0-4 year age group (Public Health England 2017).

3.4   The Cochrane review of home safety education and provision of safety equipment for injury prevention

In 2007 a Cochrane review of home safety education and provision of safety equipment for injury prevention was published (Kendrick, Coupland et al. 2007b) to evaluate the effectiveness of home safety education with or without low cost, discounted or free equipment in reducing child injury rates or increasing practices in homes to prevent childhood injuries, and to evaluate this effectiveness by social group. Outcome measures assessed were self-reported or medically attended injury following an unintentional injury in the home, possession and use of home safety equipment and safety practices to prevent injuries in the home. The injury prevention measures were grouped into categories: thermal; poisonings; falls; electrical; lacerations and bruising; suffocation; drowning. The authors of the review contacted the authors of the papers identified in the review to ask if they would be willing to provide the individual participant data (IPD) for their studies. IPD meta-analyses are described as the gold standard with many advantages over aggregate/summary data meta-regressions (Stewart 1995). IPD meta-analysis is described in more detail in section 4.9. Studies in this review included individually and cluster randomised controlled trials (RCTs), non-randomised controlled trials and controlled before and after (CBA) studies (studies with a concurrent control group which have data collected on outcome measures at baseline and follow-up). Participants were children and young people (aged 19 years and under) and their families. Socio-economic characteristics that were thought to be associated with an increased risk of childhood injury were also recorded and these included child age, gender, ethnic group, family type (single or two parent), housing tenure and parental unemployment. Meta-

analysis and meta-regression analyses were conducted comparing the intervention group with the control arm.

The review was updated in 2012 as part of the Keeping Children Safe at Home Programme described in section 1.3.3. Part of this thesis is based on my effectiveness analyses conducted for the falls injuries included in the review update (Kendrick, Young et al. 2012) and are described in sections 4.1 to 4.9 and the results presented in sections 5.1 to 5.7.

3.5   The Keeping Children Safe at Home (KCSH) Programme

The KCSH Programme was a multicentre collaborative research programme to reduce childhood injuries funded by the National Institute for Health Research (NIHR). Its aim was to increase evidence–based NHS injury prevention by assessing the cost-effectiveness of interventions to prevent falls, poisonings, and scalds, developing Injury Prevention Briefings (IPBs) for cost-effective interventions and evaluating the implementation of one IPB in Children's Centres. Figure 3-2 illustrates the different components of the programme.

A series of case-control studies (Figure 3-2: Question 1) were undertaken to assess the effectiveness of a range of potential interventions to prevent falls, poisonings and scalds injuries. These types of injuries are fairly rare so RCTs would need to be unfeasibly large to show a significant effect. The results from these case-control studies, along with the results from a prospective study to investigate the NHS costs and consequences of falls, poisonings and scald injuries (Figure 3-2: Question 2), a survey to identify what injury prevention work was being undertaken (Figure 3-2: Question 3), and a systematic review (Figure 3-2: Question 5 Study H), fed into effectiveness and cost-effectiveness analyses based on decision analytic models developed separately for each injury type (Figure 3-2: Question 5). The findings were incorporated into the production of Injury Prevention Briefings ((Kendrick, Ablewhite et al. 2017).

The effectiveness and cost-effectiveness analysis reported in this thesis aimed to answer part of Question 5 in Figure 3-2; that is, the effectiveness analysis of interventions to prevent unintentional falls injuries in the home, and the development and evaluation of decision analytic models to evaluate the most cost-effective strategies. Other members of the programme team, including other PhD students, looked at other injury outcomes. This thesis focuses on the outcomes related to falls injuries: possession and use of home safety equipment (stair gates, window locks, non-slip bath mats, safe rugs) and safety practices (use of baby walkers and not leaving children alone on a high surface). Some of the methodology described and applied in this thesis was developed as part of the programme.
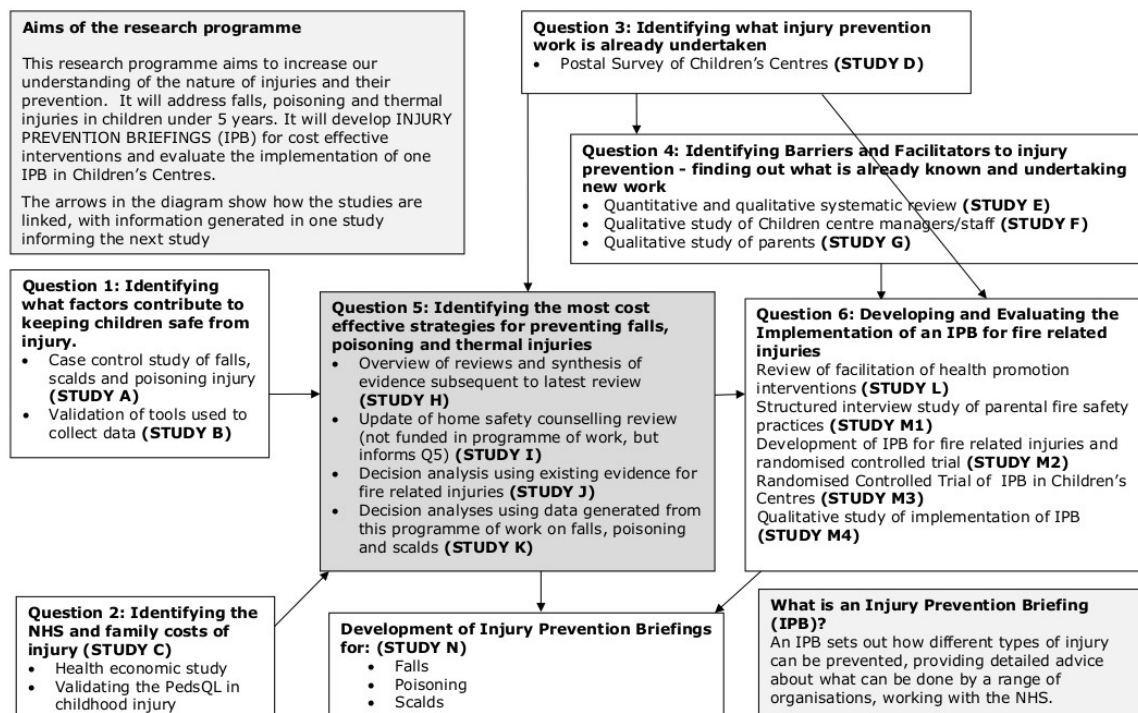
Figure 3-2 NIHR Keeping Children Safe at Home Programme: Its structure, the research questions asked, and the relationships between the different components

3.5.1 Systematic review of studies to prevent falls injuries

Sixteen studies were identified in the systematic review for the falls outcomes, Figure 3-3, published in a review paper, reported in the updated Cochrane review and the KCSH programme report (Young, Wynn et al. 2013, Kendrick, Young et al. 2012,

Kendrick, Ablewhite et al. 2017). Primary studies, overviews of reviews, systematic reviews and meta-analyses of experimental and controlled observational studies reporting interventions aimed at primary or secondary prevention of falls at home among children were eligible. Details of the eligibility and search criteria, the risk of bias analysis and details of the studies identified, included and excluded, are described in the review paper (Young, Wynn et al. 2013). Only 3 primary studies reported interventions to prevent falls or fall injuries. Other studies were identified that reported interventions to promote possession and use of safety equipment aimed at reducing falls injuries: safety gates (16 studies); non-slip bathroom items (5 studies); window safety devices (10 studies); furniture corner covers (4 studies); high chair harnesses (2 studies). Studies were also identified reporting interventions to: reduce baby walker use (9 studies); promote stairway safety (6 studies); reduce tripping hazards (4 studies); prevent children being left unattended on high surfaces (3 studies). There were 6 studies that reported a range of falls prevention practices through a falls prevention score. Some of these studies were only included in a narrative review (Young, Wynn et al. 2013) but 16 studies, Figure 3-3, were identified for inclusion in the meta-analyses reported in Chapter 5. IPD was obtained for 13 of these 16 studies and for the other three aggregate data was extracted from the published article. The characteristics of these studies are summarised in Table 3-1.

There were 12 RCTs, three non-randomised controlled trials and one controlled before-and-after study. For RCTs, allocation concealment, blinding of outcome assessment and completeness of follow-up (80% or more in both intervention arms) were used as markers of trial quality. For non-randomised studies blinding of outcome assessment, completeness of follow-up (80% or more in both intervention arms) and assessment of the distribution of confounders (baseline socio-demographic or economic characteristics, safety practices or injury rates) were used as markers of quality. Studies were considered to be balanced in terms of confounders if the prevalence of these did not differ by more than 10% between the intervention arms. Study quality assessment is given in Table 3-2.

Intervention strategies identified in the 16 studies in the review included: usual care, education, free or low cost safety equipment, home safety inspection, and fitting. The

control intervention from individual studies was classed as usual care if the study reported the control group as 'usual safety education', 'standard safety practice or advice' or 'no safety education' (i.e. no or do-nothing intervention control groups). Education was taken to mean that provided in addition to usual or standard safety education delivered by face-to-face contact with a trained health professional or by an educational leaflet. Free or low cost safety equipment included the provision of falls-related equipment such as safety gates, window locks, non-slip bath mats; some interventions also provided other home safety equipment not aimed at falls prevention (e.g. smoke alarms, cupboard locks etc). Home safety inspection refers to home visits including inspections carried out by trained health and other professionals. Finally "fitting" refers to installation of safety equipment by a trained professional.

Table 3-1 presents details of the interventions and the numbers in each arm for the 16 studies included in the effectiveness analyses in Chapter 5 and in the cost-effectiveness analysis in Chapter 7. Most studies included a package of multiple intervention components to prevent multiple home injuries, for example in Phelan et al, 2010, the intervention included home safety inspection, provision and fitting of free safety equipment when child is aged 3-6 months (stair gates, non-slip matting under rugs, window guards, repair of stair handrails, cupboard/drawer locks, door knob covers, storage bins, socket covers, smoke detectors, CO detectors, stove guards, stove locks) and safety advice handout. For the analyses presented in Chapters 5 and 7 these studies are considered solely as including interventions to increase the use of safety equipment or promote safe behaviour practices to prevent falls injuries and their wider effect on other injury prevention will be discussed in Chapter 8.

Table 3-3 describes the demographic and socio-economic characteristics of the 16 studies, averaged over the intervention groups. Not all studies report all six characteristics. Eight (50%) of the 16 studies were based in the USA and four(25%) in the UK. The average age of the child ranges from 8 months to 31 months and no studies were focussed on just male or female children. Some studies were aimed at populations with high percentages of families residing in rented accommodation, single parent families, black or minority ethnic groups or at least one parent unemployed, whereas others have very low percentages (Table 3-3). These six

characteristics are known to be risk factors for childhood injuries in the home (section 3.1) and hence were considered in the analyses in Chapters 5 and 7.

Searches for the systematic overview of reviews

Screened for inclusion:
10226 from bibliographic databases
27 Conference abstracts
24 from hand searching "Injury Prevention"
22 from reference lists
13 Already had
203 Other electronic sources

147 Papers assessed for inclusion

122 Excluded papers and reasons
93 Study design
1 Participants
16 Intervention
2 Outcomes
3 Study design of included studies
7 Paper unobtainable
12 Duplicates

13 Included reviews
1 Meta-analysis
12 Systematic reviews
Sourced from
6 Already had
4 Bibliographic databases
1 Hand searching
2 Other electronic sources

24 Included primary studies identified from reviews

Searches for the systematic review

Screened for inclusion:
15206 from bibliographic databases
87 Conference abstracts
11 from hand searching "Injury Prevention"

93 Papers assessed for inclusion

76 Excluded papers and reasons
25 Study design
16 Intervention
4 Outcomes
14 Paper already identified
17 Paper unobtainable/no published paper
12 Duplicates

5 Included primary studies from systematic review
All 5 sourced from bibliographic databases

29 Primary studies included from all searches
20 RCTs
5 NRCTs
3 CBAs
1 Observational design

13 Excluded studies from Meta-Analyses
2 Safety equipment outcomes reported by fewer than 3 studies
2 Report medically attended falls injuries rather than safety equipment
2 Prevention practices reported by fewer than 3 studies
3 Report falls hazards prevention scores
3 No effect size or p-values
1 Study design

16 Included studies in Meta-Analyses
12 Possession of a fitted stair gate
9 Possession or use of a baby walker
6 Possession of window locks
3 Child not left on a high surface
6 Possession of bathmats /decals

Figure 3-3 PRISMA flow chart for the systematic overview of reviews and systematic review of primary studies

Table 3-1 Characteristics of studies included in the meta-analyses for the falls prevention interventions

Author	Title	Population	Intervention (i) and Control (C)	Number in each arm	Outcomes and follow up period	Data included in meta-analyses
STUDIES INCLUDED IN META-ANALYSES						
RANDOMISED CONTROLLED TRIALS						
Phelan, USA, 2010	A randomized controlled trial of home injury hazard reduction: The HOME Injury Study	Pregnant women, aged 18 years and over, < 19 weeks gestation, attending pre-natal practices in Cincinnati, USA.	I= home safety inspection, provision and fitting of free safety equipment when child is aged 3-6 months (stair gates, non-slip matting under rugs, window guards, repair of stair handrails, cupboard/drawer locks, door knob covers, storage bins, socket covers, smoke detectors, CO detectors, stove guards, stove locks) and safety advice handout. C= prior to child's birth family given targeted home repairs to control lead hazards (e.g. paint stabilisation, water filters)	I=181 C=174	Outcomes measured at 12 and 24 months: Falls (use of baby walker, window locks, safety gate, non-slip bath mat) plus Thermal injuries, Poisoning, Electrical, Lacerations and bruising, Suffocation, Medically attended injuries (Unpublished data)	IPD
Babul, Canada, 2007	A randomized trial to assess the effectiveness of an infant home safety programme	Parents of new born infants at a general hospital serving mainly urban or suburban communities	I_1= Home visit + home safety inspection + free safety kit (smoke alarm, coupon for 50% discounted stair gate, corner cushions, cabinet locks, blind cord windups, water temperature card, door stoppers, socket covers, poison control centre sticker + safety brochure + home safety checklist for parents) I_2 = free safety kit (as above) C = usual care	I_1=202 I_2=206 C=192	Outcomes measured at 12 months of age: Falls (use of baby walker, left child alone on high surface) plus Thermal injuries, Suffocation, Poisonings, Drowning, Medically attended injuries	Summary
Kendrick, UK, 2005	A randomised controlled trial of the effectiveness of an educational package in reducing baby walker use	Women of at least 28 weeks gestation registered at participating general practices	I = midwife and health visitor advice to discourage walker use, information cards, fridge magnets, checklists for use in child health surveillance visit at 3-4 months. Encouraging use of stair gates and fire guards amongst walker users. C = usual care	I = 539 C = 635	Outcomes measured when child 9 months of age: Falls (use of baby walker, safety gate) plus Thermal injuries	IPD

Author	Title	Population	Intervention (i) and Control (C)	Number in each arm	Outcomes and follow up period	Data included in meta-analyses
McDonald, USA, 2005	Evaluation of kiosk-based tailoring to promote household safety behaviours in an urban pediatric primary care practice	Parents of children aged 6 weeks to 24 months, attending well child clinic	I = tailored safety advice in well child clinic + feedback report to paediatrician to encourage safety counselling + information on safety equipment savings at child safety centre C = usual care	I = 70 C = 74	Outcomes measured over 1 month: Falls (use of safety gate) plus Poisoning, Thermal injuries	IPD
Watson, UK, 2005	Providing child safety equipment for the prevention of injuries: a randomised controlled trial	Families with children < 5 years on caseloads of health visitors in deprived areas	I = health visitor safety consultation, free fitted safety equipment (stair gates, fire guards, cupboard and drawer locks, smoke alarms, window locks) C = usual care	I = 1711 C = 1717	Medically attended injuries measured over 24 months Other outcomes measured at 12 & 24 months : Falls (use of safety gate, window locks) plus Lacerations and bruising, Poisoning, Thermal injuries, (unpublished data)	IPD
Posner, USA, 2004	A randomised clinical trial of a home safety intervention based in an emergency department setting	Caregivers of children < 5 years attending ED for home injury	I = home safety counselling by trained lay personnel, home safety kit (cupboard and drawer locks, socket covers, bath tub spout covers, non-slip bath decals, bath water thermometer, poison control centre number stickers, free small parts tester) + home safety literature C = home safety literature	I = 69 C = 67	Outcomes measured over 10 weeks: Falls (use of baby walker, safety gate, non-slip bath decals, never leaves child alone on high surface) plus Lacerations and bruising, Drowning, Poisoning, Suffocation, Thermal injuries, Electrical injury, Drowning, Safety score (unpublished data)	IPD
Sznajder, France, 2003	Home delivery of an injury prevention kit for children in 4 French cities: a controlled randomised trial	Socio-economically disadvantaged families, with medical or psychological difficulties which place them at high risk	I = home safety counselling by health professionals, safety leaflets, free home safety kit (cupboard and drawer locks, door handle covers, furniture corner protectors, socket covers, non-slip bath mat, fitted smoke alarm, poison control centre number stickers)	I = 50 C = 50	Outcomes measured over 2 months: Falls (use of baby walker, safety gate, non-slip bath mats, high chair safe, floor safety, risk of falling from window)	IPD

Author	Title	Population	Intervention (i) and Control (C)	Number in each arm	Outcomes and follow up period	Data included in meta-analyses
			C = home safety counselling + safety leaflets		Lacerations and bruising, Electrical injury, Poisoning, Suffocation, Thermal injuries	
Gielen, USA, 2002	Effects of improved access to safety counselling, products and home visits on parents' safety practices	1st and 2nd year paediatric residents and their patient-parent dyads Low income population of parents of children aged 0-6 months	I = safety counselling by paediatricians + referral to child safety centre + home visit C = safety counselling by paediatricians + referral to child safety centre	I = 94 C = 93	Outcomes measured over 12 months: Falls (use of safety gate) plus Poisoning, Thermal injuries, Safety score	Summary
Hendrickson, USA, 2002	A safety home visit in a low income community	Mothers with children aged 1-4 years, predominantly Mexican/Mexican American	I = safety counselling from researchers, plus identification of home hazards + safety education + provision of safety equipment (door knob covers, smoke detectors or new batteries if smoke alarm already in situ, fire extinguisher, cabinet latches and outlet covers) C = none of the above	I = 41 C = 41	Outcomes measured over 6 weeks: Falls (use of baby walker, safety gate, non-slip bath mats, high chair safe, floor safety, window locks, hand rail on stairs) plus Electrical injury, Poisoning, Suffocation, Thermal injuries (unpublished data)	IPD
Nansel, USA, 2002	Baby be Safe	Parents of children aged 6-20 months attending well child check	I = computer generated tailored safety advice in well child clinic C = computer generated generic safety advice in well child clinic	N = 213 at baseline, not specified by treatment arm At follow up: I = 85 C = 89	Outcomes measured over 3 weeks: Falls (use of baby walker, safety gate) plus Drowning, Poisoning, Thermal injuries, Safety scores (unpublished data)	IPD
King, USA, 2001	Effectiveness of home visit to prevent childhood injury	Children <8 years attending A&E for injury or medical complaint	I = home safety inspection + information on correcting any deficiencies, discount vouchers for safety equipment, demonstrations of use of safety devices + information on preventing specific injuries provided by researcher	I = 601 C = 571	Medically attended injuries measured over 36 months. Other outcomes measured over 12 months: Falls (use of baby walker, safety gate, safe windows)	Summary

Author	Title	Population	Intervention (i) and Control (C)	Number in each arm	Outcomes and follow up period	Data included in meta-analyses
			C = home safety inspection & safety pamphlet.		plus Poisoning, Suffocation, Thermal injuries	
Clamp, UK, 1998	A randomised controlled trial of GP safety advice for families with children under 5	Families with children< 5 years registered at one GP surgery	I = general practitioner safety advice, leaflets & low cost safety equipment (smoke alarms, window locks, cupboard and drawer catches, socket covers, door slam devices, fire guards, stair gates) C = usual care	I = 83 C = 82	Outcomes measured over 6 weeks: Falls (use of safety gate, window locks) plus Lacerations and Bruises, Electrical injury, Poisonings, Thermal injuries	IPD
NON RANDOMISED CONTROLLED TRIALS						
Nansel, USA, 2008	Preventing unintentional paediatric injuries: a tailored intervention for parents and providers	Parents of children aged ≤ 4 years attending well child visits at 3 paediatric clinics with mainly low to middle income patients	I_1 = tailored injury prevention education I_2 = tailored injury prevention education and provider tailored information C = general education	I_1= 107 I_2= 100 C= 98	Outcomes measured at 1 month: Falls (use of baby walker, safety gate, never leaves child on high surface) plus Thermal injuries, Poisoning, Electrical injuries, Drowning (unpublished data)	IPD
Tan, Singapore, 2004	Effectiveness of nurse counselling in discouraging the use of infant walkers	Caregivers and infants aged 4-5 months attending three health clinics	I = structured nurse counselling + leaflets aimed at discouraging walker use C_1 = no nurse counselling C_2= no nurse counselling and no baseline data collection	n = 716 at baseline, not specified by treatment arm. At follow up: I = 228 C_1 = 214 C_2 = 271	Outcomes measured when child 9 months of age: Falls (use of baby walker) plus Baby walker injuries (unpublished data)	IPD
Kendrick, UK, 1999	Preventing injuries in children: cluster randomised controlled trial in primary care	Children aged 3-12 months registered at 36 GP practices	I = health visitor safety advice at child health surveillance, low cost equipment (stair gates, fire guards, cupboard and drawer locks, smoke alarms), home safety checks and first aid training C = usual care	I = 1100 C = 1019	Medically attended injuries and other outcomes measured over 25 months: Falls (use of safety gate, window locks, safe rugs)	IPD

Author	Title	Population	Intervention (i) and Control (C)	Number in each arm	Outcomes and follow up period	Data included in meta-analyses
					plus Poisoning, Thermal injuries, Drowning, Electrical injury, Lacerations and bruising, Suffocation (unpublished data)	
CONTROLLED BEFORE-AND-AFTER STUDIES						
Petridou, Greece, 1997	Controlled evaluation of a community injury prevention project in 2 Greek islands	Population of two Greek islands, Naxos (intervention) and Spetses (control)	I = community intervention including safety seminars for parents, workshops with teachers promoting school safety, courses with primary and secondary school children on safety and resuscitation, leaflets; plus focused intense intervention: lay home visitors, weekly visits to discuss home safety in households with children (≤18 years) or older people (≥ 65 years) C = none of the above	I = 172 households C = 177 households	Outcomes measured over 20 months: Falls (use of baby walker, safe stairs, balconies) plus Electrical injury, Poisoning, Thermal injuries, Hazard score (unpublished data)	IPD

Table only includes outcomes reported for children aged 0-19 years, outcomes reported for wider age groups, including children, but not reported separately for children are excluded. Outcomes reported by controlled before and after studies for the follow up period only are excluded. Studies reporting medically attended injuries are also included in tables relating to specific injury mechanisms if they reported injuries related to that mechanism.

Where a study has more than one article, only the title of one article is given, but references are provided for all relevant studies.

Table 3-2 Characteristics of included studies with respect to quality criteria

1st Author	Year	Design	Allocation conceal-ment adequate	Outcome assessors blinded	Outcomes measured on 80% of participants in each arm	Treatment arms balanced for confounders	Comments, including allocation level for cluster allocated studies
Phelan, USA	2010	RCT	yes	no	yes	n/a	
Nansel, USA	2008	Non-RCT	n/a	no	no	no	Participants randomly assigned to generic advice group and to tailored advice group, then remainder allocated to tailored advice + provider feedback group. Parents in tailored advice + provider feedback group older, more likely to be Caucasian and had lower educational level than those in the generic advice group.
Babul, Canada	2007	RCT	yes	no	no	n/a	
Kendrick, UK	2005	RCT (C)	yes	no	yes	n/a	Allocation at level of general practices
McDonald, USA	2005	RCT	yes	unclear	no	n/a	
Watson, UK	2005	RCT	yes	yes for injury outcomes, no for safety practices	yes for injury outcomes, no for safety practices	n/a	
Posner, USA	2004	RCT	yes	yes	no	n/a	
Tan, Singapore	2004	Non RCT (C)	n/a	unclear	yes	yes	Allocation at level of week of clinic attendance. Sequential allocation to treatment group.
Sznajder, France	2003	RCT	yes	no	yes	n/a	
Gielen, USA	2002	RCT (C)	unclear	unclear	no	n/a	Allocation at level of paediatricians.
Hendrickson, USA	2002	RCT	no	no	yes	n/a	
Nansel, USA	2002	RCT	yes	unclear	yes	n/a	
King, USA	2001	RCT	yes	yes	yes	n/a	
Kendrick, UK	1999	Non RCT (C)	n/a	no	yes for injury outcomes no for safety practices	yes	Allocated at level of GP practice. Randomised practices to intervention group and matched control group practices on deprivation score
Clamp, UK	1998	RCT	yes	no	yes	n/a	
Petridou, Greece	1997	CBA (C)	n/a	no	yes	yes	Allocation at level of islands.

(C) = clustered allocation

Table 3-3 Demographic and socio-economic characteristics of studies included in the meta-analyses

1st Author	Year	Age in years (mean (SD))	% male	% residing in rented accommodation	% single parent families	% from black or ethnic minority group	% with at least one parent unemployed
Phelan, USA	2010	0 §	46	-	18	30	17
Nansel, USA	2008	1.2 (1.3)	52	71	32	66	-
Babul, Canada	2007	1.0 (0)	52	39	11	-	-
Kendrick, UK	2005	-	-	20	5	4	-
McDonald, USA	2005	0.81 (0.60)	48	83	54	93	-
Watson, UK	2005	2.59 (1.45)	51	46	28	15	70
Posner, USA	2004	2.26 (1.31)	57	55	-	84	34
Tan, Singapore	2004	-	-	79	-	-	-
Sznajder, France	2003	1.36 (2.06) *	-	-	13	-	34
Gielen, USA	2002	-	-	-	87	94	77
Hendrickson, USA	2002	-	62	-	27	88	74
Nansel, USA	2002	0.95 (0.31)	48	73	19	95	-
King, USA	2001	-	59	-	-	-	-
Kendrick, UK	1999	0.67 (0.22)	52	33	12	7	11
Clamp, UK	1998	2.59 (1.66)	-	21	10	1	12
Petridou, Greece	1997	-	-	-	4	-	-

- = not reported * = refers to youngest child in family § all households recruited when baby was born

3.6   Summary

This chapter has introduced the motivating case-study for the thesis. Unintentional injuries are the leading cause of death in children in industrialised countries. Accidents and injuries in the home in the UK are a leading cause for deaths and admissions to hospital. Falls, particularly falls down stairs, off furniture and on steps, are the main cause of injury-related hospital admissions for children under five. There are a range of home safety prevention recommendations and safety equipment to prevent falls but a lack of evaluation of their effectiveness. The Keeping Children Safe at Home Programme aimed to increase evidence-based NHS injury prevention by assessing the effectiveness and cost-effectiveness of interventions to prevent a range of home accidents including falls, poisonings and scalds.

As part of the Programme a systematic review identified 16 studies with interventions aimed at increasing the possession and use of safety equipment and promoting injury prevention behaviours to prevent falls injuries. The data from these studies are analysed in effectiveness and cost-effectiveness analyses described in Chapters 4 and 6 and in the results presented in Chapters 5 and 7.

4   Evidence Synthesis

This chapter describes the general concepts and methods used for evidence synthesis. It starts with an introduction to meta-analysis of studies that make a pairwise comparison of an intervention group vs a control/usual care group followed by sub-group analyses and meta-regression. Then network meta-analysis (NMA)/mixed treatment comparisons (MTC) methods, that allow the comparison of more than two interventions in a network of evidence, are described. Methods are initially presented for aggregate data only and then expanded to incorporate individual participant data (IPD). Pairwise meta-analysis methods are described firstly using a frequentist approach, this was the approach used for the pairwise analysis in the Cochrane Review update (Kendrick, Young et al. 2012), and then using a Bayesian MCMC approach which in turn is used for the more advanced evidence synthesis methods. The Bayesian MCMC approach is introduced in section 4.1. The methods described in this chapter are applied to the falls accident prevention evidence, detailed in chapter 3, and the results are presented in chapter 5.

4.1   Bayesian Markov chain Monte Carlo (MCMC) approach to evidence synthesis

There are two distinct approaches to statistical inference, the frequentist and the Bayesian approaches. They have much in common but differ in terms of how they interpret probability and uncertainty regarding the model parameters. In a frequentist analysis the parameters of interest, the overall effect size and between study variance in a meta-analysis, are treated as fixed unknown quantities that are estimated from the data through the likelihood with uncertainty expressed in terms of hypothetical repeated sampling from a population. In a Bayesian approach the parameters are considered to be random quantities. Prior probability distributions can be specified for the parameters, representing external information, which are then combined with the

likelihood (equation (4.1)) to obtain a joint posterior probability distribution (Sutton, Abrams 2001).

Bayes theorem is used to combine the *prior beliefs* on the parameter of interest θ, $p(\theta)$, with the information contained in the observed data y, the *likelihood $p(y|\theta)$*, to obtain a *posterior* summary of all the available information upon which inference is based, $p(\theta|y)$ (Ntzoufras, 2009; Lunn *et al.*, 2012; Welton *et al.*, 2012). This is illustrated in equation (4.1).

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$
(4.1)

This can be written as equation (4.2) because the denominator, $p(y)$, does not depend on the parameter θ.

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$
(4.2)

From equation (4.2) the posterior distribution is proportional to the likelihood multiplied by the prior distribution and its measure of central location will lie between the two distributions. Prior distributions can be flat representing weak prior evidence and hence the information provided by the data dominates (with similar results to the frequentist analysis), or the prior can dominate the likelihood if the prior evidence is strong. If the prior distribution and likelihood are conjugates then integration can be used to find the posterior distribution which will be in the same family as the prior distribution (Lunn 2013). If they are not conjugate then one approach is to use Markov chain Monte Carlo (MCMC) simulation methods, in a software package such as WinBUGS, to estimate the posterior distribution of the model parameters, e.g. mean, variance. MCMC draws samples by running Markov chains for a long time until (hopefully) the parameter estimates converge to a stationary distribution (Gilks, Richardson et al. 1996). The Gibbs sampler is one of the most widely used algorithms for simulating Markov chains, it is a special case of the general Metropolis and Hastings algorithm (Lunn 2013) and is implemented in WinBUGS (Spiegelhalter, Thomas et al. 2003). Often there are multiple parameters to be estimated and the methods can be

extended to estimate the parameters using the posterior summaries from the marginal distributions.

The advantages of using a Bayesian approach include:

- Efficient use of all available evidence relevant to the parameter(s) of interest
- Interpretation of parameters as random variables can be useful to predict for future research.
- Results are reported with 95% credible intervals (CrI) which are easier to interpret than confidence intervals as they have a direct probability interpretation – you can state there is a probability of 0.95 that the true value of the parameter lies in the credible interval where the width of the CrI is based on posterior standard deviation.
- If there are no other available evidence about the parameters external to the data, then flat or 'vague' prior distributions can be specified over plausible ranges supported by the parameters of the model. In that case, any flat or 'vague' prior distribution containing a minimal amount of information will be dominated by the data through the likelihood and a Bayesian analysis should produce results close to those obtained from a frequentist analysis.

Priors, distributions and convergence

The choice of the prior distribution can be based on external evidence, such as expert opinion (subjective) or from previous analyses and evidence (objective), or it can be a very wide flat distribution representing a lack of prior knowledge (vague/non-informative). The evidence synthesis models (Chapter 5) and decision modelling analyses (Chapter 7) presented in this thesis are evaluated using a Bayesian MCMC simulation approach with vague prior distributions (Dias, Welton et al. 2011, Lambert, Sutton et al. 2005, Spiegelhalter, Abrams et al. 2004). A sensitivity analysis on the choice of prior distribution, particularly for the scale parameters (standard deviation, variance, precision), is recommended as vague priors, especially when the number of

studies is small, can still influence the estimates and is discussed further in section 4.3.6.

Initial values have to be specified for all parameters to be estimated using MCMC simulation. It is recommended that multiple chains are run with different but sensible starting values to ensure that the chains converge and are not affected by the initial values. WinBUGS can generate its own starting values but they can be extreme and lead to numerical errors. The initial stage of a chain, before convergence to the stationary posterior distribution, is called the burn-in and these simulations should be discarded. After convergence the chains are updated for a large number of simulations to obtain summary statistics from the posterior distribution. To assess convergence: (i) the posterior distribution(s) should be examined visually for spikes and unwanted peculiarities (Dias, Welton et al. 2011), (ii) history plots should be examined to ensure that there is only random scatter around a stable mean values, and (iii) the autocorrelation statistic, which measures the correlation between sampled values, should reduce to zero as the lag time between values increases. The deviance or residual deviance statistics can be used as a measures of goodness of fit and a DIC statistic can be used to compare models (described further in section 4.4). Brooks-Gelman-Rubin diagnostic plots (Gelman, Rubin 1992, Brooks, Gelman 1998) can be used to formally assess convergence when running multiple chains with different starting values (an example can be seen in Figure 5-6). There are two lines representing the within-chain variability and the between-chain variability, that should converged to stability. A third line, the ratio of the within- and between-chain variability, should converge around one. Chains can be compared by overlaying the history plots. As a guide the Monte Carlo error, which reflects both the number of simulations and the degree of autocorrelation, should be no more than 5% of the posterior standard deviation of the parameters of interest (Lunn 2013).

The Bayesian MCMC simulation approach will be used for the evidence synthesis described in this chapter and implemented in chapter 5 and for the decision modelling described in chapter 6 and implemented in chapter 7.

4.2 Meta-analysis

A meta-analysis is the use of statistical techniques to combine the results from two or more studies, identified in a systematic review, to give an overall estimate of an effect size (Moher, Liberati et al. 2009). The Cochrane Handbook (Higgins, Green 2011) lists the following advantages of performing a meta-analysis over a narrative review of the studies identified in a systematic review:

- an increase in power (by combining a number of smaller studies) to detect a real effect as statistically significant if it exists;
- an increase in precision due to the increase in the number of subjects;
- broader questions can be answered than addressed by the individual studies by combining studies with different subject characteristics;
- a formal assessment of conflicting studies.

The validity of the meta-analysis will depend on the quality of the studies identified in a systematic review, that is, the search strategy needs to match the research question and yield a reasonably complete and unbiased collection of the relevant studies (and providing the included studies are valid) (Borenstein, Hedges et al. 2009). Meta-analysis combines evidence "usually from RCTs" but non-RCTs and controlled before-and-after studies can be included provided they have been assessed for limited selection bias (NICE 2012, Higgins, Green 2011). Any studies identified but not included in the meta-analysis should be included in a narrative review. Many public health intervention evaluations use the lack of RCT evidence as the main reason for not conducting a meta-analysis (Achana, Hubbard et al. 2014). Outcomes for the studies can be dichotomous, continuous, ordinal or some other outcome measure.

4.3 Pairwise meta-analysis

A pairwise meta-analysis is the most commonly used evidence synthesis method, the methods are well developed and easy to apply using statistical software. It compares the effectiveness of two interventions, usually an enhanced intervention compared to a control intervention, e.g. usual care, that have been compared in two or more

studies. Often fairly heterogeneous interventions are combined to form these two groups.

In this thesis the outcome of interest for a study i is dichotomous, e.g. uptake of intervention yes/no), so the methods focus on an outcome with only two possibilities ("event" and "no event") with the odds ratio (OR) as a measure of effect size, Y_i, as shown in the 2x2 table in Table 4-1.

Table 4-1 2x2 table representing the outcome from a single study with two possible outcomes

Study i	Event	No event	Total
Intervention	r_{i1}	$n_{i1} - r_{i1}$	n_{i1}
Control/usual care intervention	r_{i2}	$n_{i2} - r_{i2}$	n_{i2}

Where n_{ij} is the number of subjects in study i on intervention j
And r_{ij} is the number of subjects with "event" in study i on intervention j

For study i ($i = 1, \dots ns$ where ns is the number of studies) the OR is given by equation (4.3).

$$\text{OR}_i = \frac{r_{i1}/(n_{i1}-r_{i1})}{r_{i2}/(n_{i2}-r_{i2})} \tag{4.3}$$

where OR compares the odds of an "event" in the intervention arm to the odds of an "event" in the control arm in study i.

The natural log of the OR is used for inference as under the large samples assumptions its sampling distribution can be assumed approximately normally distributed with standard error for the $log\ OR$ (using Woolf's method) given in equation 4.4. The $log\ OR$ will be referred to as the effect size in the thesis.

$$\text{se}(\log \text{OR}_i) = \sqrt{\frac{1}{r_{i1}} + \frac{1}{(n_{i1} - r_{i1})} + \frac{1}{r_{i2}} + \frac{1}{(n_{i2} - r_{i2})}} \tag{4.4}$$

Where empty cells exist, e.g. no events, 0.5 is often added to all cells in the table due to the problems in computing the effect size and standard error. This can bias the study estimate towards no difference and overestimate the variance of the study estimate thus down-weighting their contribution in the meta-analysis (Higgins, Green 2011, Sweeting, Sutton et al. 2007) hence should be used with caution.

An effect size Y_i, log (OR), and variance is computed for each study and then an overall weighted mean of these effect sizes is calculated. More weight is assigned to the more precise/informative studies. The mechanism to assign weights depends on the assumptions about the distribution of effect sizes from which the studies were sampled (Borenstein, Hedges et al. 2009) and is described in sections 4.3.1 and 4.3.3.

Pairwise meta-analysis models are defined as either fixed-effects (section 4.3.1) or random-effects (section 4.3.4). The results are usually displayed in a forest plot in which the individual study results together with the overall effect estimate and confidence interval are displayed. Forest plots are presented for the pairwise meta-analysis in section 5.2

4.3.1   Fixed effects model

Under a fixed-effect model no heterogeneity between studies is assumed; all studies are assumed to be estimating the same underlying true effect size, d, and the estimates only differ because of random variation. A fixed effects meta-analysis model, combining ns studies, is given equation 4.5.

$$Y_i = d + V_i \qquad i = 1, \dots number\ studies\ ns \qquad (4.5)$$

where Y_i is the observed within study effect estimate ($log\ OR_i$) for study i, d is the overall true effect size ($log\ OR$) and V_i is the within-study variance for study i from the true effect d.

Study effects are weighted to create the overall pooled effect estimate. There are different methods to estimate the weights: inverse variance; Mantel-Haenszel; Peto (Deeks, Altman et al. 2008). The inverse variance weighted method is described below.

Inverse variance-weighted method

The inverse variance-weighted method is the simplest and most commonly used method. It uses weights, w_i, which are the inverse of the variance of the study effect size, Y_i, $w_i = \frac{1}{V_i}$.

Larger studies which have smaller standard errors have a higher weighting than smaller studies and this minimises the imprecision of the pooled effect estimate.

4.3.2 Fixed effects model using MCMC simulation

For a fixed effects meta-analysis using MCMC simulation, a prior distribution must be specified for d, the true effect size and parameter of interest.

For a meta-analysis on the $log\ OR$ scale the prior distribution for d is often specified as $d \sim Normal(0, 10^5)$ (Welton 2012). This is a vague prior distribution, very wide and flat which allows the data to dominate (Lunn 2013).

4.3.3 Heterogeneity

Heterogeneity will always exist between studies due to clinical and methodological diversity and if it is substantial then it needs to be measured and accounted for (Higgins, Green 2011). In a fixed effects model it is assumed that all the studies are estimating the same overall true effect. The I^2 heterogeneity statistic can be calculated to investigate inconsistency of the findings between studies (equation (4.6)).

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

where (4.6)

$$Q = \sum w_i(Y_i - d)^2$$

with df = degrees of freedom for $Q = (ns - 1)$, where ns is the number of studies

Q is a chi-square statistic used to test if there is evidence of heterogeneity beyond chance; it is the weighted sum of squared differences between the study means and the fixed effect estimates.

I^2 is interpreted as the percentage of variability attributable to the heterogeneity between studies rather than sampling error. There are thresholds to assist in interpreting I^2, with a value of 75% or over indicating considerable heterogeneity and less than 40% indicating that heterogeneity may be unimportant (Higgins, Green 2011). I^2 should not be used to solely decide if studies should be pooled in a fixed effects meta-analysis but the clinical relevance of the heterogeneity should also be assessed . I^2 does not depend on the number of studies in a meta-analysis but is affected by the amount of evidence and has been shown to increase artificially as the number of participants in the studies increases, particularly when a large study follows a small study. The between study variance, τ^2, can be used directly to quantify heterogeneity as it is measured on the same scale as the outcome and it is not inflated by increasing numbers of participants (Rucker, Schwarzer et al. 2008b, Higgins, Thompson 2002). τ^2 is discussed further in section 4.3.6.

4.3.4   Random effects model

A fixed effects meta-analysis ignores heterogeneity among results of studies and gives an estimate of the overall effect estimate, d the $log\ OR$, that represents a typical intervention effect assuming that any observed differences between studies are due to chance. A random effects meta-analysis can be used to incorporate heterogeneity among studies by assuming that there may be a distribution of intervention effects that cannot be explained by study characteristics (these will be discussed later in section 0) (Higgins, Green 2011). The overall effect estimate, d, is the average of a distribution of effect sizes, δ_i. The random effects model is given in equation ((4.7)).

$$Y_i = \ \delta_i + \ V_i \qquad\qquad i = 1, \dots \text{number studies } ns \qquad\qquad (4.7)$$

with $\delta_i \sim \text{N}(d, \tau^2)$

where Y_i is the observed within study effect estimate in study i, δ_i is the study specific effect *(log OR_i)*, V_i is the within study variance for study i from the overall effect size d and τ^2 is the between study variance.

Study specific effects, δ_i, are assumed not to be equal but exchangeable. This means that they are "similar"; the study "labels", i, convey no information and they are independently and identically distributed. The common distribution is usually chosen to be a normal distribution with mean d and variance τ^2 (Lunn 2013, Dias, Sutton et al. 2013a). This is equivalent to a fixed effects model if $\tau^2 = 0$. Random effects models tend to be more conservative and give wider confidence intervals.

The total variance for a random effects model (equation (4.7)) is $V_i + \tau^2$, where V_i is the within study variance and τ^2 is the between study variance.

The most commonly used frequentist method to fit the random effects model uses the inverse variance-weighted method (DerSimonian, Laird 1986). The weights used are $w_i = \frac{1}{V_i + \tau^2}$.

This method gives an efficient estimate of the overall intervention effect but can be inefficient in estimating the between-study variance if the studies are not all of similar size and if the number of studies is small (Dersimonian, Laird 2015, Jackson, Bowden et al. 2010a) and is discussed further in the next section.

Random effects models are used for the pairwise meta-analysis in section 5.2 due to the observed heterogeneity between the identified studies.

4.3.5 Alternative specification of the fixed and random effects models using a logit model

Rather than model Y_i, the $log\ OR_i$, as described in equation (4.7), a meta-analysis formulation that models the number of events out of the total number of subjects in the two arms using the logit function (Smith, Spiegelhalter et al. 1995, Higgins, Whitehead 1996, Simmonds, Higgins 2016) is possible. This method can naturally be extended to IPD analysis and network meta-analysis (sections 4.9 and 4.11).

Let r_{ik} be the numbers of subjects with the outcome of interest, where $k = 1$ is the control arm and $k = 2$ is the intervention arm in study i out of the total numbers of subjects n_{ik}, then $r_{ik} \sim Bin(n_{ik}, p_{ik})$ and p_{ik} is the probability of an event in arm k of study i (Table 4-1).

A logit link, which maps the probabilities (on a $0-1$ scale) into a continuous measure on the scale $-\infty \ to \ +\infty$, is used where $logit(p) = log(p/(1-p))$ and the random effects model can be specified as in equation (4.8).

$$r_{ik} \sim Bin(n_{ik}, p_{ik})$$

$$\begin{aligned} logit \ p_{i1} &= \mu_i \\ logit \ p_{i2} &= \mu_i \ + \ \delta_i \end{aligned} \qquad (4.8)$$

$$\delta_i \sim N(d, \tau^2)$$

where δ_i is the $log \ OR_i$ study i, d is the true effect size $(log \ OR)$, τ^2 is the between study variance, μ_i is the log odds of an event in the control group in study i

The simplest and most commonly used method to estimate τ^2, the between study variance, is the DerSimonian-Laird approach (DerSimonian, Laird 1986, Dersimonian, Kacker 2007). This method performs well for large sample sizes and the but alternative methods such as using maximum likelihood (ML), restricted maximum likelihood (REML) or profile likelihood are preferable for smaller sample sizes particularly if inferences about the between study variance are important (Jackson, Bowden et al. 2010b). Alternatively a Bayesian approach using MCMC could be used and is described in section 4.3.6.

In a fixed effects analysis $\tau^2 = 0$ and hence $\delta_i \ = \ d$ for all i.

4.3.6 Random effects model using MCMC simulation

For the random effects analysis described in section 4.3.4 using MCMC simulation, a prior distribution is required for τ^2, the between study variance, as well as d, the true effect size. For the analysis described in section 4.3.5, a prior distribution is also required for μ_i, the log odds of an event in the control group in study i. Commonly used vague prior distributions (Welton 2012) are

$$\tau \sim Uniform(0,10)$$

$$d \sim Normal(0, 10^5)$$

$$\mu_i \sim N(0, 10^5)$$

τ^2 estimates how much variability there is between estimates from the population of studies. τ can be interpreted in terms of the "range" of the ORs; 95% of d's lie in the range $d \pm 1.96\tau$ so $exp(3.92\tau)$ is the "range" of the odds ratios. A value of $exp(3.92\tau) > 10$ ($\tau > 0.59$) is considered to be a high value of between study standard deviation τ (Spiegelhalter, Abrams et al. 2004). If there are only a small number of studies (<10) a sensitivity analysis is highly recommended as the choice of prior distribution can be influential. The following prior distributions could be considered (Lambert, Sutton et al. 2005):

$$\tau^2 \sim Uniform(10^{-3}, 4)$$

$$\log(\tau^2) \sim Uniform(-10, 1.386),$$

$$1/\tau^2 \sim Gamma(0.001, 0.001)$$

A predictive distribution can be calculated and estimates the underlying effect in a new study, $\delta^{new} \sim N(d, \tau^2)$. It may be a more appropriate summary of the intervention effect than the overall mean effect as the average of the individual study effects may not accurately represent the different study populations as it does not account for between study heterogeneity. This is especially the case if there is high degree of

heterogeneity (Spiegelhalter, Abrams et al. 2004, Higgins, Thompson et al. 2009). The calculation of δ^{new} is described further in section 4.4.

Posterior distributions for the study level intervention effects, $\delta_i = log\ OR_i$ for study i can be estimated. Under the assumption of exchangeable δ_is, each posterior distribution borrows strength (precision) from the others via their joint influence on the estimation of the underlying population parameter (Lunn 2013). As a result, the uncertainty around the intervention effect estimates is spread more evenly across the studies and there is "shrinkage to the mean". The individual study estimates contribute to the population intervention effect estimate proportional to the study size. The more extreme study effect estimates, which typically come from small studies because of sampling variation, get pulled (shrunk) towards the population mean because the larger studies, which tend not to have the extreme estimates, contribute more to locating the population mean effect. Shrinkage is lowest for the largest studies (Lunn 2013).

Pairwise meta-analyses are repeated using a Bayesian MCMC simulation approach in section 5.2 and compared to the frequentist analysis results from Stata.

4.4   Assessing model fit and inconsistency for pairwise meta-analysis

4.4.1   Model fit for pairwise meta-analysis

Models should be assessed on how well the predictions from a model fit the observed data. Using WinBUGS to fit the model, an estimate of the posterior distribution of the deviance statistic can be obtained. The deviance statistic, D, measures the fit of the predicted model to the observed data using the likelihood function and can be calculated as a function of the model parameters for each MCMC simulation. The posterior mean deviance, $\overline{D}$, is used to measure how much the model predictions vary from the observed data. The smaller the value of $\overline{D}$ the better the fit but this can be difficult to interpret. A more useful measure of fit statistic is the overall residual deviance, $\overline{D}_{res}$. $\overline{D}_{res}$ is the difference between the posterior mean deviance for the model fitted and the deviance of the saturated model in which the predictions equal

the observed data (McCullagh, Nelder 1989). If the model is an accurate fit then $\bar{D}_{res}$ should be approximately equal to the number of unconstrained data points (i.e. in a meta-analysis this will equate to the number of studies × number of arms). For a binomial response, equation (4.9) gives the calculation of $\bar{D}_{res}$.

$$
\begin{aligned}
\bar{D}_{res} &= \sum_{ik} 2\left(r_{ik}\log\left(\frac{r_{ik}}{\hat{r}_{ik}}\right) + (n_{ik} - r_{ik})log\left(\frac{(n_{ik} - r_{ik})}{(n_{ik} - \hat{r}_{ik})}\right)\right) \\
&= \sum_{ik} D_{ik}
\end{aligned}
\tag{4.9}
$$

where $\hat{r}_{ik} = n_{ik}p_{ik}$ is the model predicted number of events, r_{ik} is the observed number of events and D_{ik} is the deviance statistic for study i, intervention arm $k = 1,2$ and is calculated for each MCMC simulation (Dias, Welton et al. 2011, Spiegelhalter, Best et al. 2002).

This statistic can be extended to include more than two interventions and hence is used in network meta-analysis (section 4.11.4).

The deviance information criteria, DIC, can be used to compare models (Spiegelhalter, Best et al. 2002). The DIC penalizes the posterior mean deviance, $\bar{D}$, of the model by the effective number of parameters, p_D, i.e the complexity of the model. For a fixed effects model, p_D is the number of study baselines, μ_i, plus the one fixed intervention effect, d. For a random effects model p_D will depend on the heterogeneity between studies and the intervention effect contribution can vary between one and the number of studies. p_D is quantified using equation (4.10).

$$
p_D = \sum_{ik} \left(\bar{D}_{ik} - \widetilde{D}_{ik} \right)
\tag{4.10}
$$

where WinBUGS uses the posterior mean of the r_{ik} instead of $\hat{r}_{ik}$ to calculate the deviance $\widetilde{D}_{ik}$. Equation (4.11) defines the DIC.

$$
DIC = \bar{D} + p_D
\tag{4.11}
$$

Comparing two models fitted to the same data, a model with a smaller value of the DIC is considered to be a better fit and differences in DIC over five are considered important (MRC Biostatistics Unit 2015b)

4.4.2   Assessing inconsistency for pairwise meta-analysis

Possible inconsistencies between different study results, such as one study demonstrates a strong intervention effect and the majority of other studies demonstrate no effect, should be investigated. Cross-validation can be used to explore the effect of omitting each study as a sensitivity analysis. Using a Bayesian MCMC approach, a predictive distribution, based on the remaining studies, can be found for what would be expected to be observed in the omitted study as if it were a new study (Dias, Sutton et al. 2012). For a random effects meta-analysis the true intervention effect for the omitted study would be drawn from the random effects distribution of effects, $\delta^{new} \sim N(d, \tau^2)$. This depends on uncertainty in the mean value d and uncertainty in where the new study lies in the random effects distribution. The new study will have the same sample size as the omitted study so the predicted probability of an "event" in the intervention arm, p_2^{new}, is given in equation (4.12).

$$logit(p_2^{new}) = logit(p_1^{new}) + \delta^{new} \qquad (4.12)$$

The baseline probability of an "event" in the control arm, p_1^{new}, is drawn from $Beta(r_1, n_1 - r_1)$, which describes the uncertainty in the proportion for a given number of "events" in the control arm.

The predicted number of positive responders in the intervention arm can then be compared to the observed number in the omitted study. A Bayesian p-value can be calculated by monitoring when r_2^{new} exceeds r_2, where r_2 is the number of "events" in the intervention group.

4.5   Reporting/publication bias

Reporting/publication biases arise when the dissemination of research findings is influenced by the nature and the direction of results. Statistically significant results are more likely to be published, published rapidly, published in English, publishing more than once in high impact journals and cited in research papers (Sterne, Sutton et al. 2011). Data that lead to negative results may be filtered, manipulated, or presented in such a way that they become positive (Sterne, Sutton et al. 2011). Studies with statistically non-significant results are as important as studies with statistically significant results (Higgins, Green 2011).

A funnel plot is a simple scatter plot of the intervention effect estimates from individual studies on a horizontal axis against some measure of each study's size or precision on the vertical axis. Figure 4-1 gives an example of a funnel plot showing no publication bias. The precision of the effect size estimate increases as the study size increases; there is more scatter at the bottom of the graph with the spread narrowing towards the top for larger studies, hence the name funnel. Often smaller studies with non-statistically significant results are not published, hence the funnel will be asymmetrical with a gap in the bottom corner of the graph and estimates of the intervention effect are likely to overestimate the effect (Egger, Davey Smith et al. 1997).
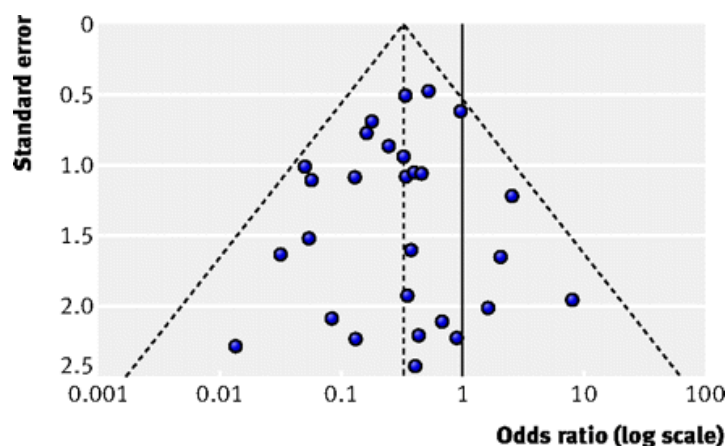
The outer dashed lines indicate the triangular region within which 95% of studies are expected to lie in the absence of both biases and heterogeneity (fixed effect summary log odds ratio±1.96×standard error of summary log odds ratio). The solid vertical line corresponds to no intervention effect. (Sterne, Sutton et al. 2011)

Figure 4-1 Example of symmetrical funnel plot

Tests for funnel plot asymmetry have low power and are only recommended if there are at least 10 studies in the meta-analysis and should not be used if the standard errors of the intervention effect estimates are all similar (Sterne, Sutton et al. 2011). To test for funnel plot asymmetry the arcsine test proposed by Rucker (Rucker, Schwarzer et al. 2008a) can be conducted together with inspection of contour enhanced funnel plots (Peters, Sutton et al. 2008). To aid visual interpretation, contour enhanced funnel plots include contour lines corresponding to perceived milestones of statistical significance, an example is given in Figure 4-2. There are several alternative tests available to assess funnel plot asymmetry, these include the tests by Harbord et al (Harbord, Egger et al. 2006) and Peters et al (Peters, Sutton et al. 2008) but the arcsine test has been shown to be preferable if there is substantial between study heterogeneity ($\tau^2 > 0.1$) however it is slightly conservative in the absence of heterogeneity (Sterne, Sutton et al. 2011). Funnel plot asymmetry is not linked solely to reporting/publication bias and can be due to a number of different biases, such as poor methodological quality, size of effect differs according to study size, sampling variation and chance, so caution is required in interpretation (Sterne, Sutton et al. 2011).
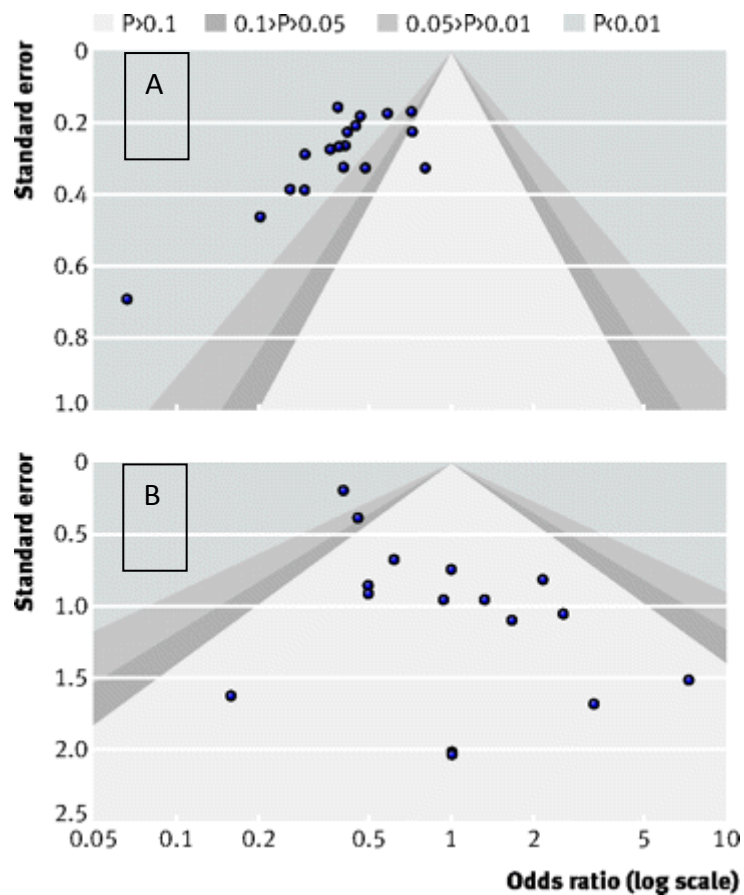
The contour lines represent statistical significance milestones of study estimates. Plot A appears to have missing studies in the middle and right of the plot, broadly in the white area of non-significance, making publication bias plausible. Plot B appears to have missing studies on the left hand side of the plot. Since most of this area contains regions of high significance, publication bias is unlikely to be the underlying cause of asymmetry. (Sterne, Sutton et al. 2011)

Figure 4-2 Contour enhanced funnel plots

4.6   Sub-group analyses and meta-regression

In section 4.3 fixed- and random-effects meta-analyses are described and methods to measure the level of heterogeneity discussed (section 4.3.3 and 4.3.6). A random effects meta-analysis can account for heterogeneity but the overall intervention effect describes an average across all included populations of participants and study types which may not be meaningful (Riley, Steyerberg 2010). If heterogeneity is identified then it can be explored using sub-group analyses or meta-regression on pre-specified characteristics of the studies that might contribute to heterogeneity. Heterogeneity can arise from clinical heterogeneity (e.g. variability in the participants, interventions and outcomes), and methodological heterogeneity (e.g. variability in study design and risk of bias) (Higgins, Green 2011). In a sub-group analysis it is assumed that all studies within the sub-groups share a common effect size or a common distribution of effect sizes for a random effects model. In meta-regression it is assumed that all studies with the same covariate value share a common effect size or a common distribution of effect sizes for a random effect model (Borenstein, Hedges et al. 2009)

4.6.1   Sub-group analyses

Sub-group analyses can be conducted to investigate heterogeneous results, or to answer specific questions about particular participant groups (e.g. males or females), types of intervention (e.g. different intensities of an intervention) or subsets of studies (e.g. geographical location) (Higgins, Green 2011). The disadvantages are that there are different estimates of the between study heterogeneity for each sub-group and it is difficult to assess if the intervention effects are the same across the sub-groups (Dias, Sutton et al. 2013b) and there are often small numbers of studies in some sub-groups. Sub-group analyses are used to explore the study binary, yes/no, characteristics in section 5.3.1.

4.6.2   Meta-regression

Meta-regression can be used to explore the effects of participant and study characteristics (covariates), such as the percentage of females or mean age of the participant, in a single analysis with a shared between study heterogeneity and an

interaction of the characteristics with the intervention. Study-level aggregated participant characteristics are usually the only information available and it is assumed that through the study design, RCTs, these characteristics are evenly distributed across the intervention groups. Participant characteristics are best explored with individual participant data (IPD) (discussed in section 4.9), if available. This section will look at aggregated study level covariates only.

Meta-regressions can use fixed or random effects models but random are preferred because they can account for heterogeneity not explained by the covariate. The random effects models in section 4.3.5 can be extended to include the covariate of interest, X_i, for study i given in equation (4.13).

$$logit \; p_{i1} \; = \; \mu_i$$

$$logit \; p_{i2} \; = \; \mu_i \; + \; \delta_i + \beta x_i$$

(4.13)

where

x_i is the study-level covariate value for study i and
β is the covariate interaction effect on the intervention
$\delta_i \sim N(d, \tau^2)$ is the study specific log odds ratio

In a frequentist setting the results of a meta-regression can be presented in a bubble plot. It is a scatter plot with the intervention effect for each study on the y-axis and the covariate used in the meta-regression on the x-axis. The size of the bubble is inversely proportional to the variance of the estimated intervention effect (Thompson, Higgins 2002) and the larger studies tend to have the bigger bubbles. Using MCMC a non-informative prior distribution is used for β, for example $\beta \sim N(0, 100^2)$ (Dias, Sutton et al. 2012) and covariates are usually centred to improve mixing and reduce autocorrelation (Welton 2012). Meta-regression is not recommended when there are only a small number of studies that report the covariate of interest (Higgins, Thompson 2004).

Meta-regression is used to explore subject characteristic covariate effects in section 5.3.2

4.7 Adjusting for baseline risk

One possible explanation for between study heterogeneity is an association between baseline risk and intervention. Baseline risk is the log odds of an event for a subject in the control arm and is an estimate of the risk for a subject if they do not receive an enhanced intervention (Higgins, Green 2011). The studies combined in a meta-analysis are often heterogeneous as described above and this may modify the effect of the intervention (Arends, Hoes et al. 2000, Achana, Cooper et al. 2012, Thompson, Smith et al. 1997). In a meta-analysis several methods have been proposed for including the baseline risk, e.g. using the observed risk of events in the usual care/control group, the observed usual care/control log odds and the average of the observed event risks in the usual care/control and intervention groups. Using the observed risk in the baseline group can be problematic because it is also part of the calculation of the odds ratio outcome leading to structural dependence and both the covariate and outcome are estimated rather than true values; this can lead to overestimation (Thompson, Smith et al. 1997). Using WinBUGS the relationship between the true control group log odds and the true odds ratio can be investigated using the model in equation (4.14).

$$logit \ p_{i1} \ = \ \mu_i$$

$$logit \ p_{i2} \ = \ \mu_i \ + \ \delta_i + \beta(\mu_i - \bar{\mu})$$

(4.14)

where

μ_i is the log-odds of an event in the control group and a covariate centred on the mean control group log odds across all studies, $\bar{\mu}$

$\delta_i \sim N(d, \tau^2)$ is the study specific log odds ratio based on the underlying log odds and not the observed baseline log odds.

In section 5.4 the usual care/control arm log odds are used as the estimate of baseline risk.

4.8 Adjusting for cluster trials

Studies identified in a systematic review may include cluster randomised trials. In a cluster trial the participants within any one cluster often tend to respond similarly so independence between the participants can no longer be assumed. Cluster trials should not be analysed assuming that the unit of allocation to intervention is the participant as this can give artificially low p-values (Higgins, Green 2011).

To account for clustering the size of each trial is reduced to its 'effective sample size' (Donner, Klar 2002). The effective sample size of a single intervention group in a cluster-randomized trial is its original sample size divided by a quantity called the 'design effect'. The design effect is $(1 + (M - 1) \, ICC)$ where M is the average cluster size and ICC is the intracluster correlation coefficient.

A common design effect is usually assumed across intervention groups. For dichotomous data both the number of participants and the number with the "event" should be divided by the same design effect. Original study analysis does not always adjust for clustering and hence does not report the ICC so, at the meta-analysis stage, an external estimate for the ICC needs to be researched.

In chapter 5 clustering was accounted for either by adjusting the data prior to analysis or adjusting within the analysis, using reported or estimated ICC.

4.9 Individual participant data

Individual participant data (IPD) relates to the data recorded for each participant in the original study. IPD meta-analyses have been described as the gold standard (Stewart 1995) and they have many advantages over meta-regression using summary/aggregated covariates, these include:

- When participant level covariates are of interest, using the IPD to regress individual participant characteristics on individual participant outcomes will produce a more

powerful and reliable analysis (Lambert, Sutton et al. 2002, Berlin, Santanna et al. 2002).

- Standardising analysis methods if they are reported differently or are missing in the individual studies, for example some studies may report mean and other median covariate values (Riley, Lambert et al. 2010).

The PRISMA-IPD statement, a stand-alone extension to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), has been developed and tailored to the specific requirements of reporting systematic reviews and meta-analyses of IPD (Stewart 2015).

IPD meta-analysis is a resource-demanding approach to evidence synthesis, it encounters problems such as uncooperative investigators unwilling to provide the IPD, incompleteness of records and having to standardise participant characteristics across studies (Rogozińska, Marlin et al. 2017).

There are two approaches to IPD meta-analyses, a two-step approach and a one-step approach. In a two-step approach firstly the effect is estimated in each study with its standard error and then a meta-analysis is conducted of the effect estimates. This approach has been shown to be most commonly used as it is the quickest and least complex of the two approaches (Simmonds, Higgins et al. 2005). In a one-step approach all the IPD data is combined and analysed simultaneously whilst accounting for the within study clustering of participants. This approach avoids some of the assumptions of meta-analysis so is useful when the studies are small and events rare (Simmonds, Higgins 2016) and will be applied in this thesis.

The one-stage method described in sections 4.9.1 - 4.9.3, takes a Bayesian MCMC simulation approach for model estimation and is used in Chapter 5. Sections 4.9.1 and 4.9.2 incorporate IPD only and IPD with aggregate data respectively and section 4.9.3 extends meta-regression to combine aggregate and ID data.

The authors of the systematic review of studies reporting interventions to prevent childhood falls in the home (Young, Wynn et al. 2013) requested IPD from the authors

of the studies identified in the review and this is incorporated into the meta-analysis and meta-regression reported in sections 5.5-5.6.

4.9.1 Meta-analysis using only IPD data

If all studies have IPD then, using a one-stage approach, the logit random effects model for a binary outcome (section 4.3.5) can be extended (Turner, Omar et al. 2000, Sutton, Kendrick et al. 2008, Simmonds, Higgins 2016) and is given in equation (4.15).

$$Y_{si} \sim Bernoulli(p_{si})$$

$$logit \ p_{si} \ = \ \mu_i \ + \ \delta_i int_{si}$$

(4.15)

where
i is 1, 2, … number of IPD studies
s represents each participant in the ith study
int_{si} is intervention for participant s in study i, 0 if control/usual care, 1 if intervention
μ_i is the estimated log odds of an event in the control group in study i
δ_i is the estimated log OR of the intervention effect in study i

Prior distributions (Sutton, Kendrick et al. 2008):

$$\delta_i \sim N(d, \tau^2), \ \mu_i \sim N(0, 10^6), \ \tau \sim unif(0,10)$$

4.9.2 Meta-analysis using IPD and aggregate data

IPD is usually not available on all studies but, to avoid possible availability bias (Riley, Simmonds et al. 2007), studies with IPD can be combined with those that only have aggregate covariates in a meta-analysis. The IPD model (section 4.9.1) can be extended to combine IPD and aggregate data sources and adjust for clustering (Sutton, Kendrick et al. 2008). This model is split into five parts that model the different types of available data (IPD and aggregate, individually allocated or clustered) and combine them together in an overall meta-analysis. The model is a simplification of the meta-regression model which is described in section 4.9.3 removing any covariate terms and hence is not presented. Prior distributions can be chosen to be vague and the δ_j are assumed to be exchangeable across all studies in all parts of the model $\delta_j \sim N(d, \tau^2)$ (Sutton, Kendrick et al. 2008).

4.9.3   Meta-regression using IPD and aggregate data

It has been discussed previously that random effects models can account for heterogeneity and meta-regression can be used to assess between-study aggregated participant level covariate effects assuming that they reflect the within-study relationship between the individual response and the individual covariate values. Some participant characteristics cannot be investigated using meta-regression when they are aggregated across the studies. They may vary substantially within a study but, when aggregated to give for example a mean, do not exhibit any variation between studies and may be prone to study-level confounding. This is known as aggregation bias or ecological bias (Berlin, Santanna et al. 2002).

A random effects IPD model can include an intervention by covariate interaction as in a meta-regression (section 4.6.2). By using IPD and aggregate data the intervention by covariate interactions can be estimated using between-study variability when only summary data are available and using within-study and between-study variability if IPD are available. The IPD components compare the intervention effect among those with and without the covariate of interest and the aggregate data components show the effect of a unit change in the mean study covariate value on the average intervention effect across studies.

An advantage of IPD is that both between- and within-study coefficients can be used to assess possible ecological bias where the study-level and participant-level results are different (Riley, Steyerberg 2010). It has been shown through simulation studies that the between study association is approximately an unbiased estimate of the within-study association (Lambert, Sutton et al. 2002). However, due to ecological bias or study-level confounding, the between-study association can be very different from the within-study association (Greenland, Morgenstern 1989).

Meta-regression for both IPD and aggregate data (Sutton, Kendrick et al. 2008), splitting the variability between and within-studies (Riley, Steyerberg 2010), is presented below with adjustments included for clustering in studies. Abo-Zaid et al in 2013 showed in a simulation study that models accounting for clustering perform consistently well, but downwardly biased effect estimates and low coverage can occur

when ignoring clustering (Abo-Zaid, Guo et al. 2013). There are five parts in the meta-regression analysis that uses a Bayesian MCMC approach in WinBUGS, prior distributions are taken from Sutton et al (Sutton, Kendrick et al. 2008):

Part 1: Individually allocated IPD studies (random effects)

A logistic regression model (equation (4.16)) is used to estimate the effect sizes from IPD studies including an intervention (int_{si}) by covariate (x_{si}) interaction and splitting the between- and within-study variance.

$$Y_{si} \sim Bernoulli(p_{si})$$

(4.16)

$$logit\ p_{si} = \mu_i + \delta_i int_{si} + \beta_{0i} x_{si} + \beta^W int_{si}(x_{si} - \bar{x}_i) + \beta^B int_{si}\bar{x}_i$$

where
  i is 1, 2, … number of individually allocated studies
  s represents each subject/participant in the ith study
  Y_{si} is the response for the sth participant in the ith study
  x_{si} is the covariate value for the sth participant in the ith study
  $\bar{x}_i$ is the mean covariate value for study i
Prior distributions

$$\mu_i \sim N(0, 10^6)$$
$$\beta_{0i} \sim N(0, 10^6)$$

Part 2: Cluster allocated IPD studies

A logistic regression model (equation (4.17)) is used to estimate the effect sizes from cluster allocated IPD studies including an intervention (int_{smi}) by covariate (x_{smi}) interaction and splitting the between- and within-study variance.

$$Y_{smi} \sim Bernoulli(p_{smi})$$

$$logit\ p_{smi} = \mu_{mi} + \delta_i int_{smi} + \beta_{0i} x_{smi} + \beta^W int_{smi}(x_{smi} - \bar{x}_i)$$

(4.17)

$$+ \beta^B int_{smi}\bar{x}_i$$

where
  s represents each subject/participant in the ith study
  i =no. individually allocated IPD studies, …., (no. individually allocated IPD studies +no. cluster allocated IPD studies)
  m represents the cluster for the sth participant in the ith study
  Y_{smi} is the response for the sth participant in the kth cluster in the ith study

x_{smi} is the covariate value for the sth participant in the ith study in the kth cluster

Prior distributions

$\mu_{mi} \sim N(\psi_i, \tau.cluster_i^2)$
$\psi_i \sim N(0, 10^6)$
$\tau.cluster_i \sim \text{Unif}(0,0.1)$
$\beta_{0i} \sim N(0, 10^6)$

Each cluster has its own control group event rate and they are assumed exchangeable within each study. Cluster effects are assumed independent between studies.

In the models in parts 1 and 2:

β_{0i} is the study specific individual level covariate effect

β^W estimates the within-study association (change in an individual's logit event risk for a one-unit increase in x_{ij}) assumed the same for all individuals in all IPD studies

β^B estimates the between-study association assumed the same for all individuals in all IPD studies and equivalent to the β^B in the aggregate data models (part 3)

Prior distributions $\beta^W, \beta^B \sim N(0, 10^6)$

The difference between β^W and β^B represents potential ecological bias and can be estimated with uncertainty in WinBUGS.

Part 3: Aggregate data (AD) studies (not clustered)

A random effects meta-regression model (section 4.2), equation (4.18) is used for studies, i, providing aggregate data only that were not cluster allocated.

$$r_{ik} \sim Bin(n_{ik}, p_{ik})$$

$$logit\ p_{i1} = \mu_i \qquad\qquad (4.18)$$

$$logit\ p_{i2} = \mu_i + \delta_i + \beta^B x_i^{agg}$$

where

k=1 for usual care/control and =2 for intervention

x_i^{agg} is the study level covariate for study i

Part 4: Aggregate data (AD) cluster allocated studies, i, are combined using equation (4.19) (assuming no adjustment for clustering prior to analysis, if they have been adjusted then parts 3 and 4 can be combined).

$$design\ effect_i = 1 + (average\ cluster\ size_i - 1) \times ICC_i$$

$$\sigma^2\ adjusted_i = \sigma_i^2 \times design\ effect_i$$

$$T_i \sim N(\delta_i', \sigma^2\ adjusted_i)$$ (4.19)

$$\delta_i' = \delta_i + \beta^B x_i^{agg}$$

where
x_i^{agg} is the study level covariate for study i
T_i is the intervention effect $(\log OR)$ for study i

This extends the random effects meta-regression model adjusting for clustering in the cluster-allocated studies assuming that the effect of clustering had not been adjusted for in the original analysis. To adjust for clustering the design effect is used to inflate the variance. Part 4 could be combined with part 3 if the data is adjusted for clustering prior to analysis.

Part 5: Model combining all estimates of intervention effect from the 4 data sources (equations (4.16)-(4.19)) is given in equation (4.20).

$$\delta_i \sim N(d, \tau^2)$$

$$d \sim N(0, 10^6)$$ (4.20)

$$\tau \sim Unif(0, 0.1)$$

where

$i = 1, \ldots .\ total\ number\ of\ studies$ (IPD, AD, cluster and non-cluster)

A random effect is placed on all intervention effect estimates, δ_i $(\log(OR_i))$, from parts 1-4 assuming exchangeability across studies and a normal distribution.

Sensitivity analyses should be conducted on the choice of prior distribution, particularly for τ^2, the initial values and the number of iterations used in WinBUGS. Continuous covariates are often centred to improve convergence and reduce autocorrelation. This analysis can be extended to include multiple covariates but will be limited by the number of studies included in the meta-analysis.

4.10 Summary of pairwise meta-analysis

Sections 4.3-4.9 describe a pairwise meta-analysis to compare the effectiveness of two interventions evidenced in two or more studies with a binary/dichotomous outcome. If the studies are heterogeneous then a random effect model (section 4.3.4) is usually preferred to a fixed effects model (section 4.3.1). Sub-group analysis and meta-regression (section 0) can be used to explore heterogeneity (section 4.3.3) arising from clinical and methodological differences in the studies. A classical/frequentist or Bayesian MCMC approach can be used but the Bayesian approach has several advantages (section 4.3.6). The model fit should be assessed (section 4.4) and possible publication bias explored (section 4.5). Section 4.9 introduces models that combine IPD, where available, and aggregate data sources in a pairwise meta-analysis and meta-regression.

To perform a pairwise meta-analysis often different types of interventions are lumped together to give the two intervention groups that are compared. The pairwise meta-analysis in chapter 5 informed the update of a Cochrane Review of "Home safety education and provision of safety equipment for injury prevention". This type of evidence lumping does not always provide the relevant information for decision makers who want to know which specific intervention work the best. A network meta-analysis (section 4.11) allows multiple interventions to be compared to each other using direct evidence, observed in the studies, and indirect evidence, where no study evidence is available.

4.11 Network meta-analysis (NMA)

An extension of standard (pairwise) meta-analysis that enables comparison of more than two evaluated interventions *simultaneously* within a single coherent analysis is network meta-analysis (NMA) (Lumley 2002), also known as mixed treatment comparison (Lu, Ades 2004, Caldwell, Ades et al. 2005). This section describes NMA methods firstly applied to aggregate data sources and then extended to include covariates (section 4.12) and IPD (section 4.13) and IPD and covariates (section 4.14).

NMA estimates the pooled effects where pairwise evidence exists (direct evidence) but also allows estimation of effects where interventions are not directly compared within any primary studies but linked through a connected network of studies (indirect evidence) and where there is consistency across the evidence base. These additional assumptions will be discussed in section 4.11.4.

NMA is being increasingly used in health technology assessment when deciding on the optimal intervention strategy for a given medical condition (Cooper, Peters et al. 2011, Cooper, Kendrick et al. 2012). The relative efficacy of interventions that we can estimate with uncertainty from the NMA can be used in a decision model to evaluate cost-effectiveness. This will be discussed in Chapter 6.

4.11.1 Networks of evidence

If, for example, we wanted to compare the following 4 interventions—usual care (A), education (B), safety equipment giveaway (C), and home safety inspection (D) — this could be achieved by using studies containing the following direct pairwise comparisons: usual care versus education (A vs B), education versus safety equipment giveaway (B vs C), and safety equipment giveaway versus home safety inspection (C vs D) (Figure 4-3I) and by tracing a comparison pathway through the direct pairwise comparisons to estimate, for example, the indirect effect of usual care versus safety equipment giveaway (A vs C) not evidenced in a study. However, the network would be disconnected, and the analysis invalid, if only studies of usual care versus education, and safety equipment giveaway versus home safety inspection existed (Figure 4-3(II)).
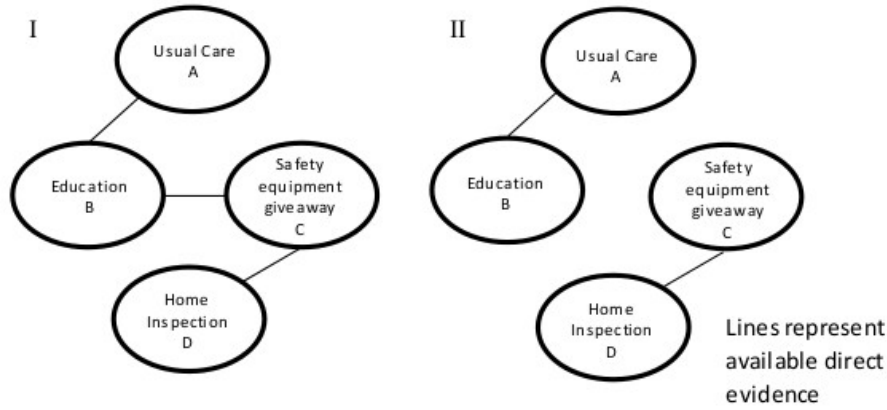
Figure 4-3 Network meta-analysis (I) Connected Network, (II) Disconnected Network

If we refer to the four interventions in Figure 4-3(I) as A, B C and D where A is the usual care reference/baseline category then three intervention effects (for a binomial outcome) can be defined representing the log odds ratios (log OR) of B, C and D relative to A: d_{AB}, d_{AC}, d_{AD} (Welton 2012). These are the basic parameters. The other three intervention effects, d_{BC}, d_{BD}, d_{CD}, can be defined in terms of these basic parameters (equation (4.21) and are referred to as functional parameters.

$$
\begin{aligned}
d_{BC} &= d_{AC} - d_{AB} \\
d_{BD} &= d_{AD} - d_{AB} \\
d_{CD} &= d_{AD} - d_{AC}
\end{aligned}
\tag{4.21}
$$

Equations (4.21) are called the consistency equations (Lu, Ades 2006). For some comparisons of interventions we may not have any studies providing evidence and we will need an "indirect" estimate for the intervention effects, for example the comparison between A and C in Figure 4-3(I) there is no direct evidence but the direct evidence from the studies comparing A to B and B to C can be used to provide an indirect estimate. Where direct evidence from studies is available it can provide a direct estimate of the intervention effect but can also be pooled with indirect evidence. These equations assume evidence consistency, no conflict in evidence, in the direct and indirect intervention effect estimates, where they exist, and checks for inconsistency are presented in section 4.11.4. This model can be extended to any connected network of evidence with any number of interventions.

69

4.11.2 Network meta-analysis model

The random effects meta-analysis model for pairwise comparisons (section 4.3.5) is extended for a network meta-analysis (Higgins, Whitehead 1996, Welton 2012, Dias, Sutton et al. 2013a). Each study i comparing intervention k to b estimates a distinct $log\ OR$, δ_{ibk}, drawn from a common distribution, $N(d_{bk}, \tau_{bk}^2)$. This is simplified by assuming exchangeability across all interventions comparisons, $\tau_{bk}^2 = \tau^2$ and τ is given a vague prior distribution.

Equation (4.22) presents the random effects NMA model with a binary outcome for the comparison of intervention k to intervention b in study i with interventions A, B, C, ... For each study r_{ik} is the number of events of interest observed on intervention k out of n_{ik} observations and p_{ik} is the probability of an event in study i on intervention k.

$$r_{ik} \sim Bin(n_{ik}, p_{ik})$$

$$logit(p_{ik}) = \begin{cases} \mu_{ib} & where\ b = A, B, C, \dots\ if\ k = b \\ \mu_{ib} + \delta_{ibk} & where\ k\ "after"\ b, k = B, C, \dots \end{cases} \quad (4.22)$$

$$\delta_{ibk} \sim N(d_{bk}, \tau^2) \sim N(d_{Ak} - d_{Ab}, \tau^2)$$

where
   μ_{ib} is the log odds of an event in study i on the baseline intervention b
   δ_{ibk} is the study-specific (i) log OR for intervention k compared to intervention b
   d_{bk} is the pooled log OR for intervention k compared to intervention b, with
      $d_{AA} = 0$
      $d_{BC} = d_{AC} - d_{AB}$
      $d_{BD} = d_{AD} - d_{AB}$
      etc
   τ^2 is the between study heterogeneity assumed constant for all intervention comparisons.

Prior distributions are vague and specified to be $\mu_{ib}, d_{Ak} \sim N(0, 10^4)$ for $k = B, C, \dots$ and $\tau \sim Unif(0, 10)$ (Dias, Sutton et al. 2012).

Intervention A is assumed to be the usual care/control reference intervention. If τ^2 is assumed zero then this model is a fixed effects NMA. NMA models can be estimated using MCMC simulation (Caldwell, Ades et al. 2005) with minimally informative prior distributions in WinBUGS (Spiegelhalter, Thomas et al. 2003). As with the pairwise meta-analysis (section 4.3.6) the predictive distribution of the $log\ OR$ can be used to

estimate what we might expect in a new study rather than using the posterior mean of the $log\ OR$. To find the prediction distributions for each pairwise comparison in a network meta-analysis, a prediction for a "new" M-arm study (where M is the number of interventions) need to be generated. One approach, shown in Dias el al (Dias, Welton et al. 2011), is to monitor $\delta_2^{new}, \ldots \delta_M^{new}$, from a multivariate normal distribution, expressed as a series of conditional univariate normal distributions.

Absolute intervention effects (probability of the event of interest) can be estimated for each intervention (derived by using an underlying rate based on the control/usual-care arm) from the NMA results which in turn can be incorporated into a probabilistic decision model, Chapter 6 discusses this further. A Bayesian approach allows interventions to be ranked according to their relative effectiveness and the probability that each intervention is best for a particular outcome is calculated using the posterior distribution of the ranks (Caldwell, Ades et al. 2005, Salanti, Ades et al. 2011). The probability of being best does not take into account uncertainty and the posterior distributions of the ranks are often overlapping and difficult to interpret. A simulation study showed that estimates of rank probabilities are highly sensitive to both the number of studies per comparison and the overall network configuration and recommended that they should be treated with caution (Kibret, Richer et al. 2014). The rankings can be presented in rankograms and interpreted using the surface under the cumulative ranking curve (SUCRA) (Salanti, Ades et al. 2011, Chaimani, Higgins et al. 2013). The SUCRA value for intervention i is the proportion of interventions worse than i and interventions can be ranked on this value, the larger the SUCRA value the better the rank of the intervention. A rankogram is a plot of the probabilities of assuming each of the possible ranks, i.e. the probability distribution of the ranking for an intervention. The NMA results can be presented using forest plots developed by Tan et al (Tan, Cooper et al. 2015) using WinBUGS and R (and the R2WinBUGS command). The forest plot (shown in Figure 5-2) presents the NMA and pairwise estimates, rankograms and SUCRA values, prediction intervals and an estimate of the between study variance.

NMA models are fitted to the studies identified for reporting interventions to increase the possession of a fitted safety gate (section 5.9), reduce the possession and use of a

baby walker (section 5.10), increase the possession of window locks and increase those never leaving their child on a high surface (section 5.11).

4.11.3 Multi-arm studies

The standard NMA random-effects model with a binary outcome (equation (4.22)) can be extended to include studies with 3 or more arms by accounting for the correlation structure (Lu, Ades 2004, Cooper, Sutton et al. 2009, Welton 2012, Dias, Sutton et al. 2013b). When a multi-arm study, for example comparing three interventions A, B and C, is included in a network meta-analysis we need to account for the correlation between the posterior distributions of the $log(OR)$ of comparisons between A and B and between A and C because they both depend on the same study baseline, μ_i. Assuming homogeneity of the variance as before then it can be shown that the correlations are equal to 0.5, hence the covariance is $\tau^2/2$ and the study i intervention effects follow a multivariate normal (MVN) distribution as shown in equation (4.23) (Higgins, Whitehead 1996).

$$\begin{pmatrix} \delta_{iAB} \\ \delta_{iAC} \end{pmatrix} \sim MVN\left(\begin{pmatrix} d_{AB} \\ d_{AC} \end{pmatrix}, \begin{pmatrix} \tau^2 & \tau^2/2 \\ \tau^2/2 & \tau^2 \end{pmatrix} \right) \qquad (4.23)$$

This structure can be used in WinBUGS to generate predictive distributions for every intervention in every study and can be extended to n-arm studies. Using this model it is assumed that the intervention effect being estimated in an AB study is exchangeable with the intervention effects being estimated in AC or BC or CD studies, even if the study does not include all interventions as it is assumed that every study is a multi-arm study with some intervention arms missing. The missing intervention arms are assumed missing at random.

4.11.4 NMA model choice, checking and consistency

The deviance and DIC can be calculated for NMA models as described in section 4.4 and used to investigate goodness of fit. To assess the goodness of fit of the model to the data, the posterior mean residual deviance should be approximately equal to the number of intervention arms across all studies (McCullagh, Nelder 1989, Congdon 2003).   Heterogeneity of the network (variability in intervention effects ($log(OR)$) within pairwise comparisons above that expected by chance) is quantified by using the between study standard deviation parameter, τ^2, where a standard deviation of above 0.5 indicates high heterogeneity and above one substantial heterogeneity as described in section 4.3.6 (Spiegelhalter, Abrams et al. 2004, Cooper NJ, Sutton AJ, Lu G,Khunti K. 2006, Cooper, Sutton et al. 2009, Borenstein, Hedges et al. 2009). This model assumes that the degree of between-study within-comparison heterogeneity is constant across all intervention comparisons in the network (exchangeable).

There have been concerns about the validity of NMA when direct and indirect evidence from different sources is not consistent and hence should not be pooled together (Song, Altman et al. 2003, Song, Xiong et al. 2011). Consistency should be explored to ensure that the indirect estimates of the intervention effect are comparable to the pairwise direct estimates available in the data. Inconsistency can only be investigated where complete loops of evidence exist, that is there is direct evidence between interventions AB, BC and AC. Inconsistency arises, for example, when the direct intervention effect estimate for AC is different to the indirect intervention effect estimate for AC obtained from AB and BC studies. Inconsistency can arise when there is an uneven balance of effect modifiers in the direct and indirect evidence (Dias, Welton et al. 2013a). The Bucher method can be used for testing consistency in single loops of evidence when there are only two-arm studies by comparing the direct evidence to the indirect evidence and calculating an estimate of the inconsistency parameter, ω, and its variance (equation (4.24)).

$$\widehat{\omega}_{AC} = \hat{d}_{AC}^{Direct} - \hat{d}_{AC}^{Indirect}$$

$$Variance(\widehat{\omega}_{AC}) = Var(\hat{d}_{AC}^{Direct}) + Var(\hat{d}_{AB}^{Direct}) + Var(\hat{d}_{BC}^{Direct})$$

(4.24)

A test of no inconsistency can be conducted using a normal based test. Where there are multiple loops in a network independent tests are conducted on each loop but this makes it difficult when three-arm studies are included (Dias, Welton et al. 2010a).

An alternative approach, used in this thesis (section 5.10), is to use node-splitting (Dias, Welton et al. 2010a) in which two posterior distributions are obtained for the mean intervention effect d_{AC} for two interventions A and C; one for the direct only evidence for X compared to Y, d_{AC}^{Direct}, and the other for the indirect evidence only from an NMA of all other studies, $d_{AC}^{Indirect}$, assuming consistency. This split is done for all pairs of interventions (nodes) where there is direct evidence. The variance of the $log(OR)$, τ^2, will have a different posterior estimate for each node-split and can be compared with those obtained from the overall NMA. The posterior distribution of the inconsistency parameter,$\omega_{AC} = d_{AC}^{Direct} - d_{AC}^{Indirect}$, can be used to test the hypothesis of no inconsistency and a p-value generated by monitoring the proportions of times ω_{AC}>0. These node-split models are implemented by calling WinBUGS from the R2WinBUGS package in R (R Development Core Team 2012). These tests tend to have low power and can only detect where inconsistency exists and so the studies included in the loops of evidence displaying inconsistency should be investigated further. Any bias in the indirect estimates must be due to bias in the direct estimates (Caldwell, Ades et al. 2005).

4.12 Network meta-analysis including covariates

There can be heterogeneity and inconsistency in a NMA due to the effect of study covariates, for example a covariate can affect all interventions by the same amount compared to the usual care/control group or there can be confounding between studies if the interventions have been studied in different populations or there is an imbalance in the distribution of the covariate between studies. This can lead to misleading comparisons in an NMA. The intervention by covariate interaction can be explored in the NMA to explain variations in interventions effects between studies within the pairwise comparisons and can also reduce inconsistency (Cooper, Sutton et al. 2009).

There are several possible NMA models that incorporate an *intervention x covariate* interaction that can be fitted that make different assumptions about the covariate effect on the intervention. Cooper et al presented the three model specifications described below with increasingly strong assumptions about the relationship between the covariate effects for each intervention, all prior distributions stated are as given in the paper (Cooper, Sutton et al. 2009). If the number of studies is small compared to the number of interventions comparisons being made then the stronger the assumptions may need to be. The NMA model in equation (4.22) has been extended to the three models (equations (4.25),(4.26) and (4.27)) presented in sections 4.12.1 - 4.12.3 to include the aggregated study i covariate, X_i, and the differences are highlighted in bold.

4.12.1 Model 1: NMA including covariates: independent regression coefficient

This model given in equation (4.25) assumes that all *intervention x covariate* interactions are different for each pairwise comparison of intervention vs the baseline intervention and independent of each other. For a network with I studies, K interventions and D data points, this model requires the estimation of a high number of parameters: i baselines (μ_i), (K-1) intervention effect means (d_k), (K-1) regression coefficients and the between study variance (τ^2).

$$r_{ik} \sim Bin(n_{ik}, p_{ik}) \quad \text{for study } i, \text{ intervention } k$$

$$logit(p_{ik}) = \begin{cases} \mu_{ib} & where\ b = A, B, C, \dots\ if\ k = b \\ \mu_{ib} + \delta_{ibk} & where\ k\ "after"\ b, k = B, C, \dots \end{cases} \quad (4.25)$$

$$\delta_{ibk} \sim N\big(d_{bk} + \boldsymbol{\beta_{bk}} X_i,\ \tau^2\big) \sim N(d_{Ak} - d_{Ab} + (\boldsymbol{\beta_{Ak}} - \boldsymbol{\beta_{Ab}})X_i, \tau^2)$$

where
- β_{bk} is the change in the log odds of an event per unit change in the covariate value X_i in study i for intervention k compared to intervention b, with $\beta_{AA} = 0$ and $\beta_{bk} = \beta_{Ak} - \beta_{Ab}$
- d_{bk} is the pooled log OR for intervention k compared to intervention b when the covariate value is zero (or the mean if the covariate has been centred on the mean value), with $d_{AA} = 0$

Prior distributions are vague and specified to be $\mu_{ib}, d_{Ak}, \beta_{Ak} \sim N(0, 10^4)$ for $k = B, C, \dots$ and $\tau \sim Unif(0, 10)$.

4.12.2 Model 2: NMA including covariates: exchangeable regression coefficient

This model given in equation (4.26) assumes that all *intervention x covariate* interactions are different for all interventions but exchangeable (section 4.3.4) so is a simplification to equation (4.25).

$$r_{ik} \sim Bin(n_{ik}, p_{ik}) \quad \text{for study } i, \text{ intervention } k$$

$$logit(p_{ik}) = \begin{cases} \mu_{ib} & where \; b = A, B, C, \dots \; if \; k = b \\ \mu_{ib} + \delta_{ibk} & where \; k \; \text{"after"} \; b, k = B, C, \dots \end{cases}$$

(4.26)

$$\delta_{ibk} \sim N(d_{bk} + \boldsymbol{\beta_{bk}X_i}, \tau^2) \sim N(d_{Ak} - d_{Ab} + (\boldsymbol{\beta_{Ak}} - \boldsymbol{\beta_{Ab}})\boldsymbol{X_i}, \tau^2)$$

$$\beta_{Ak} \sim N(B, \sigma_B^2)$$

where

$$d_{AA}, \beta_{AA} = 0$$

Prior distributions are vague and specified to be $\mu_{ib}, d_{Ak}, B \sim N(0, 10^4)$ for $k = B, C, ..$ and $\tau, \sigma_B \sim Unif(0, 10)$.

4.12.3 Model 3: NMA including covariates: common regression coefficient

This model given in equation (4.27) assumes that all *intervention x covariate* interactions are identical, share a common beta, for each pairwise comparison of intervention vs the baseline intervention and is a further simplification of equations (4.25) and (4.26). This model estimates only one additional parameter value, the common β, and hence demands the least amount of data.

$$r_{ik} \sim Bin(n_{ik}, p_{ik}) \quad \text{for study } i, \text{ intervention } k$$

$$logit(p_{ik}) = \begin{cases} \mu_{ib} & where \; b = A, B, C, \dots \; if \; k = b \\ \mu_{ib} + \delta_{ibk} & where \; k \; \text{"after"} \; b, k = B, C, \dots \end{cases}$$

(4.27)

$$\delta_{ibk} \sim \begin{cases} N(d_{Ak} + \boldsymbol{\beta X_i}, \tau^2) \sim N(d_{Ak} - d_{AA} + \boldsymbol{\beta X_i}, \tau^2) & if \; b = A \\ N(d_{bk}, \tau^2) \sim N(d_{Ak} - d_{Ab}, \tau^2) & if \; b \neq A \end{cases}$$

where

$$d_{AA} = 0$$

Prior distributions are vague and specified to be $\mu_{ib}, d_{Ak}, B \sim N(0, 10^4)$ for $k = B, C, ..$
and $\tau \sim Unif(0, 10)$.

These models (equations (4.25)-(4.27)) are applied to the accident prevention data in section 5.12. The posterior mean residual deviance can be used measure the goodness of fit and the DIC can be used to compare competing models.

4.13 Network meta-analysis using IPD and aggregate data

In section 4.9 a pairwise meta-analysis was introduced using IPD, where available, and aggregate data. In this section a NMA is extended to include the available IPD using methods developed by Saramago et al (Saramago, Sutton et al. 2012) to extend the pairwise IPD and aggregate data meta-regression described in section 4.9 (Sutton, Kendrick et al. 2008). The random effects model is described in four parts representing the different available types of data (IPD for individually allocated studies, IPD for cluster allocated studies, aggregated studies (both individually and cluster allocated)) and the combination of the effect estimates for interventions A, B, ... All prior distributions are as given by Saramago et al (Saramago, Sutton et al. 2012).

Part 1: Individually allocated IPD studies

$$Y_{sik} \sim Bernoulli(p_{sik})$$

$$logit \, p_{sik} = \begin{cases} \mu_{ib}^{IPD} & b = A, B, .. \text{ if } k = b \\ \mu_{ib}^{IPD} + \delta_{ibk} & where \, k \text{ "after" } b, k = B, C, ... \end{cases}$$

(428)

Where
  s = 1, 2, … (no. participants in the ith individually allocated IPD study)
  i = 1, 2, … individually allocated IPD studies
  k = 1, 2, …number of interventions
  Y_{sik} is the response for the sth participant in the ith study on intervention k

Prior distributions

$$\mu_{ib}^{IPD} \sim N(0, 10^6)$$

Part 2: Cluster allocated IPD studies

$$Y_{smik} \sim Bernoulli(p_{smik})$$

$$logit\ p_{smik} = \begin{cases} \mu_{mib}^{c.IPD} & b = A, B,.. \ if\ k = b \\ \mu_{mib}^{c.IPD} + \delta_{ibk} & where\ k\ "after"\ b,\ k = B, C, \dots \end{cases}$$ (4.29)

where

    m = 1, 2, …number of clusters in the ith study

    i = (no. individually allocated IPD studies), …., (no. individually allocated IPD studies +no. cluster allocated IPD studies)

    Y_{smik} is the response for the sth participant in the kth cluster in the ith study on the kth intervention

Prior distributions

$$\mu_{mib}^{c.IPD} \sim N(\psi_i, \tau.cluster_i^2) \qquad \text{assuming exchangeability within studies and independence between studies}$$

$$\psi_i \sim N(0, 10^6)$$
$$\tau.cluster_i \sim Unif(0,10)$$

Part 3: Aggregate data (AD) studies (both clustered and individually allocated)

$$r_{ik} \sim Bin(n_{ik}, p_{ik})$$

$$logit\ p_{ik} = \begin{cases} \mu_{ib}^{AD} & b = A, B,.. \ if\ k = b \\ \mu_{ib}^{AD} + \delta_{ibk} & where\ k\ "after"\ b,\ k = B, C, \dots \end{cases}$$ (4.30)

where

    i = (no. individually allocated IPD studies +no. cluster allocated IPD studies+1), … (total no. studies)

    r_{ik} is the number of events in the ith study on the kth intervention

This model assumes that the data has been adjusted for clustering prior to model fitting unlike the model in section 4.9, but it can be extended to perform the adjustment within the model.

Part 4: Combining estimates

$$\delta_{ibk} \sim N(d_{bk}, \tau^2) \sim N(d_{Ak} - d_{Ab}, \tau^2)$$ (4.31)

where

i = 1, 2, … (total no. studies)

A random effect is placed on all intervention effect estimates, δ_i ($\log OR_i$), from parts 1-3 assuming exchangeability across studies and a normal distribution.

4.14 Network meta-analysis using IPD and AD including a covariate

The models in section 4.13 can be extended, using the methods described by Saramago et al, to include a covariate to explain between-study heterogeneity and reduce inconsistency in a network (Saramago, Sutton et al. 2012). The models include individual participant level covariate terms in parts 1 and 2 and a study level covariate value in model 3. In section 4.12, where a covariate is added to the NMA model, three different assumptions were considered about the *intervention x covariate* interactions (Cooper, Sutton et al. 2009). These assumptions can be applied to the IPD and AD NMA models although the assumption that the interactions are independent for all interventions requires a lot of data and hence for the application in section 5.13, because there are only a few studies for each covariate, are not considered. For the individually allocated IPD data (clustered allocated in brackets) three additional terms are added to the models, assuming exchangeable *intervention x covariate* interactions between interventions (but could be modified to assume they are the same) and mean centring the covariate values:

(1) a study specific individual-level covariate regression term $\beta_{0i} x_{si}$ ($\beta_{0i} x_{smi}$), where x_{si} (x_{smi}) is the binary covariate value (e.g. 1=male, 0=female) for participant s in study i (cluster m);

(2) within study interaction term $\beta_{bk}^w (x_{si} - \bar{x}_i)$ ($\beta_{bk}^w (x_{smi} - \bar{x}_i)$), where $\bar{x}_i$ is the mean covariate value for study i and $\beta_{bk}^w \sim N(\beta_{Ak}^w - \beta_{Ab}^w, \sigma_{Bw}^2)$ with prior $\sigma_{Bw} \sim \text{Unif}(0,10)$;

(3) between study interaction term $\beta_{bk}^B \bar{x}_i$

These three terms are similar to the terms added to the IPD and AD meta-regression model presented in section 4.9.3 and the difference in β_{bk}^W and β_{bk}^B estimates ecological bias (Riley, Steyerberg 2010).

4.15 Summary

This chapter presents methods to evaluate the effectiveness of interventions in increasing order of complexity. Initially a pairwise meta-analysis is described to compare a single enhanced intervention arm to a control/usual care arm (sections 4.2-4.8). A network meta-analysis is introduced where multiple interventions can be compared in a network of evidence. NMA provides estimates of effects where pairwise evidence between interventions is available and also where interventions are not directly compared but linked through a network of evidence (section 4.11). If there is heterogeneity between studies then covariates can be included in a meta-regression analysis to explain the heterogeneity (section 4.6). IPD, where available, can be incorporated into the analysis to increase power and avoid ecological bias when covariates are aggregated (sections 4.9, 4.13 and 4.14). All models make assumptions that should be stated and model fit should be investigated. WinBUGS is used to analyse the data using a Bayesian MCMC simulation approach (section 4.1) as it provides a more flexible approach when fitting complex models and model fit can be assessed.

5 Effectiveness analyses applied to accident prevention.

This chapter presents the results of the meta-analysis methods described in chapter 4 applied to studies identified in the systematic review of studies reporting interventions to increase safety practices to prevent falls in children under 5 in the home. The chapter starts by describing the studies identified in the review. The results of pairwise meta-analyses, using a frequentist and Bayesian MCMC simulation approach, and meta-regression on the aggregate data and using IPD where available are presented. This is followed by the results of network meta-analyses. The results are reported in the following published documents and presentations:

- NIHR report for the Keeping Children Safe at Home programme (Kendrick, Ablewhite et al. 2017),
  https://www.journalslibrary.nihr.ac.uk/pgfar/pgfar05140/#/abstract
  Contribution included: pairwise meta-analysis, meta-regression with IPD and network meta-analysis for falls prevention interventions.
- Update of the Cochrane review on home safety education and provision of safety equipment for injury prevention (Kendrick, Young et al. 2012),
  http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD005014.pub3/pdf
  Contribution included: pairwise meta-analyses and meta-regressions using IPD.
- Exploration of synthesis methods in public health evaluations of interventions paper (Achana, Hubbard et al. 2014) Joint First Author (Appendix U)
  Contribution included: discussion on the review, pairwise meta-analysis, meta-regression with IPD and network meta-analysis for interventions to increase the possession and use of safety gates to prevent falls.
- NMA evaluating the effectiveness of interventions to prevent falls paper (Hubbard, Cooper et al. 2015). (Appendix U) This paper focused on the network meta-analyses of interventions to prevent falls injuries. Contribution included: all statistical analyses, writing the paper.
- Presentation at the Population Health Conference 2012 in Birmingham (Appendix T).

- Poster presentation at ISCB (International Society for Clinical Biostatistics) conference Birmingham 2016 and at the University of Leicester Festival of Postgraduate Research 2017 (Appendix S).

The objective of the meta-analyses was to evaluate the effectiveness of different interventions to increase the possession of safety equipment by households to prevent falls in children under 5 in the home or increase falls prevention behaviours.

The analyses were conducted using Stata (StataCorp. 2017), R (R Core Team (2013)) and using a MCMC approach in WinBUGS (Lunn, Thomas et al. 2000). Analyses and reporting adhere to guidelines including the PRISMA and PRISMA-IPD statements (Stewart, Clarke et al. 2015), NICE Evidence Synthesis of Treatment Efficacy in Decision Making: A Reviewer's checklist (Ades, Caldwell et al. 2013) and the criteria for reporting the results of NMA (Bafeta, Trinquart et al. 2014). Sensitivity analyses on the choice of initial values and prior distributions were conducted for models fitted in WinBUGS. Convergence is assessed by examination of the trace and autocorrelation plots and the Rubin-Gelman statistic. The number of simulations and length of the burn-in varies depending on the complexity of the analysis but the effect of changing it is explored in sensitivity analyses. Not all model convergence assessments and sensitivity analyses conducted are reported in this thesis but examples have been given to show the process and any issues have been described.

5.1   Studies in the meta-analyses

A published overview of systematic reviews and a systematic review of primary studies (Young, Wynn et al. 2013) identified 16 primary studies eligible for inclusion in the meta-analyses in this chapter. They have been described in section 3.5.1 and Table 3-1.

The main events of interest reported for households in these studies are: possession of fitted safety gate(s) on the stairs (12 studies); possession (5 studies) or use (4 studies) of a baby walker; possession of window locks or windows with limited opening (6 studies); never leaving a child alone on a high surface (3 studies); and possession of a bath-mat (3 studies) or decals (1 study). For the baby walker outcome, baby walker

possession and use are combined as one of the included studies (Kendrick, Illingworth et al. 2005) found that 94% of those owning a walker used it and 96% of those who used a walker owned one. Table 5-1 presents the studies included in the preventions of falls meta-analyses and the reported number of events.

One of the studies, Babul 2007 (Babul, Olsen et al. 2007), is a three-arm study and is reported in Table 5-1 as three separate comparisons. Four of the studies, Kendrick 2005 (Kendrick, Illingworth et al. 2005), Kendrick 1999 (Kendrick, Marsh et al. 1999), Petridou 1997 (Petridou, Tolma et al. 1997) and Gielen 2002 (Gielen, McDonald et al. 2002), are cluster randomised and so the numbers are adjusted for clustering (section 4.8) as the original studies did not adjust for clustering. ICCs calculated from IPD are available for the studies by Kendrick 1999 and 2005. Gielen and Petridou are adjusted using an ICC estimated from studies with similar allocation (published or IPD available). For Gielen the midpoint of published ICCs for injury outcomes at GP, midwife or health visitor level is used and for Petridou the midpoint of published ICCs for injury outcomes at health authorities, local authorities or town level is used (Kendrick, Young et al. 2012).

There were six covariates of interest: age of the child, gender, accommodation tenure, single parent, black or minority ethnic, parents unemployed. Not all studies reported all covariates (Table 3-3). Individual participant data was provided for thirteen studies (Table 3-1).

This chapter will focus on the effectiveness of interventions to increase the possession of a fitted safety gate to prevent stair falls to illustrate the methods as this is the outcome that had the most evidence (12 studies out the of the 16 presented in Table 5-1 with 5,206 participants). The other outcome results will be summarised and discussed in more detail where they illustrate something different to the safety gate outcome, including the baby walker outcome where there is a three arm study.

Table 5-1 Summary of Studies and their data included in the meta-analyses of the interventions to prevent falls injuries in children under 5 (continued overleaf)

First author,Year	Study design	Follow-up (months)	Study quality	Comparison (intervention number)	Intervent ion ±	Safety gate		Baby walker		Window locks		High surfaces		Bath mats	
						Number of households with safety gates*	Total number of households	Number of households with or using baby walkers*	Total number of households	Number of households with window locks	Total number of households	Number of households not leaving child on high surface	Total number of households	Number of households with bath mats*	Total number of households
Nansel 2002[e]	RCT	0.75	A=Y, B=Y, F=Y	Usual care (1) **vs.** Education (2)	1 2	70 76	89 85	30 19	89 85						
Kendrick 2005	RCT	9	A=Y, B=N, F=Y		1 2	418(348.44) 373(310.93)	524(436.80)[a] 452(376.78)	230(105.27) 131(59.96)	543(248.52)[a] 463(211.90)						
Nansel 2008[f]	Non-RCT	1	B=N, F=N, C=N		1 2	29 60	38 69	12 13	38 69			21 55	24 62		
Tan 2004[g]	Non-RCT	5	B=U, F=Y, C=Y		1 2			393 143	480 228						
Babul 2007 [d]	RCT	10	A=Y, B=N, F=N	Usual care (1) **vs.** Education + Low/free equipment (3)	1 3			31 22	148 162			69 89	148 161		
Clamp 1998	RCT	1.5	A=Y, B=N, F=Y		1 3	50 52	69 64			72 80	82 83				
McDonald 2005	RCT	1	A=Y, B=U, F=N		1 3	10 23	41 54								
Babul 2007 [d]	RCT	10	A=Y, B=N, F=N	Usual care (1) **vs.** Education + Low/free equipment + Home safety inspection (4)	1 4			31 26	148 173			69 84	148 170		
Kendrick 1999	Non-RCT	25	B=N, F=N, C=Y		1 4	241(214.26) 251(223.15)	364(323.61)[b] 364(323.61)			339 323	366 362				
Hendrickson 2002	RCT	1.5	A=N, B=N, F=Y		1 4					21 24	39 34				
Watson 2005	RCT	12	A=Y, B=N, F=N	Usual care (1) **vs.** Education + Low/free equipment + Fitting (5)	1 5	328 408	718 742			493 550	741 767				
Petridou 1997	CBA	20	B=N, F=Y, C=Y	Usual care (1) **vs.** Education + Home safety inspection (6)	1 6									64(50.44) 66(48.91)	128(100.12)* 131(97.83)

Table 5-1 (continued) Summary of Studies and their data included in the meta-analyses of the interventions to prevent falls injuries in children under 5

First author,Year (Reference No.)	Study design	Follow-up (months)	Study quality	Comparison (intervention number)	Intervention ±	Safety gate		Baby walker		Window locks		High surfaces		Bath mats	
						Number of households with safety gates*	Total number of households	Number of households with or using baby walkers	Total number of households	Number of households with window locks	Total number of households	Number of households not leaving child on high surface	Total number of households	Number of households with bath mats	Total number of households
Phelan 2010]	RCT	12	A=Y, B=N, F=Y	Usual care (1) **vs.** Education + Low/free equipment + Fitting + Home safety inspection (7)	1 7	78 131	147 146	29 24	138 140	145 146	150 149			59 56	149 150
Posner 2004	RCT	2.5	A=Y, B=Y, F=N	Education (2) **vs.** Education + Low/free equipment (3)	2 3	25 28	47 49	4 4	8 7			44 12	50 17	34 44	47 49
Sznajder 2003	RCT	1.5 to 2	A=Y, B=N, F=Y	Education (2) **vs.** Education + Low/free equipment + Fitting (5)	2 5	45 44	50 47	14 19	50 47					37 31	49 48
Gielen 2002	RCT	12	A=U, B=U, F=N	Education + low/free equipment (3) **vs.** Education + low/free equipment + Home safety inspection (4)	3 4	11(12.85) 13(10.87)	48(47.44) c 48(47.44)								
Babul 2007d	RCT	10	A=Y, B=N, F=N		3 4			22 26	162 173			89 84	161 170		
King 2001	RCT	12	A=Y, B=Y, F=Y	Education + Low/free equipment + Home safety inspection (4) **vs.** Education + Home safety inspection (6)	4 6	158 166	482 469	29 33	482 469	285 299	469 482				

Abbreviations: RCT, randomized clinical trial; CBA, controlled before-and-after study; A = adequate allocation concealment (RCT only); B = blinded outcome assessment; C, prevalence of confounders does not differ by more than 10% between treatment arms (non-RCT and CBA only); F = at least 80% participants followed up in each arm (not CBA); Y= yes; N = no; U = unclear

*Numbers adjusting for clustering in parentheses

a ICC calculated from IPD from Kendrick 2005[20]

b ICC calculated from IPD from Kendrick 1999[39]

c ICC calculated from IPD from Kendrick 1999[39] and Kendrick 2005[20]

d Babul has been included in the NMA as a three-arm trial but is listed above as three separate comparisons.

e Two intervention arms are combined (tailored advice and tailored advice + care provider feedback)

f Generic safety advice is counted as usual care

g Two control arms are combined (usual care and usual care + baseline questionnaire)

± Interventions: 1. Usual care, 2. Education, 3. Education + low cost (i.e. voucher) / free equipment, 4. Education + low cost equipment (i.e. voucher) / free equipment + home safety inspection, 5. Education + low cost (i.e. voucher) / free equipment + fitting, 6. Education + home safety inspection, 7. Education + low cost (i.e. voucher) / free equipment + fitting + home safety inspection

5.2   Pairwise meta-analysis

A pairwise meta-analysis is used to address the question "do enhanced interventions offered in a healthcare setting increase the possession of a fitted safety gate to prevent falls down stairs compared with a usual care intervention". The enhanced interventions and usual care interventions both consisted of a range of interventions which are lumped together to form the two groups that can be compared in a pairwise meta-analysis and are very heterogeneous (Figure 5-1). The results from the pairwise analysis are reported in the update of the Cochrane review (Kendrick, Young et al. 2012). In this section the results of a random effects model fitted using a frequentist approach in Stata and a Bayesian MCMC simulation approach in WinBUGS is presented using the methods described in section 4.3. The Bayesian MCMC simulation approach has then been extended to more complex methods in the rest of this chapter.
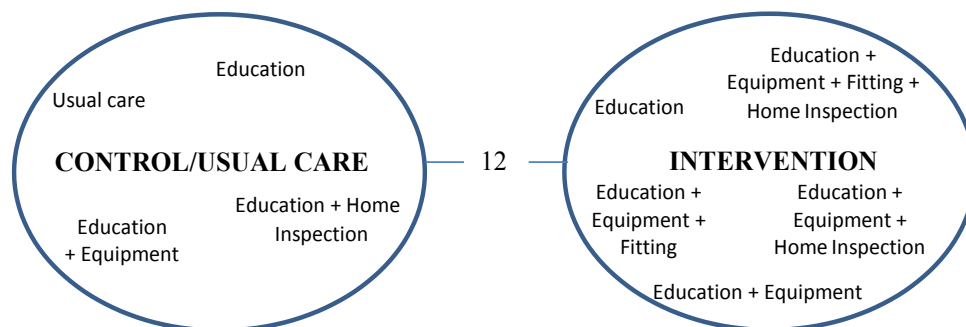
Figure 5-1 Pairwise meta-analysis comparison group interventions to increase the possession of fitted safety gates

A random effects model is used to estimate an overall pooled odds ratio (OR) due to the heterogeneity observed in the interventions and study designs (Table 3-1 and Table 5-1). Model fit (section 4.4) for the fixed and random effects models fitted in WinBUGS (the WinBUGS code is presented in Appendix A) is presented in Table 5-2 and shows that the random effects model is a better fit with a lower DIC (difference=33.6). Studies with a clustered design are adjusted for clustering before analysis (Table 5-1).

Table 5-2 Model estimates and fit for fixed and random effects pairwise meta-analysis (WinBUGS)

	$\bar{D}$	$\tilde{D}$	p_D	DIC	$\bar{D}_{res}$
Random Effects $OR = 1.65(95\%CrI: 1.07, 2.56)$ $\tau^2 = 0.34(95\%CrI: 0.10, 1.17)$	140.2	112.0	21.2	161.4	23.0
Fixed Effects $OR = 1.38(95\%CrI: 1.22, 1.55)$	182.0	169.0	13.1	195.0	64.8

Using Stata, enhanced healthcare interventions are shown to be effective in increasing the possession of a fitted safety gate (OR=1.61, 95%CI: 1.19, 2.17), the forest plot is given in Figure 5-2. There is statistically significant heterogeneity (p<0.001) with an $I^2 = 75.8\%$ and estimated between-study variance $\tau^2 = 0.17$. The WinBUGS results, using a burn-in period of 10,000 iterations followed by 20,000 iterations are similar giving OR=1.65 (95%CrI: 1.07, 2.56) and τ^2=0.34 (95%CrI: 0.11, 1.17) (Table 5-2) and there were no problems with convergence (Figure 5-3). Vague prior distributions were used so the results are expected to be similar as the data dominates (section 4.1).
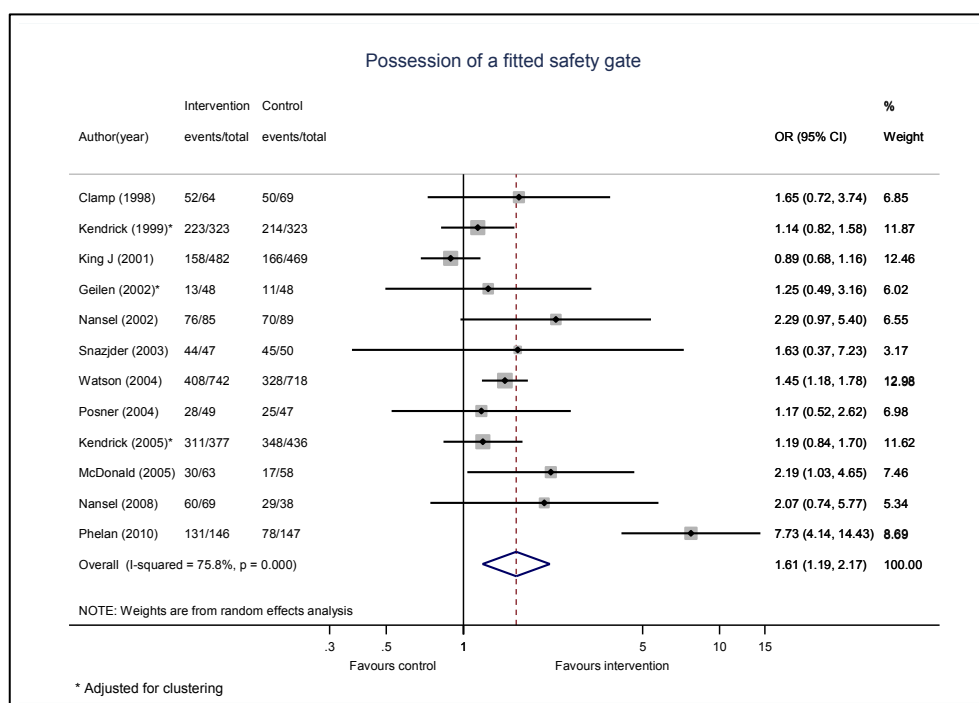
Figure 5-2 Forest plot (Stata) showing the overall pooled OR for possession of a fitted safety gate

The Phelan study (Figure 5-2 and study 12 in Figure 5-4) shows a much higher OR than the other studies (Phelan, Khoury et al. 2010). When this study is removed from the meta-analysis in Stata the overall OR reduces to 1.28(95%CI: 1.07, 1.53) but the enhanced intervention is still statistically significantly more effective than the usual care intervention. This study compares usual care to an intervention that included education and home safety visits where multiple free accident prevention devices are installed; the most intensive intervention of any of the studies. Cross-validation is used in WinBUGS (Appendix A) to examine the effect of removing this study. A Bayesian p-value of 0.0005 indicated that this study is giving outlying results compared to the other studies and hence the need for more complex methods to account for the different intervention intensities; network meta-analysis will be presented from section 5.9.

Due to there being moderate evidence of heterogeneity between studies a predictive distribution (section 4.3.6) from WinBUGS is used to estimate the underlying effect in a new study. The results are illustrated in the caterpillar plot in Figure 5-4 along with the individual shrunken study (i) intervention effect estimates ($logOR_i$) (section 4.3.6). The estimate of the effect in a new study using the predictive distribution (OR(new study)=1.66 (95%CrI: 0.42, 6.82)) has a much wider credible interval as it takes into account between study heterogeneity, τ^2. The heterogeneity is investigated using sub-group analyses and meta-regression in section 5.3.
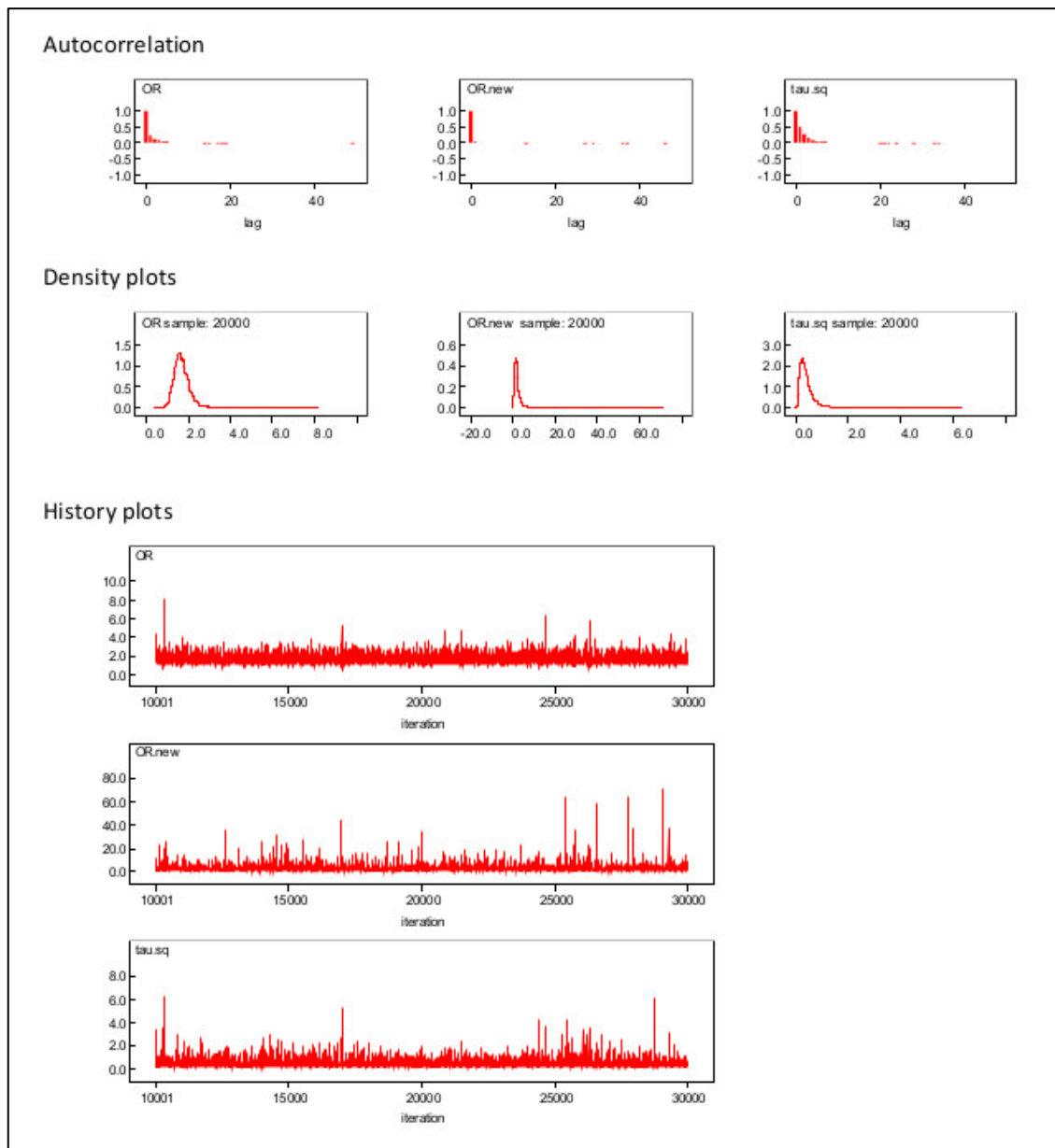
Figure 5-3 WinBUGS assessment of convergence for the random effects meta-analysis of studies with interventions to increase the possession of a fitted safety gate
 (OR is the estimate of the overall odds ratio, OR.new is the estimate of the OR for the predictive distribution of the effect in a new study and tau.sq is estimate the between study variance)
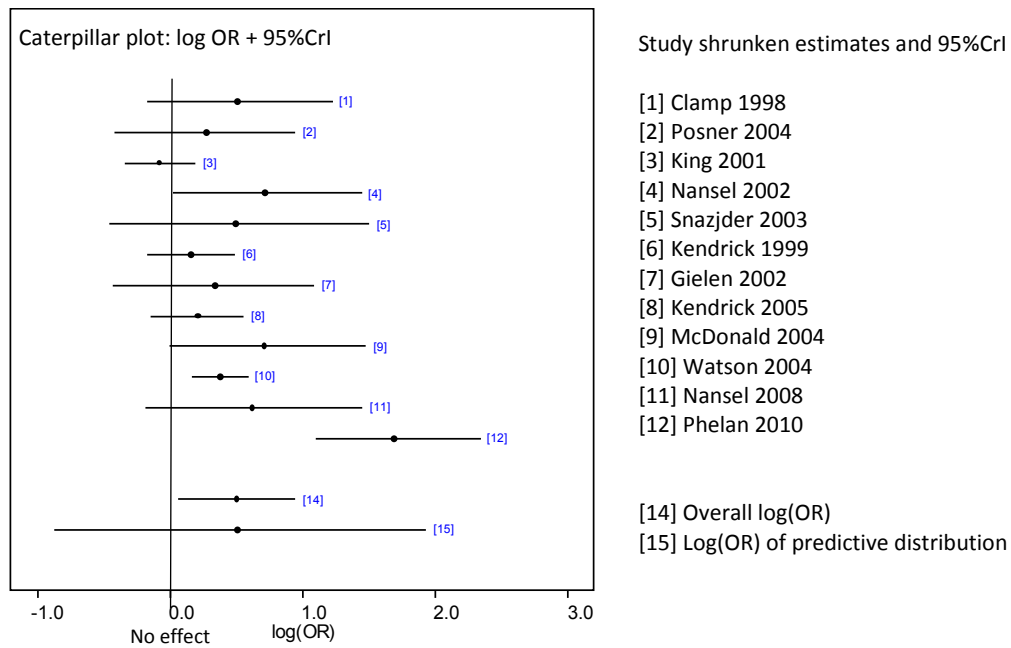
Figure 5-4 Caterpillar plot from the WinBUGS results for the meta-analysis of studies with interventions to increase the possession of a fitted safety gate
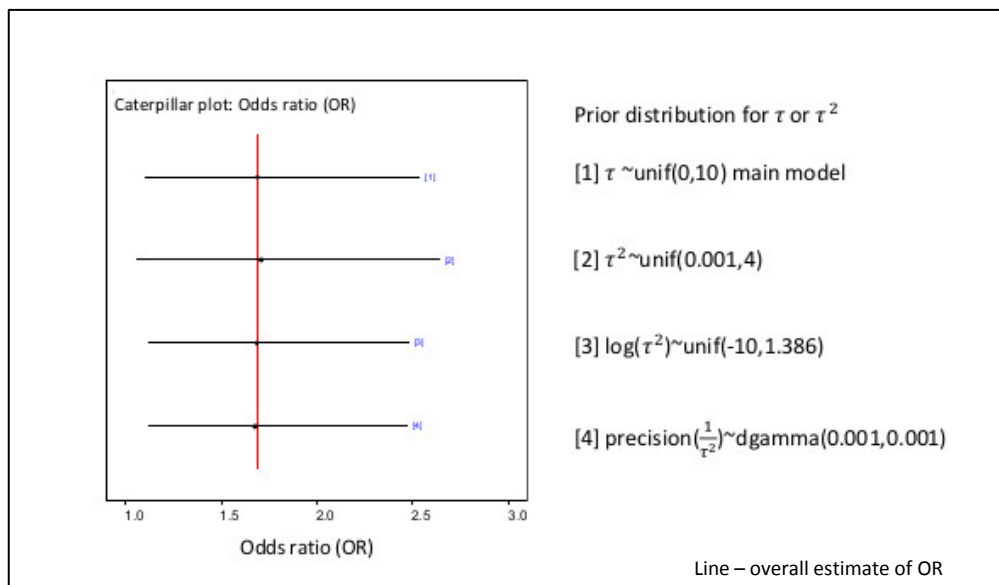
Figure 5-5 Sensitivity analysis on the prior distribution for the between study variance , τ^2, OR with 95%CrI

For the WinBUGS analysis, sensitivity analyses on different prior distributions for between study variance, all give similar OR and CrI (Figure 5-5). When three chains with different initial values were used, the ratios of the between- and within-variability measures converged to zero and the between- and within-variability

measures converged to stability in the Brooks-Gelman-Rubin diagnostic plots (Figure 5-6). The Monte Carlo errors were less than 5% of the posterior standard deviation for all estimates.
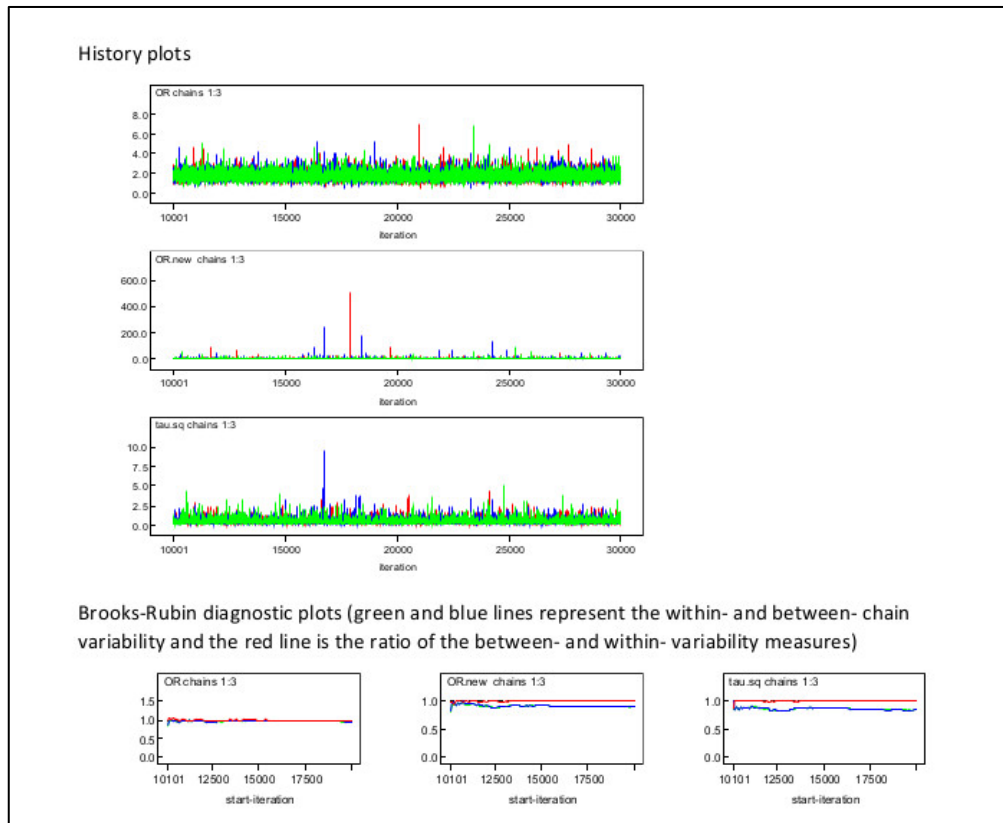
Figure 5-6 Sensitivity analysis on the initial values using three chains in WinBUGS for three chains with different initial values

5.2.1   Publication bias

To investigate for publication bias, Stata is used to produce a funnel plot and contour enhanced funnel plot Figure 5-7.

The plots show a lack of studies on the left-hand side of the funnel plots indicating that there could be issues with heterogeneity and/or publication bias. This could be due to the high OR for the Phelan (2010) study. There was no statistically significant evidence of publication bias (Peter's test (p=0.36) and Rucker's test (p=0.53)) (Appendix B) and

when the Phelan study was excluded the evidence of publication bias was reduced so may have been due to heterogeneity in the study interventions (Figure 5-7).
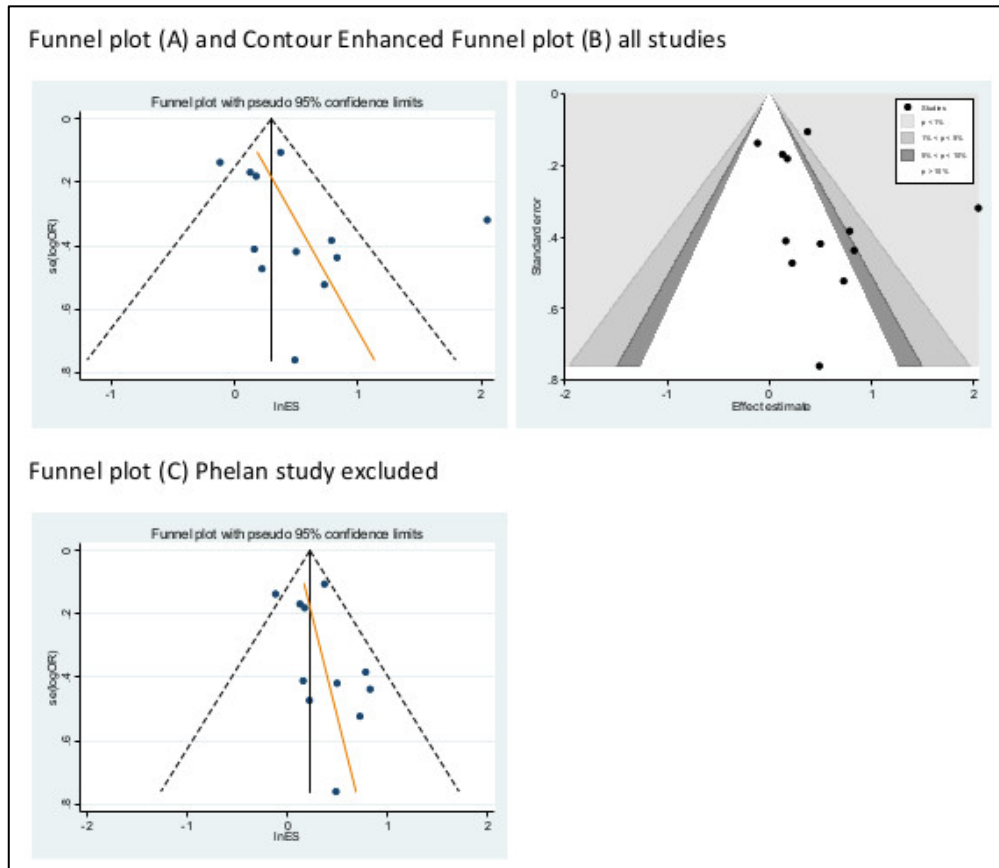
Figure 5-7 Funnel Plot (A and C) and Contour Enhanced Funnel (B) plot for the meta-analysis of studies with interventions to increase the possession of a fitted safety gate
(lnES and Effect estimates are the log(OR)values for the studies)

5.3 Sub-group analyses and meta-regression

5.3.1 Sub-group analyses

Using Stata, sub-group analyses (section 4.6.1) are conducted to investigate any differences in the types of studies included in the meta-analyses and to see if the differences accounted for heterogeneity in the overall results. The results are shown in Table 5-3.

Table 5-3 Sub-group meta-analyses using Stata

Subgroup		Studies	OR (95% CI)	Heterogeneity I^2 (p-value) τ^2
Provided a safety gate	Yes	4 Clamp(1998), Kendrick(1999), Watson(2004), Phelan(2010)	2.05 (1.08, 3.89)	89.8% (<0.001) 0.36
	No	8 King J (2001)*, Gielen(2002)*, Nansel(2002), Snazjder(2003), Posner(2004), Kendrick(2005), McDonald(2005)*, Nansel(2008)	1.28 (0.97, 1.68)	32.4% (0.17) 0.044
Setting of intervention	Home	4 King J (2001), Gielen(2002), Snazjder(2003), Phelan(2010),	1.95(0.57, 6.64)	92.4% (<0.001) 1.36
	Clinical	8 Clamp(1998), Kendrick(1999), Nansel(2002), Posner(2004), Watson(2004), Kendrick(2005), McDonald(2005), Nansel(2008),	1.38(1.20, 1.60)	<0.01% (0.543) <0.01
RCT	Yes	10 Clamp(1998), King J (2001), Gielen(2002), Nansel(2002), Snazjder(2003), Posner(2004), Watson(2004), Kendrick(2005), McDonald(2005), Phelan(2010)	1.67 (1.17, 2.41)	79.4% (<0.001) 0.22
	No	2 Kendrick(1999), Nansel(2008)	1.25 (0.81, 1.92)	16.1% (0.28) 0.03
RCT with adequate concealment	Yes**	9 Clamp(1998), King J (2001), Nansel(2002), Snazjder(2003), Posner(2004), Watson(2004), Kendrick(2005), McDonald(2005), Phelan(2010)	1.72(1.17, 2.53)	81.7% (<0.001) 0.24
RCT with blinding	Yes	2 King J (2001), Posner(2004)	0.92(0.71, 1.18)	<0.01% (0.52) <0.01
	Not blinded	8 Clamp(1998), Gielen(2002), Nansel(2002), Snazjder(2003), Watson(2004), Kendrick(2005), McDonald(2005), Phelan(2010)	1.96(1.28, 3.00)	76.4% (<0.001) 0.24

Table 5-3 (continued) Sub-group meta-analyses using Stata

Subgroup		Studies	OR (95% CI)	Heterogeneity I^2 (p-value) τ^2
RCT and 80%+ follow-up	Yes	6 Clamp(1998), King J (2001), Nansel(2002), Snazjder(2003), Kendrick(2005), Phelan(2010)	1.88(0.98, 3.63)	87.9% (<.001) 0.52
	No	4 Gielen(2002), Posner(2004), Watson(2004), McDonald(2005)	1.46(1.21, 1.77)	<0.01% (0.68) <0.01

* gave voucher for free equipment rather than directly giving a safety gate
** only one RCT without adequate blinding

Families in studies with interventions that included provision of a free safety gate are more likely to possess fitted safety gates than in the control/usual care group of these studies (OR=2.05 (1.08, 3.89)) but this is not shown in studies where the intervention did not include the free safety gate. There is no statistically significant difference between the enhanced intervention and usual care groups in the home setting (OR=1.95(0.57, 6.64)) but there is increased possession in the enhanced intervention group in the clinical setting (OR=1.38(1.20, 1.60)). In the ten RCTs, families in the enhanced intervention group are more likely to possess a fitted safety gate (OR=1.67 (1.17, 2.41)) and this is also shown in the nine RCTs with adequate concealment. Only two of the RCTs are blinded and these did not provide any evidence of a difference in the groups and similarly with the RCTs with 80% or higher follow-up. The four RCTs with less than 80% follow-up did show some evidence of a difference between the groups (OR=1.46(1.21, 1.77)) but this result is questionable as RCTs with 80% or higher follow-up did not demonstrate this difference. All sub-group analyses that included the Phelan (2010) study showed some heterogeneity between the study results. Using sub-group analysis it is difficult to assess the effect of the categorical covariate on the overall effect size, separate effect sizes are given and the difference between the groups with associated uncertainty is not estimated, hence meta-regression is used in the next section.

5.3.2 Meta-regression

Meta-regression (section 4.6.2) is used to investigate the effect of the enhanced intervention compared to usual care by participant characteristic covariates. The covariates chosen are known to be associated with risk of injury: child age (study mean or median), gender (% male child), ethnic group (% black or minority ethnicity (BME)) and single parent family (% single parent). Young mothers (% where mother was under 20 at the birth of the child) is only reported in four studies so is not included in the meta-regression analyses. As indicators of deprivation, housing tenure (% residing in rented accommodation) along with parental unemployment (% with at least one parent not in paid employment) are used. Studies reporting these characteristics are shown in Table 3-3. Not all covariates are recorded for all studies so a complete case analysis is conducted with only studies reporting the covariate included and covariates are reported for the study as a whole and not split by intervention group.

The covariates were centred on the mean covariate value due to poor mixing in the trace plots and high autocorrelation in the estimates in WinBUGS. With sufficient iterations the model is likely to converge but centring moves the intercept to the mean of the covariate and reduces the correlation between the intercept and slope parameters in the meta-regression model so the model is more efficient (Welton 2012). The results using WinBUGS are given in Table 5-4 and Figure 5-8 displays bubble plots produced when Stata is used to repeat the meta-regression; the size of the bubbles is proportional to the size of the study and the value of beta in Table 5-4 is an estimate of the slope of the line on the bubble plot (Figure 5-8). These models are compared to a meta-analysis on the same data without the covariate using the DIC (Table 5-4) and there is very little difference in the fit of the models for any of the covariates. There is statistically significant evidence that the intervention effect varied in terms of the child gender but not for the other covariates. The effectiveness ($log\ OR$) of an enhanced intervention compared to usual care/control decreases with an increasing percentage of boys in the study (beta=-0.120 (95%CI: -0.212, -0.028)).

Table 5-4 Pairwise meta-regression using aggregate data comparing interventions to increase possession of a fitted safety gate in WinBUGS

Covariate (mean centred)	Number of studies	Mean (min, max) covariate value	Beta (95% CrI) Underlying effect of the covariate on the log OR	DIC no covariate	DIC with covariate
Gender - % male	8	51.6% (46, 59)	-0.12 (-0.21, -0.03)	113.31	112.06
% BME ethnic group	10	48.9% (1, 95)	0.002 (-0.010, 0.014)	135.29	136.14
% Single-parent family	10	27.8% (5, 87)	-0.001 (-0.022, 0.020)	132.263	133.06
% Rented accommodation	8	50.3% (20, 83)	0.010 (-0.002, 0.021)	105.40	104.20
% one or more parents unemployed	7	36.4% (11, 77)	-0.007 (-0.037, 0.022)	94.434	94.703
mean/median age	12	1.48 years (0, 4.2)	-0.23 (-0.61, 0.15)	161.32	162.11
% teenage mother*	4	18.3% (15, 24)			

* meta-regression not used due to small number of studies
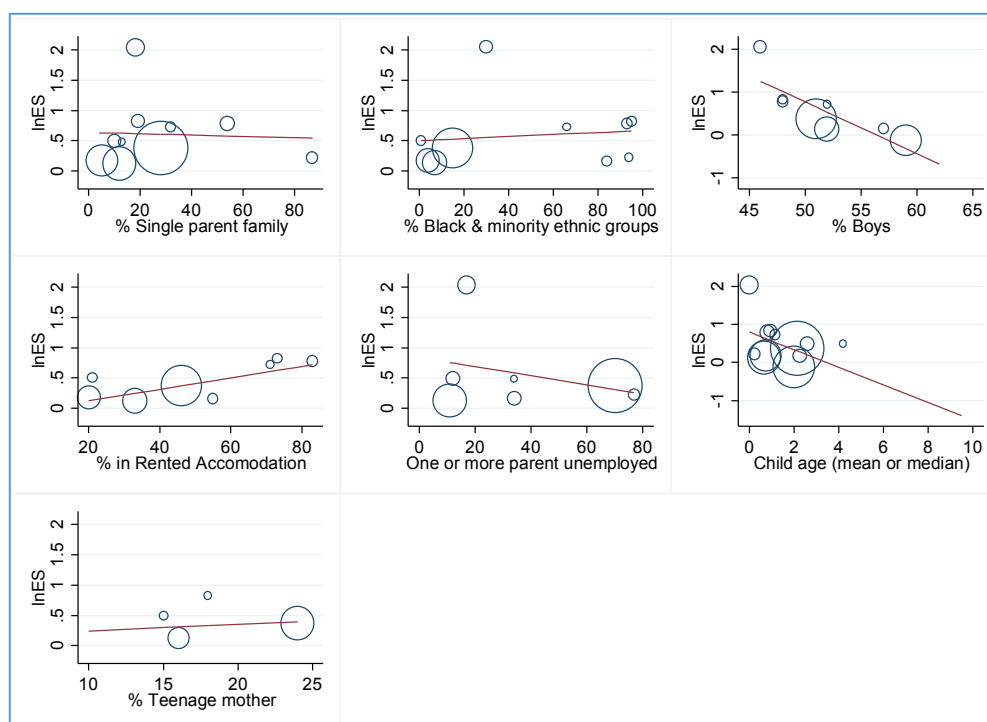
Figure 5-8 Bubble plots for the pairwise meta-regression using aggregate data

5.4 Adjusting for baseline risk

WinBUGS is used to adjust for baseline risk (section 4.7) using the observed usual care/control group log odds of the possession of a fitted safety gate (Table 5-5). There is a lot of variation in the proportion of households possessing a fitted safety gate in

the usual care arm, ranging from 23% to 90%. The coefficient of the baseline risk,
beta, is 0.0082 (95%CrI: -0.433, 0.476) showing that the baseline risk has very little
influence on the log odds ratio and hence differences in the baseline possession of
fitted safety gates does not significantly affect the intervention effect. The DIC for this
model 161.8 is very similar to the pairwise random effects meta-analysis model that
had a DIC of 161.4. The results of this analysis are questionable due to the wide range
of usual care/control interventions used in the studies that could have also affected
the probability of possession and hence not representative of the baseline possession
for households in this group.

Table 5-5 Baseline risk for the possession of a fitted safety gate

Author	Year	Probability of possession of a fitted safety gate in usual care/control group	Usual care/control group log odds
Clamp	1998	0.72	0.97
Kendrick	1999	0.66	0.67
King J	2001	0.35	-0.60
Nansel	2002	0.79	1.30
Gielen	2002	0.23	-1.21
Snazjder	2003	0.90	2.20
Posner	2004	0.53	0.13
McDonald	2004	0.24	-1.13
Watson	2004	0.46	-0.17
Kendrick	2005	0.80	1.37
Nansel	2008	0.76	1.17
Phelan	2010	0.53	0.12

5.5 Meta-analysis using IPD and aggregate data

IPD is available for ten of the 12 studies in the meta-analysis, two of which are cluster
allocated studies (Table 3-1). IPD meta-analysis is seen as the gold standard (section
4.9) and for this application most studies had IPD which is unusual. The IPD required a
lot of data manipulation to get all study data in the same form to use in WinBUGS and
not all covariates were recorded in each study. WinBUGS is used to conduct a pairwise
meta-analysis including the IPD combined with aggregate data and accounting for
cluster allocated studies (section 4.9). Enhanced healthcare interventions are shown to
be effective in increasing the possession of a fitted safety gate ($OR =$
$1.62, 95\%\text{CrI}: 1.00, 2.73$) with an estimated between-study variance $\tau^2 =$

0.37 (95%CrI: 0.11, 1.51). The OR is very similar to the aggregate analysis with slightly more uncertainty and heterogeneity (section 5.2).

5.6 Meta-regression using IPD and aggregate data

Meta-regression combining cluster and individually allocated IPD and study level aggregate data (section 4.9.3) is used to assess the effect of the covariates considered in section 5.3.2. Ecological bias, study-level and subject-level analyses giving different results, is accounted for in the analysis by splitting the between and within study variability (section 4.9.3). The numbers of studies reporting covariates and providing IPD are given in Table 5-6.

Table 5-6 Studies providing IPD and the covariates recorded for the possession of a fitted safety gate outcome

Author	IPD	Cluster study	Age	%Male	%Rented	%Single parent	%BME**	%Un-employed
Clamp(1998)	✓		✓		✓	✓	✓	✓
Kendrick(1999)	✓	✓	✓	✓	✓	✓	✓	✓
King J(2001)			✓	✓				
Nansel(2002)	✓		✓	✓	✓	✓	✓	
Gielen(2002)		✓	✓			✓	✓	✓
Snazjder(2003)	✓		✓			✓		✓
Posner(2004)	✓		✓	✓	✓		✓	✓
McDonald(2004)	✓		✓	✓	✓	✓	✓	
Watson(2004)	✓		✓	✓	✓	✓	✓	✓
Kendrick(2005)	✓	✓	✓		✓	✓	✓	
Nansel(2008)	✓		✓	✓*	✓*	✓	✓	
Phelan(2010)	✓		✓§	✓		✓	✓	✓

* Nansel(2008) IPD did not include rented and gender so they were included as aggregate covariates
** BME – Black or minority ethnicity, §all households recruited when baby was born

Due to the smaller numbers of studies for some covariates some rules were determined with other members of the KCS programme team on when to use random effects and when to split the variance to account for ecological bias. The rules applied, in order of complexity, are:

- Five or more studies with at least one IPD: random effects model for simultaneously analysing IPD and summary-level data, investigating splitting variance between and within studies.

- Four studies with at least two IPD or five or more aggregate only studies: random effects model for simultaneously analysing IPD and summary-level data.

- One IPD and two to four aggregate; two IPD and one aggregate; three IPD and no aggregate only studies: fixed effects model for simultaneously analysing IPD and summary-level data.

- Aggregate data only, five or more studies: meta-regression

Appendix C summarises the number of studies and type of data, aggregate/IPD/cluster allocated, for each outcome/covariate combination.

To improve convergence speed covariates are centred on the overall mean covariate across studies. The age of the child covariate was evaluated at ages 0 and 4 to be consistent with the other injury prevention outcomes included in the Cochrane Review (Kendrick, Young et al. 2012).

Table 5-7 presents the results for the meta-regression analyses using IPD and aggregate covariate values. There is little difference between the between- and within-study results for all covariates, hence little evidence of ecological bias, so the results without splitting have also been included. There is little evidence that the enhanced intervention varied in effect with child age, ethnic group or parental unemployment. An enhanced intervention is significantly more effective in increasing possession of a fitted safety gate amongst households in rented accommodation (OR=1.98(95%CrI:1.48 to 2.66)) and compared to households in the owner-occupied group (interaction OR=1.62(95%CrI:1.18 to 2.24)) however the results should be treated with caution due to the multiple tests conducted.

These results were presented in the update of the Cochrane Review (Kendrick, Young et al. 2012) along with the other outcomes reported in section 5.7.

Table 5-7 Meta-regression ORs (95% credible interval) for possession of a fitted safety gate outcome by child age, gender, and social group

Outcome: possession of a fitted safety gate				
Between & within study variance combined		**Between & within study variance split**		
Gender		*Gender*		
Boys (exp(d+β))	1.64(0.85 to 3.31)	*Within*		
Girls (exp(d))	1.92(0.99 to 3.85)	Boys (exp(d*+βw))	1.04(0.47 to 2.40)	
Interaction term (exp(β))	0.86(0.62 to 1.18)	Girls (exp(d*))	1.22(0.58 to 2.66)	
		Interaction term (exp(βw))	0.86(0.62 to 1.18)	
Cluster IPD: 1§		*Between*		
Non-Cluster IPD: 5		Boys (exp(d*+βb))	0.42(0.04 to 4.25)	
Cluster AD: 0		Girls (exp(d*))	1.22(0.58 to 2.66)	
Non-Cluster AD: 2		Interaction term (exp(βw))	0.34(0.03 to 4.40)	
		Difference (β.diff)	0.92(-1.64 to 3.44)	
Ethnic group		*Ethnic group*		
Black & minority ethnic groups	1.98(1.17 to 3.34)	*Within*		
White	1.65(1.01 to 2.76)	Black & minority ethnic groups	2.04(0.86 to 5.07)	
Interaction term	1.19(0.77 to 1.85)	White	1.70(0.83 to 3.64)	
		Interaction term	1.20(0.73 to 1.96)	
Cluster IPD: 2§		*Between*		
Non-Cluster IPD: 7		Black & minority ethnic groups	1.91(0.79 to 4.33)	
Cluster AD: 1		White	1.70(0.83 to 3.64)	
Non-Cluster AD: 0		Interaction term	1.12(0.32 to 3.68)	
		Difference	0.07(-1.22 to 1.45)	
Family type		*Family type*		
Single-parent family	2.03(1.16 to 3.62)	*Within*		
Two-parent family	1.82(1.12 to 3.02)	Single-parent family	2.25(1.00 to 3.17)	
Interaction term	1.11(0.75 to 1.65)	Two-parent family	1.99(0.98 to 4.17)	
		Interaction term	0.78(0.10 to 5.68)	
Cluster IPD: 2§		*Between*		
Non-Cluster IPD: 7		Single-parent family	1.57(0.32 to 7.55)	
Cluster AD: 1		Two-parent family	1.99(0.98 to 4.17)	
Non-Cluster AD: 0		Interaction term	1.13(0.76 to 1.69)	
		Difference	0.36(-1.68 to 2.46)	
Housing tenure		*Housing tenure*		
Non-owner occupied	1.98(1.48 to 2.66)	*Within*		
Owner occupied	1.22(0.96 to 1.61)	Non-owner occupied	1.73(0.98 to 3.07)	
Interaction term	1.62(1.18 to 2.24)	Owner occupied	1.10(0.71 to 1.74)	
		Interaction term	1.58(1.13 to 2.22)	
Cluster IPD: 2§		*Between*		
Non-Cluster IPD: 5		Non-owner occupied	2.36(1.19 to 4.67)	
Cluster AD: 0		Owner occupied	1.10(0.71 to 1.74)	
Non-Cluster AD: 1		Interaction term	2.13(0.75 to 6.04)	
		Difference	-0.31(-1.40 to 0.78)	

Table 5-7 (continued) Meta-regression ORs (95% credible interval) for possession of a fitted safety gate outcome by child age, gender, and social group

Outcome: possession of a fitted safety gate			
Between & within study variance combined		**Between & within study variance split**	
Unemployed		*Unemployed*	
1 or more parents unemployed	2.08(0.77 to 5.86)	*Within*	
Parents employed	1.82(0.67 to 5.01)	1 or more parents unemployed	2.94(0.36 to 15.49)
Interaction term	1.15(0.77 to 1.71)	Parents employed	2.53(0.32 to 12.41)
		Interaction term	1.17(0.78 to 1.75)
Cluster IPD: 1§		*Between*	
Non-Cluster IPD: 4		1 or more parents unemployed	1.30(0.16 to 20.13)
		Parents employed	2.53(0.32 to 12.41)
Cluster AD: 1		Interaction term	0.49(0.03 to 35.38)
Non-Cluster AD: 0		Difference	0.85(-3.43 to 3.82)
Age		*Age*	
OR at age 0	1.40(1.02 to 2.06)	*Within*	
OR at age 4	1.26(0.81 to 2.02)	OR at age 0	2.36(1.16 to 5.02)
Interaction term	0.97(0.84 to 1.13)	OR at age 4	1.91(0.72 to 5.30)
		Interaction term	0.95(0.79 to 1.13)
Cluster IPD: 1§		*Between*	
Non-Cluster IPD: 7		OR at age 0	2.36(1.16 to 5.02)
Cluster AD: 2		OR at age 4	0.85(0.20 to 3.40)
Non-Cluster AD: 2		Interaction term	0.78(0.47 to 1.24)
		Difference	0.20(-0.29 to 0.72)

d overall intervention effect, d* is the uncentred effect estimate, β covariate effect, β_w is the within study association, β_b is the between study association, β.diff is the difference $\beta_w - \beta_b$ which represents the ecological bias.

§ no. of studies included in the analysis

5.7　Pairwise meta-analysis for the other outcomes

Sections 5.2 - 5.6 described the pairwise meta-analyses for the studies reporting interventions to increase the possession of a fitted safety gate. These analyses are also repeated for the studies reporting interventions to increase the possession of safety equipment to prevent falls in children or increase fall prevention behaviour: no possession or use of a baby walker, possession of window locks, possession of a bath mat or decals and never leaving a child unattended on a high surface.

Studies reporting data on these outcomes are summarised in Table 5-1 and again the types of intervention were very heterogeneous so to conduct a pairwise meta-analysis they were lumped together into two intervention groups, usual care vs enhanced

intervention. The Babul (2007) study, reporting interventions to reduce the use of a baby walker and leaving children unattended on high surfaces, is a three-arm study (Babul, Olsen et al. 2007). For the pairwise meta-analysis the two enhanced intervention groups were combined. The results of the pairwise meta-analysis undertaken in Stata are summarised in Table 5-8 and the forest plots are presented in the Appendix E using Stata. IPD was available for some of the studies (Table 3-1) and this was included in meta-regression analyses where possible; there was a limited number of studies for some covariates and outcomes (Appendix C).

Table 5-8 Pairwise meta-analysis results for other falls prevention outcomes (using Stata)

	No. of studies	Overall OR (95%CI) Random effects model	Heterogeneity I^2 (p-value) τ^2
No possession or use of baby walker	9	1.57 (1.18, 2.09)	51.3% (0.04) 0.088
Possession of window locks	6	1.17 (0.87, 1.57)	51.8% (0.07) 0.055
Possession of bath mats or decals	4	1.10 (0.68, 1.78)	47.6% (0.12) 0.114
Never leaving a child unattended on a high surface	3	1.20 (0.84, 1.72)	<0.1% (0.773) <0.001

Households in the enhanced intervention group are less likely to possess or use a walker than control group households (OR 1.57, 95% CI: 1.18, 2.09) (OR represents the ratio of no possession or use of baby walkers in the enhanced intervention group compared to the usual care group). There is no significant evidence of publication bias. There is statistically significant heterogeneity between the study effect sizes (p=0.04) but, using meta-regression on the aggregate mean covariate values, there are no statistically significant effects by gender, ethnic group, family type, unemployed or housing tenure. Age of child was not investigated as baby walkers are not used by older children. When the IPD are included in the meta-regression, Appendix F, there are no statistically significant covariate effects for the possession or use of baby walker outcome.

Households in the home safety interventions arms were not significantly more likely to possess window locks than usual care group households (OR 1.17, 95% CI: 0.87, 1.57) and the effect sizes did not vary significantly between studies. Intervention arm

households were not significantly more likely to have non-slip bath mats or decals (OR 1.10, 95% CI: 0.68, 1.78) or to never have left a child unattended on a high surface (OR 0.84, 95% CI 0.58 to 1.20) and there was no significant heterogeneity between study effect sizes. Meta-regression analyses are conducted using aggregate covariate values and also using including the available IPD; for some outcome covariate combinations only a fixed effects model can be fitted due to the small numbers of studies and splitting the between and within study variance is only possible for the baby walker and window lock outcomes. The results are given in Appendix F and Appendix G and show that the enhanced intervention is more effective in increasing possession of window locks in households with a male child compared to households with a female child (OR=1.72(95%CrI: 1.16, 2.57) this is because for boys the enhanced intervention is slightly (but not significantly) more effective (OR=1.45 (95%CrI: 0.80, 2.92) but for girls the enhanced intervention is slightly (but not significantly) less effective (OR=0.85 (95%CrI: 0.46, 2.57) than usual care. However, the results should be treated with caution due to the multiple tests conducted.

5.8   Summary of pairwise meta-analyses

Sections 5.2-5.7 describe the results of pairwise meta-analyses to compare the effectiveness of two intervention groups, usual care vs an enhanced intervention group, using the evidence identified in two or more studies to increase the possession of safety equipment to prevent falls or increase the use of safety behaviour practice. Random effects models were fitted using WinBUGS with vague prior distributions to allow the data to dominate and showed that households in the enhanced intervention arm were more likely to possess a fitted safety gate and less likely to have or use a baby walker, than families in the usual care arm. Households in the enhanced intervention group were not statistically significantly more likely to possess of bath mats or decals, window locks or never leave a child on a high surface. Sub-group analysis and meta-regression, including available IPD, were used to explore heterogeneity arising from clinical and methodological differences in the studies. An increasing proportion of male children in a study reduces the effectiveness of the enhanced intervention to increase possession of a fitted safety gate when only

aggregate data was used but is not statistically significant when IPD is included. An enhanced intervention is significantly more effective in increasing possession of a fitted safety gate amongst households in rented accommodation compared to households in the owner-occupied group.

The enhanced intervention for one study, Phelan (2010), is much more effective than the enhanced interventions in the other studies. This study has the most components, including a home safety inspection, provision and fitting of free safety equipment (including safety gates and window guards) and a safety advice handout.

To perform the pairwise meta-analyses different types of interventions are lumped together to give the two intervention groups, usual care and enhanced intervention (Figure 5-1). For example the enhanced intervention for the Phelan study consists of several components and is treated the same as the enhanced intervention for Kendrick (2005) where households received home safety advice from a midwife and health visitor (Table 3-1). This type of evidence lumping will not provide the relevant information for decision makers who want to know which specific intervention work the best. A network meta-analysis (NMA) allows multiple interventions to be compared to each other with direct evidence, observed in the studies, and indirect evidence, where no study evidence available and is described in section 4.11.

5.9   Network meta-analysis (NMA) to compare the effectiveness of interventions to increase the possession of fitted safety gates

A network meta-analysis is used to answer the question "Which intervention(s) are most effective in increasing the possession of safety equipment to prevent falls in children or increase fall prevention behaviour", this is in contrast to the pairwise meta-analysis which addresses the question "do enhanced interventions offered in a healthcare setting increase the possession of safety equipment to prevent falls in children or increase fall prevention behaviour compared with a usual care intervention". Where there are multiple interventions identified in a systematic review, the NMA answers the more relevant question for decision makers. Firstly the

NMA uses aggregate data (section 4.11) and then covariates (section 4.12) and IPD (section 4.13) separately and together (section 4.14) are included.

5.9.1 Possession of fitted safety gates NMA

The twelve studies used in the pairwise meta-analysis (Table 5-1) are included in the network meta-analysis to investigate which intervention(s) are most effective in increasing the possession of fitted safety gates to prevent falls. Ten (83%) are RCTS and two (17%) are non-RCTs. The two interventions compared in the pairwise meta-analysis (usual-care vs enhanced intervention) are extended to seven interventions that include usual care along with combinations of education, low cost/free equipment, home safety inspection and fitting.

A recent publication (James, Yavchitz et al. 2018) reported that there was a lack of a consensual methods to support the node-making process which could lead to different choices of nodes and different NMA results. They identified two methods that authors of NMA use to support the process which were: use a previous published classification or rely on expert consensus and recommended reporting of NMAs could be improved with a transparent and reproducible node-making process. The seven interventions given below were identified by the authors of the systematic review (Young, Wynn et al. 2013). The authors included experts in the area of accident prevention who were able to identify appropriate combinations of the elements aiming to increase the use of fitted safety gates. The usual care intervention was often not clearly detailed in the original studies and was assumed to have included generic/standard leaflets or advice but no home safety visits or equipment. Education represents enhanced education usually tailored to the family and often face-to-face. The equipment is not always relevant to the possession of a fitted safety gate outcome but is related to home safety, for example provision of cupboard locks or smoke alarms, and this will be discussed further in section 5.9.3.

The seven interventions considered are:

1. Usual care
2. Education

3. Education + low cost (i.e. voucher) / free equipment

4. Education + low cost equipment (i.e. voucher) / free equipment + home safety
   inspection

5. Education + low cost (i.e. voucher) / free equipment + fitting

6. Education + home safety inspection

7. Education + low cost (i.e. voucher) / free equipment + fitting + home safety
   inspection

A description of the studies and interventions is presented in Table 5-9. The adjusted
effective sample sizes are used for any cluster allocated studies. Figure 5-9 (Stata) and
Figure 5-10 display network diagrams for this NMA show different ways of presenting
the network of evidence. The diagrams show a connected network. When this NMA
was first conducted the Stata code had not been released to draw the networks or
conduct an NMA. The NMA model for this network is fitted using a MCMC approach
with minimally informative prior distributions in WinBUGS to obtain pooled estimates
of 21 possible pairwise intervention comparison effects. The effect estimates are
expressed as odds ratios (comparing the higher numbered intervention to the lower
numbered intervention (section 5.9)) with 95% credible intervals (CrI) using both a
combination of direct and indirect evidence, and indirect evidence only. The WinBUGS
code is given in Appendix H.
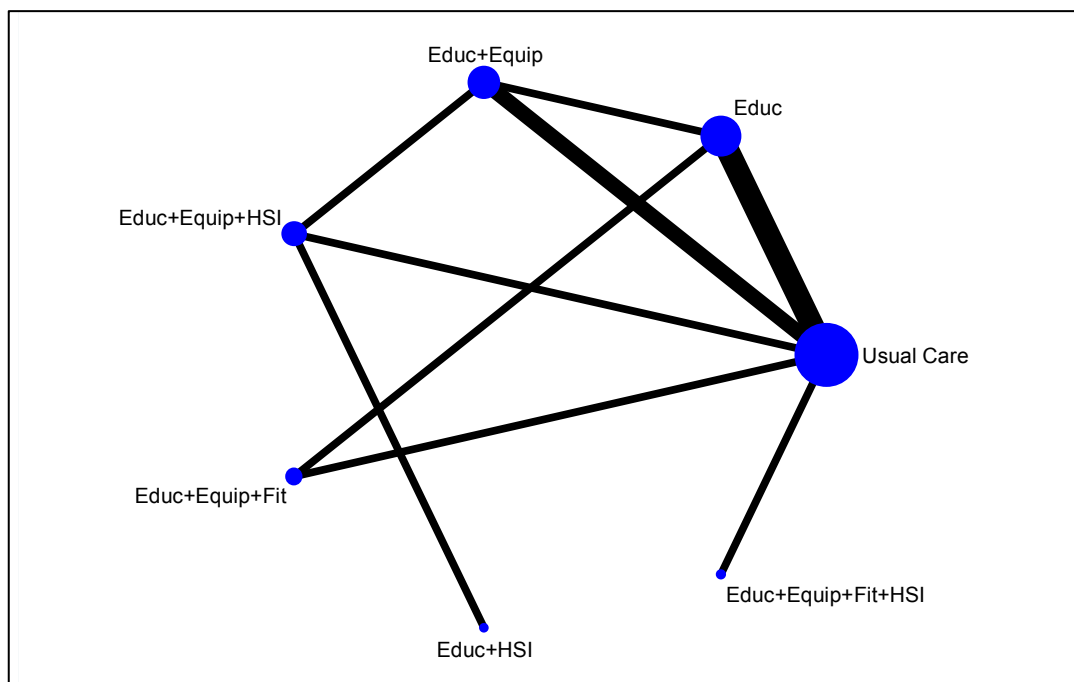
Each intervention is a node in the network. The links between the nodes are pairwise intervention comparisons. The thickness of lines represents the evidence.
Educ – Education, Equip – free/low cost equipment, HIS –home safety inspection, Fit - fitting

Figure 5-9 Network diagram (Stata) of interventions to increase the possession of a fitted safety gate.
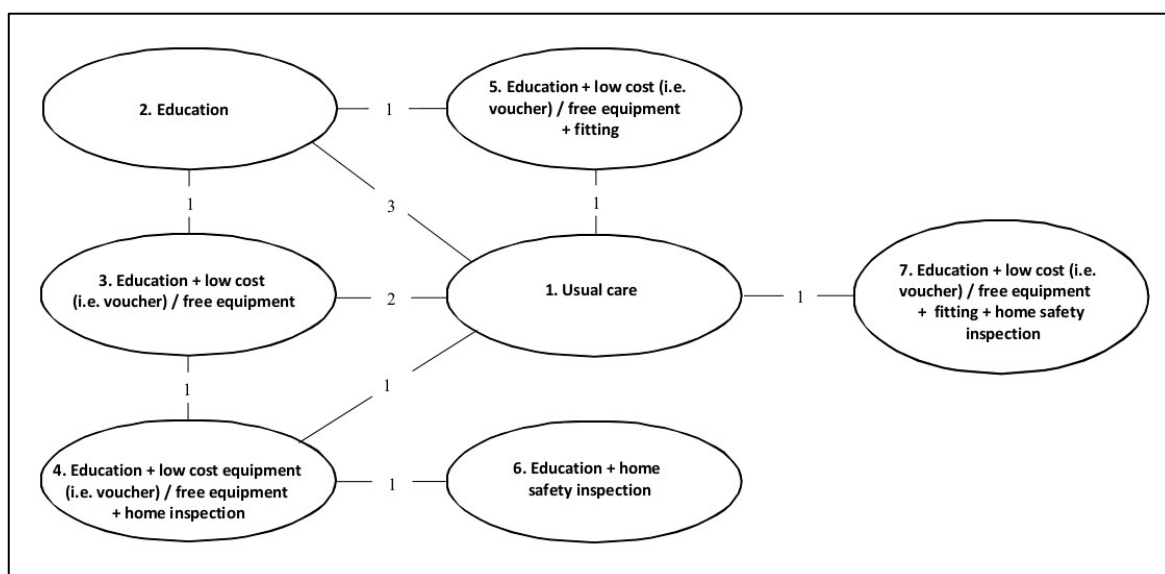
Each intervention is a node in the network. The links between the nodes are pairwise intervention comparisons. The number on the lines represents the number of studies reporting evidence.

Figure 5-10 Network diagram of interventions to increase the possession of a fitted safety gate.

The NMA results are presented in Table 5-10 and in a caterpillar plot in Figure 5-11. For completeness, pooled estimates from the direct evidence only are presented for each pairwise comparison where study data is available (using a fixed effect meta-analysis model when only two studies are available for a particular pairwise comparison, and a random effects model where three or more studies are available, and where only one study had evaluated a particular pairwise comparison the results from this study alone). From the NMA results, intervention effectiveness is ranked based on absolute intervention effects (derived by using an underlying rate based on the usual-care arms) and the probability that each intervention is best for a particular outcome is calculated and the results are presented in Table 5-11.

Table 5-11 shows the most intensive intervention (7=education + low cost/free equipment + home safety inspection + fitting) is most likely to be effective (median rank=1 (95%CrI: 1,2), probability best = 0.97), in increasing possession, with, for example, families in the intensive intervention group more likely to possess a fitted safety gate compared to those in the usual care group (OR=7.80 (95% CrI: 3.18, 21.3)). The odds ratios comparing intervention 7 to all other interventions show clearly that this intervention is most effective (Figure 5-11) but the credible intervals are wide due to the lack of evidence. There is only one study comparing these two interventions directly and no other studies trial the most intensive intervention so inconsistency cannot be checked but the pairwise estimate for this effect (OR=7.73 (95%CrI: 4.14, 14.4)) is similar to the estimate from the NMA. The usual care intervention had a median rank=7 (95%CrI: 5,7) with a probability best<0.001. The CrIs for the ranking overlap considerably and hence are difficult to interpret and the probability best does not have an estimate of uncertainty so should be treated with caution.

Table 5-9 Summary of Studies and their data included in the NMA of the Possession of a Fitted Safety gate

Comparison (intervention number)	First author,Year (Reference)	Study design	Follow-up (months)	Study quality	Safety gate Number of safety gates *	Total number of households
Usual care (1) vs. Education (2)	Nansel 2002(Nansel, Weaver et al. 2002)[e]	RCT	0.75	A=Y, B=Y, F=Y	70 76	89 85
	Kendrick 2005(Kendrick, Illingworth et al. 2005)	RCT	9	A=Y, B=N, F=Y	418(348.44) 373(310.93)	524(436.80)[a] 452(376.78)
	Nansel 2008(Nansel, Weaver et al. 2008)[f]	Non-RCT	1	B=N, F=N, C=N	29 60	38 69
Usual care (1) vs. Education + Low/free equipment (3)	Clamp 1998(Clamp, Kendrick 1998)	RCT	1.5	A=Y, B=N, F=Y	50 52	69 64
	McDonald 2005(McDonald, Solomon et al. 2005)	RCT	1	A=Y, B=U, F=N	10 23	41 54
Usual care (1) vs. Education + Low/free equipment + Home safety inspection (4)	Kendrick 1999(Kendrick, Marsh et al. 1999)	Non-RCT	25	B=N, F=N, C=Y	241(214.26) 251(223.15)	364(323.61)[b] 364(323.61)
Usual care (1) vs. Education + Low/free equipment + Fitting (5)	Watson 2005(Watson, Kendrick et al. 2005)	RCT	12	A=Y, B=N, F=N	328 408	718 742
Usual care (1) vs. Education + Low/free equipment + Fitting + Home safety inspection (7)	Phelan 2010(Phelan, Khoury et al. 2010)	RCT	12	A=Y, B=N, F=Y	78 131	147 146
Education (2) vs. Education + Low/free equipment (3)	Posner 2004(Posner, Hawkins et al. 2004)	RCT	2.5	A=Y, B=Y, F=N	25 28	47 49
Education (2) vs. Education + Low/free equipment + Fitting (5)	Sznajder 2003(Sznajder, Janvrin et al. 2003)	RCT	1.5 to 2	A=Y, B=N, F=Y	45 44	50 47
Education + low/free equipment (3) vs. Education + low/free equipment + Home safety inspection (4)	Gielen 2002(Gielen, McDonald et al. 2002)	RCT	12	A=U, B=U, F=N	11(12.85) 13(10.87)	48(47.44) [c] 48(47.44)
Education + Low/free equipment + Home safety inspection (4) vs. Education + Home safety inspection (6)	King 2001(King, Klassen et al. 2001)	RCT	12	A=Y, B=Y, F=Y	158 166	482 469

Abbreviations: NMA, network meta-analysis; RCT, randomized clinical trial; A = adequate allocation concealment; B = blinded outcome assessment; C, prevalence of confounders does not differ by more than 10% between treatment arms (non-RCT only); F = at least 80% participants followed up in each arm; Y= yes; N = no; U = unclear.

[a] ICC calculated from IPD from Kendrick 2005   [b] ICC calculated from IPD from Kendrick 1999   [c] ICC calculated from IPD from Kendrick 1999 and Kendrick 2005

[d] Babul has been included in the NMA as a three-arm trial but is listed above as three separate comparisons. [e] Two intervention arms are combined (tailored advice and tailored advice + care provider feedback) [f] Generic safety advice is counted as usual care   [g] Two control arms are combined (usual care and usual care + baseline questionnaire

* (Numbers adjusting for clustering in parentheses)

An alternative presentation of the results is given in Figure 5-12. The results compare the lower intervention number to the higher intervention number in each comparison, hence the OR in Figure 5-12 are the reciprocal of those given in Table 5-10. The results are slightly different due to the different seed starting value in WinBUGS. Effect estimates are given for both the direct only (pairwise M-A) and combined direct and indirect estimates from the NMA. Prediction intervals and credible intervals are presented as error bars. SUCRA percentages (section 4.11.2) are presented for each intervention with the most effective intervention being the most intensive (intervention 7) with a SUCRA of 99%. Usual care has a SUCRA of 9%. Another alternative plot produced by the Tan et al software is given in Appendix I.
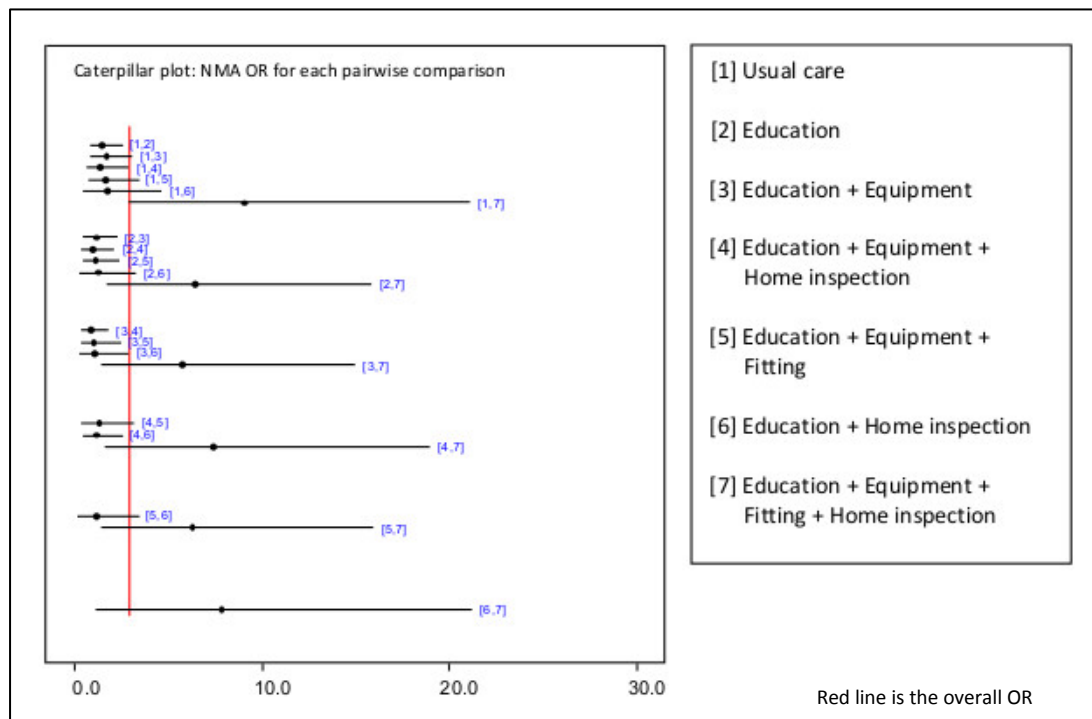
Figure 5-11 Caterpillar plot showing the OR for the pairwise comparisons using direct (where available) and indirect evidence

Table 5-10 Results for the NMA of interventions to increase possession of a fitted safety gate expressed as Odds Ratios [a,b]

Safety gates NMA

Interventions	Usual care (1)	Education (2)	Education + Equipment (3)	Education + Equipment + Home inspection (4)	Education + Equipment + Fitting (5)	Education + Home inspection (6)	Education + Equipment + Fitting + Home inspection (7)
Usual care (1)		1.43 (0.90 , 2.49)	1.63 (0.93 , 3.03)	1.28 (0.69 , 2.79)	1.52 (0.84 , 3.38)	1.43 (0.56 , 4.42)	7.80[c] (3.08 , 21.3)
Education (2)	1.48 (0.97 , 2.25)		1.14 (0.56 , 2.23)	0.90 (0.41 , 2.07)	1.07 (0.51 , 2.41)	1.01 (0.33 , 3.25)	5.46[c] (1.75 , 16.1)
Education + Equipment (3)	1.92[c] (1.05 , 3.51)	1.17 (0.52 , 2.63)		0.78 (0.38 , 1.77)	0.94 (0.42 , 2.41)	0.88 (0.32 , 2.80)	4.77[c] (1.56 , 15.2)
Education + Equipment + Home inspection (4)	1.13 (0.82 , 1.58)		1.25 (0.49 , 3.17)		1.20 (0.45 , 3.25)	1.12 (0.52 , 2.49)	6.13[c] (1.75 , 18.7)
Education + Equipment + Fitting (5)	1.45[c] (1.18 , 1.79)	1.63 (0.37 , 7.23)				0.94 (0.27 , 3.28)	5.07[c] (1.47 , 15.9)
Education + Home inspection (6)				1.12 (0.86 , 1.47)			5.48[c] (1.23 , 20.7)
Education + Equipment + Fitting + Home inspection (7)	7.73[c] (4.14 , 14.4)						

Abbreviations: CrI, credible interval; CI, confidence interval; NMA, network meta-analysis; OR, odds ratio.

[a] values above the diagonal are results from the NMA, OR with 95%CrI; those below the line are direct estimates from a trial or, where more than one are available, a meta-analysis with 95%CI. Blank cells indicate that no direct evidence on specific pairwise comparisons is available.

[b] Column and row headings signify intervention or comparison (intervention number).

[c] CrI does not contain 1

5.9.2 Evaluation of the NMA model

Overall, the NMA model fitted the data well with the posterior mean residual deviance, 22.51, being close to the number of data points in the network, Table 5-11, indicating good model fit to the data.

Table 5-11 Assessment of best intervention and model fit for the NMA of interventions to increase the possession of a fitted safety gate

Intervention	Possession of a safety gate	
	Probability intervention is best	Median intervention rank (95% CrI)
Education + Equipment +Fitting + Home safety inspection (7)	0.97	1(1, 2)
Education + Home safety inspection (6)	0.013	4(2, 7)
Education + Equipment + Fitting (5)	0.008	4(2, 7)
Education + Equipment (3)	0.004	3(2, 7)
Education (2)	0.002	4(2, 7)
Education + Equipment + Home safety inspection (4)	0.001	5(2, 7)
Usual care (1)	<0.001	7(5, 7)
Model fit		
Posterior mean residual deviance	22.51 (cf 24 data points)	
Between-study standard deviation	0.23 (0.015, 0.87)	

The between study variance (τ^2) is estimated to be 0.23 (95%CrI: 0.015, 0.87). The uncertainty reflects the relatively low number of studies providing direct evidence for each pairwise comparison but shows fairly low heterogeneity between studies. This model assumes that the degree of between-study within-comparison heterogeneity is constant across all intervention comparisons in the network.

Consistency is checked between the direct and indirect evidence by using node-splitting methodology in R and WinBUGS (section 4.4.2). This can only be done when a pair of interventions is part of a closed loop in the network. Any closed loops in the networks are checked for consistency between the direct and indirect evidence. For the safety gate outcome there are closed loops for interventions 1 to 5. Interventions 6 and 7 are not part of closed loops (Figure 5-10). There is no statistically significant

evidence of inconsistency in this analysis, the p-values for all the closed loops of evidence are >0.05 (Table 5-12). The posterior distributions for the direct, indirect and combined $log\ OR$ estimates are plotted in Figure 5-13 for the usual care (intervention 1) and education (intervention 2) comparison.

Network Meta-Analysis Summary Forest Plot Matrix for Possession of safety gates

E+Eq+Fitting+HI (SUCRA=99%)	0.21 (0.07 to 0.64) NA	0.20 (0.06 to 0.67) NA	0.19 (0.05 to 0.85) NA	0.19 (0.06 to 0.56) NA	0.17 (0.05 to 0.58) NA	0.13 (0.05 to 0.34) 0.13 (0.07 to 0.24)
	E+Equipment (SUCRA=59%)	0.94 (0.40 to 2.45) NA	0.88 (0.30 to 2.96) NA	0.88 (0.45 to 1.77) 0.85 (0.38 to 1.91)	0.78 (0.37 to 1.82) 1.25 (0.49 to 3.17)	0.62 (0.33 to 1.09) 0.52 (0.11 to 2.36)
		E+Eq+Fitting (SUCRA=53%)	0.95 (0.26 to 3.56) NA	0.93 (0.41 to 2.10) 0.61 (0.14 to 2.72)	0.84 (0.31 to 2.31) NA	0.66 (0.29 to 1.26) 0.69 (0.56 to 0.85)
			E+HI (SUCRA=48%)	0.99 (0.29 to 3.16) NA	0.88 (0.39 to 2.06) 0.89 (0.68 to 1.16)	0.70 (0.20 to 1.93) NA
				Education (SUCRA=47%)	0.90 (0.40 to 2.17) NA	0.70 (0.39 to 1.11) 0.64 (0.19 to 1.73)
					E+Eq+HI (SUCRA=34%)	0.79 (0.34 to 1.47) 0.88 (0.63 to 1.23)
						Usual Care (SUCRA=9%)

Key: Odds Ratio with 95% CrI & 95% PI (log scale). Heterogeneity: between-study variance = 0.05; 95% CrI (0.000 to 0.835)
NMA results in black; Pairwise MA results in grey. 95% CrI and PI presented as error bars. Interventions are displayed sorted by median rank. SUCRA refers to the surface under the cumulative ranking line.

Figure 5-12 NMA of the interventions to increase possession of safety gates, alternative presentation

5.9.3 Sensitivity analysis

The effect of study design on the NMA results is assessed by repeating the above analysis using only data from the 10 RCTs, excluding the non-RCTs. The result is similar with the most intensive intervention identified as being the most likely to be effective (probability best = 0.87), with an estimated odds ratio for possession versus usual care of 7.93 (95% CrI: 2.76, 23.6).

Table 5-12 Checking consistency in NMA for possession of a fitted safety gate (Random Effects Model)
Posterior Means (Mean) and Standard Deviations (Sd) of the Log-Odds Ratios Using the Full Network, Direct and Indirect Evidence on Each Pairwise

Pair-wise contrast	Combined evidence from network meta-analysis		Direct evidence		Indirect evidence		Inconsistency estimate ω_{xy} and p-value*		
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	p-value
Usual care (1) vs. Education (2)	0.3684	0.2572	0.4422	0.3477	0.1618	0.6023	0.2805	0.695	0.65
Usual care (1) vs. Education + low cost/free equipment (3)	0.4960	0.3049	0.6697	0.4344	0.2661	0.5182	0.4036	0.6759	0.49
Usual care (1) vs. Education + low cost /free equipment + Home safety inspection (4)	0.2649	0.3405	0.1263	0.4722	0.8511	0.7399	-0.7249	0.8774	0.33
Usual care (1) vs. Education + low cost /free equipment + fitting (5)	0.4446	0.3327	0.3717	0.47	0.9691	1.002	-0.5974	1.113	0.56
Education (2) vs. Education + low cost /free equipment (3)	0.1276	0.3493	0.1562	0.6338	0.08906	0.5158	0.0671	0.8141	0.92
Education (2) vs. Education + low cost /free equipment + fitting (5)	0.07624	0.3800	-0.03673	0.5449	-0.03673	0.5449	0.5773	1.073	0.56
Education + low/free equipment (3) vs. Education + low cost /free equipment + Home safety inspection (4)	0.8583	0.4113	0.2347	0.6478	-0.5104	0.5763	0.7451	0.8688	0.32

* p-values correspond to 2 x (*probability of direct estimate > indirect estimate*)
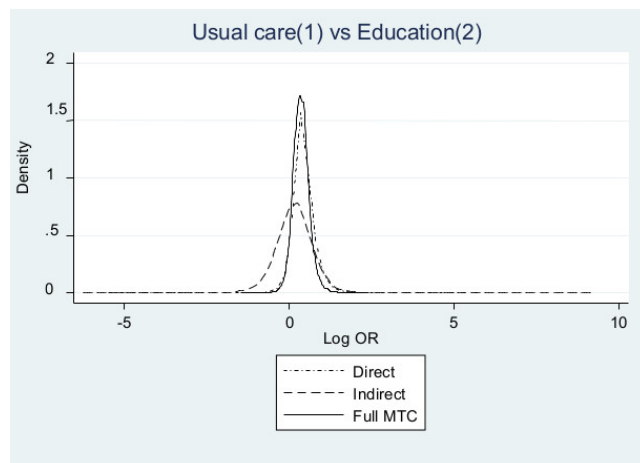
Figure 5-13 Posterior density plot overlaying the direct, indirect and combined estimates of the log odds ratio

The network is extended to comprise 9 different intervention groups by splitting the low cost/free safety equipment giveaway included in interventions into relevant and not relevant/not stated (Appendix K). The findings from this analysis (Appendix K) are similar in that the most intensive intervention clearly is the most effective in increasing the possession of a safety gate and it also showed that there is very little difference between the interventions with low cost/free relevant or not relevant equipment.

5.10 Network meta-analysis (NMA) for the interventions to decrease the possession of a baby walker including a 3-arm study

The 9 studies used in the pairwise meta-analysis (Table 5-1) are included in the network meta-analysis to investigate which intervention(s) are most effective in reducing the possession and use of a baby walker to prevent falls. Nine studies are identified, seven (78%) are RCTs and two (22%) are non-RCTs (Table 5-1). There is no directly relevant equipment in the interventions for this outcome. One RCT is a three-arm study (Babul, Olsen et al. 2007). Figure 5-14 displays a connected network diagram for this NMA with the three arm trial indicated by *.
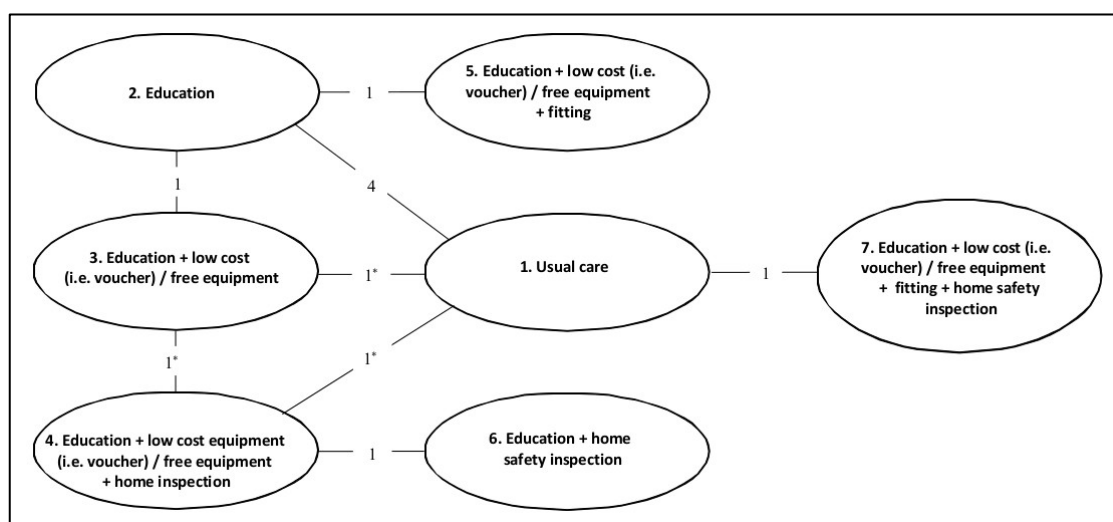
Each intervention is a node in the network. The links between the nodes are pairwise intervention comparisons. The number on the link lines represents the number of studies reporting evidence.
* three-arm study.

Figure 5-14 Network Diagram of Interventions to reduce the possession and use of baby walkers

A standard NMA random-effects model with a binary outcome is fitted that allows studies with 3 or more arms to be included by accounting for the correlation structure (section 4.11.3). The NMA estimated the 21 possible pairwise comparisons between the 7 seven interventions (including usual care). The pooled estimates, along with the available direct within-trial estimates are reported in Table 5-14. Education only is most likely to be effective in reducing possession or use of a baby walker (probability best = 0.65, Table 5-13), with an estimated odds ratio versus usual care of 0.48 (95% CrI: 0.31, 0.84). The only other significant difference is that education only is slightly more effective than education with equipment but this is borderline.

The effect of study design on the results of the NMA results is assessed by repeating the above analysis using only data from the 7 RCTs. The result is similar with the education only intervention identified as being the most likely to be effective (probability best = 0.45), with an estimated odds ratio versus usual care of 0.58 (95% credible interval: 0.21, 1.87) that is no longer significant.

The between-study variance (τ^2) is estimated to be 0.24 (95% credible interval: 0.0094, 1.14), the posterior mean residual deviance is close to the number of data points (i.e. 17.95 compared with 19 data points) indicating good model fit to the data and there is no evidence of inconsistency (Table 5-15).

Table 5-13 Assessment of best intervention and model fit for the NMA of interventions to reduce the possession and use of a baby walker

Intervention	Possession or use of a baby walker	
	Probability intervention is best	Median intervention rank (95% CrI)
Education (2)	0.65	1(1, 4)
Education + Equipment + Fitting (5)	0.13	3(1, 7)
Education + Equipment +Fitting + Home safety inspection (7)	0.13	3(1, 7)
Education + Home safety inspection (6)	0.049	6(1, 7)
Education + Equipment + Home safety inspection (4)	0.027	5(1, 7)
Education + Equipment (3)	0.007	6(2, 7)
Usual care (1)	0.001	4(2, 7)
Model fit		
Posterior mean residual deviance	17.95 (cf 19 data points)	
Between-study standard deviation	0.24 (0.0094, 1.14)	

Table 5-14 Results of a NMA of interventions to reduce possession or use of a baby walker expressed as Odds Ratios (95% CrI)[ab]

	Usual care (1)	Education (2)	Education + Equipment (3)	Education + Equipment + Home inspection (4)	Education + Equipment + Fitting (5)	Education + Home inspection (6)	Education + Equipment + Fitting + Home inspection (7)
Usual care (1)		0.48[c] (0.31 , 0.84)	1.51 (0.56 , 3.65)	1.07 (0.37 , 2.89)	0.85 (0.29 , 3.35)	1.28 (0.29 , 5.06)	0.78 (0.27 , 2.22)
Education (2)	0.46[c] (0.36 , 0.58)		3.15[c] (1.02 , 8.38)	2.25 (0.66 , 6.24)	1.80 (0.53 , 6.24)	2.68 (0.55 , 10.93)	1.63 (0.47 , 4.96)
Education + Equipment (3)	0.59 (0.33 , 1.08)	1.33 (0.17 , 10.3)		0.70 (0.26 , 2.02)	0.56 (0.13 , 2.97)	0.84 (0.21 , 3.52)	0.52 (0.13 , 2.22)
Education + Equipment + Home inspection (4)	0.67 (0.38 , 1.19)		0.89 (0.48 , 1.64)		0.79 (0.16 , 4.71)	1.20 (0.43 , 3.26)	0.73 (0.17 , 3.24)
Education + Equipment + Fitting (5)		1.75 (0.75 , 4.08)				1.51 (0.20 , 9.30)	0.91 (0.16 , 4.47)
Education + Home inspection (6)				0.85 (0.51 , 1.42)			0.61 (0.10 , 3.69)
Education + Equipment + Fitting + Home inspection (7)	0.78 (0.43 , 1.42)						

Abbreviations: CrI, credible interval; CI, confidence interval; NMA, network meta-analysis;

[a] values above the diagonal are results from the NMA, OR with 95%CrI; those below the line are direct estimates from a trial or, where more than one are available, a meta-analysis with 95%CI. Blank cells indicate that no direct evidence on specific pairwise comparisons is available.

[b] Column and row headings signify intervention or comparison (intervention number).

[c] CrI does not contain 1

Table 5-15 Possession or use of a baby walker (Random Effects NMA)

Pair-wise contrast	Combined evidence from network meta-analysis		Direct evidence		Indirect evidence		Inconsistency estimate ω_{xy} and p-value*		
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	p-value
Usual care (1) vs. Education (2)	-0.729	0.253	0.756	0.357	-0.159	1.47	0.915	1.52	0.47
Usual care (1) vs. Education + low/free equipment (3)	0.405	0.479	-0.517	0.803	0.423	1.40	-0.940	1.62	0.47
Education (2) vs. Education + low/free equipment (3)	1.13	0.519	0.945	1.56	-1.28	0.837	-0.334	1.31	0.46

Posterior Means (Mean) and Standard Deviations (Sd) of the Log-Odds Ratios Using the Full Network, Direct and Indirect Evidence on Each Pairwise Comparison

* p-values correspond to 2 x (*probability of direct estimate > indirect estimate*)

5.11 NMA for other injury prevention outcomes

5.11.1 Possession of window locks

Five (83%) of the 6 studies for the possession of window locks outcome are RCTs and 1 (17%) is a non-RCT (Table 5-1). Figure 5-15 displays a network diagram for this NMA showing a connected network.
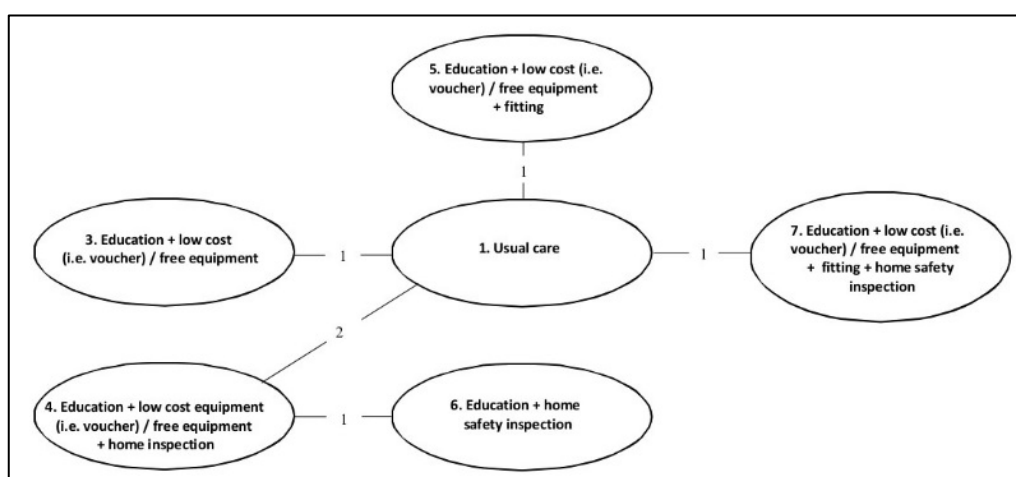
Figure 5-15 Network diagram of interventions to increase possession of window locks

The NMA estimated the 15 possible pairwise comparisons between 6 interventions (including usual care and excluding the education only intervention). The pooled estimates, along with the available direct within-trial estimates are reported in Appendix J. Education + low cost / free equipment + fitting is most likely to be effective in families using window locks (probability best = 0.26), but there is very little difference between any of the interventions. The effect of study design on the results on the NMA results is assessed by repeating the above analysis using only data from the 5 RCTs. The result is similar with very little difference between the interventions.

5.11.2 Child not left on a high surface

Only three studies reported the numbers who never left child on high surface (Table 5-1). Two (67%) of the 3 studies are RCTs and 1 is a non-RCT. One RCT is a three-arm study (Babul, Olsen et al. 2007). Figure 5-16 displays a network diagram for this NMA showing a connected network. There is no relevant equipment for this outcome.

The NMA estimated the 6 possible pairwise comparisons between 4 interventions (including usual care). The pooled estimates, along with the available direct within-trial estimates are reported in Appendix J. There is very little difference between any of the interventions but education only is the least likely to be effective in preventing children being left on high surfaces.
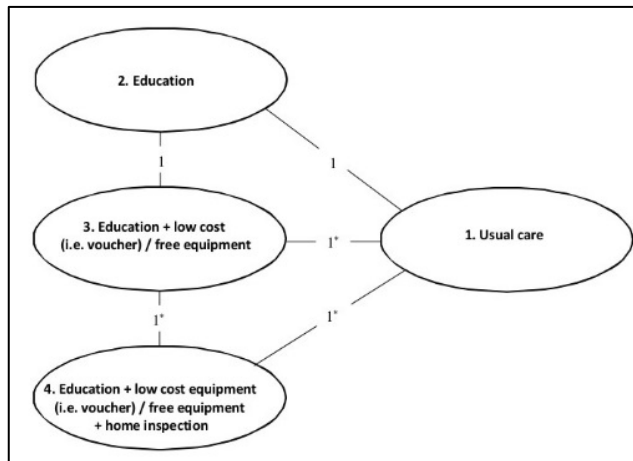
Figure 5-16 Network diagram of interventions to reduce leaving child on high surface

5.11.3 Possession of bath mats or decals

Four studies reported the possession of bath mats (Table 5-1). Three (75%) of the studies are RCTs and 1 (25%) is a CBA study. Figure 5-17 displays a network diagram for this NMA showing two unconnected networks of 3 interventions so NMA cannot be used for this outcome.
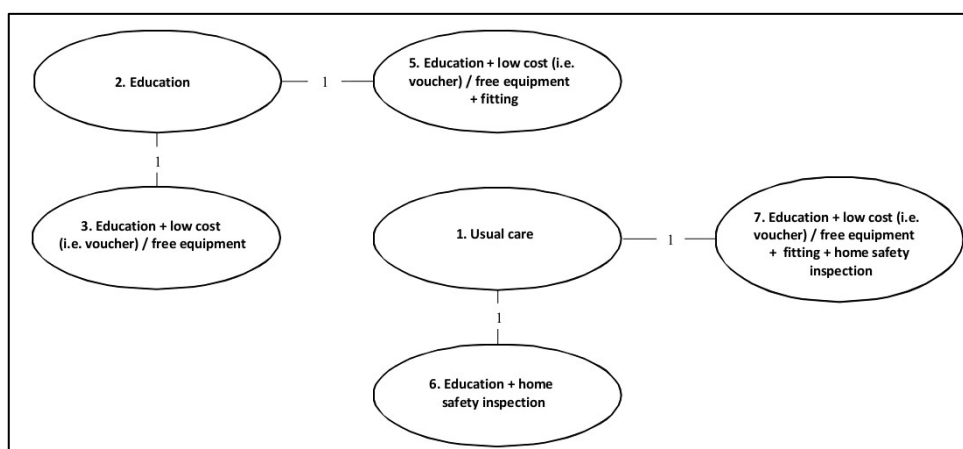
Figure 5-17 Network of evidence for the bath mats/decals outcome

5.12 Network meta-analysis including a covariate – interventions to increase possession of a fitted safety gate

As discussed in section 5.8, the studies included in the meta-analyses and network meta-analyses display some heterogeneity and hence the NMA for the possession of a fitted safety gate will be extended to include the covariates, where there is sufficient evidence. The new networks for the interventions to increase the possession of a fitted safety gate are presented for each covariate in Figure 5-18. Not all of the studies report all of the covariates, section 5.3, so the networks are reduced and not all of the seven interventions are included (Figure 5-18).

Table 5-16 gives the results from the network meta-analyses when covariates were added to the model. The three models described in section 4.12 are fitted using MCMC simulation in WinBUGS (code given in Appendix L). Model 1 (independent regression coefficients) requires too many parameter estimates for the number of studies providing data and failed to converge for all of the covariates. There were problems with autocorrelation and convergence for model 2 (exchangeable regression coefficients) for the covariates representing housing tenure and unemployment shown in Table 5-16, even after increasing the number of iterations, so the results should be treated with caution. The results labelled "no covariate" were obtained by fitting the NMA model to the same studies used in the covariate models, i.e. studies are excluded that do not report the covariate.

Models 2 and 3 give very similar DIC values to the model with no covariate, and none of the covariate effects are statistically significant. There appears to be little advantage to including the covariate in the NMA model. The assumption of a common regression coefficient (model 3) makes a very strong assumption about the regression coefficient and the uncertainty around the variance is wider than the no covariate model.
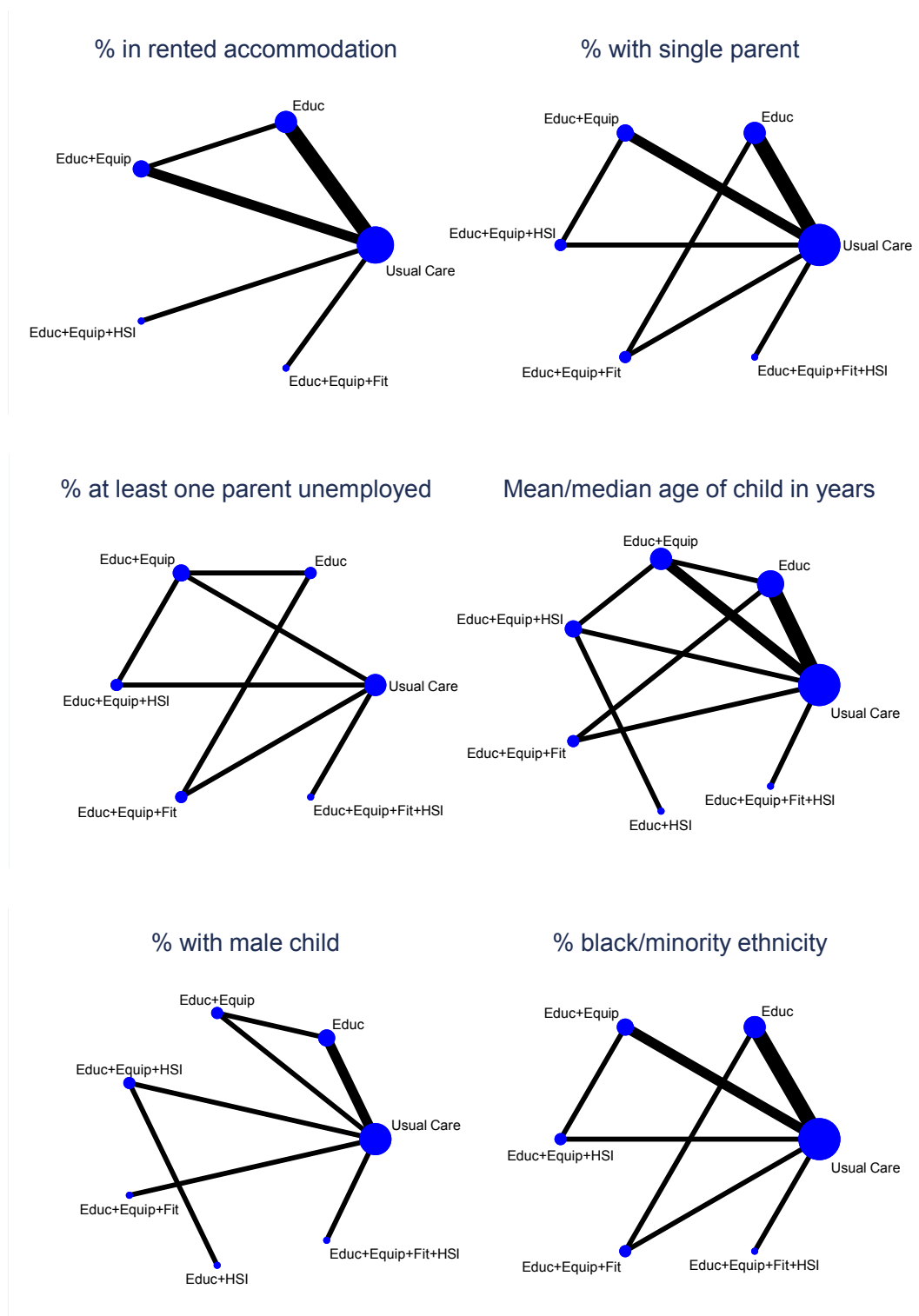
Figure 5-18 Network diagrams for the possession of a fitted safety gate by covariates

Table 5-16 NMA for the possession of safety gates adjusting for covariates

	Residual deviance $\bar{D}$	pD	DIC	Beta (95%CrI)[#] σ_B(95%CrI)	Between study sd τ (95% CrI)
Gender (percentage male) 8 studies, 7 interventions, mean=51.6%					
No covariates	14.8	14.2	112.0		0.37 (0.021 to 1.71)
Model 2: Exchangeable regression coefficients	15.5	16.0	114.5	$\beta_2 = -0.024$ (-0. 90 to 0.89) $\beta_3 = -0.010$ (-0.84 to 0.85) $\beta_4 = -0.027$ (-2.38 to 1.99) $\beta_5 = -0.021$ (-2.38 to 2.22) $\beta_6 = -0.012$ (-2.46 to 2.12) $\beta_7 = -0.005$ (-2.54 to 2.25) $\sigma_B = 0.35$ (0.013 to 1.83)	0.87 (0.046 to 1.94)
Model 3: Common regression coefficient	15.1	15.7	113.9	β =-0.029 (-1.47 to 2.73)	1.07 (0.030 to 4.64)
Ethnic group (percentage black or minority ethnicity) 10 studies, 6 interventions, mean=48.9%					
No covariates	18.6	13.8	131.9		0.26 (0.015 to 1.13)
Model 2: Exchangeable regression coefficients	18.5	18.5	133.2	$\beta_2 = 0.018$ (-0.038 to 0.071) $\beta_3 = 0.011$ (-0.035 to 0.054) $\beta_4 = 0.016$ (-0.028 to 0.059) $\beta_5 = 0.010$ (-0.090 to 0.078) $\beta_6 = 0.013$ (-0.12 to 0.14) $\sigma_B = 0.016$ (0.0007 to 1.55)	0.33 (0.012 to 1.61)
Model 3: Common regression coefficient	18.3	17.7	132.2	β =0.014 (-0.020 to 0.048)	0.28 (0.020 to 1.46)
Family type (percentage single parent) 10 studies, 6 interventions, mean=27.8%					
No covariates	18.6	17.2	131.9		0.26 (0.015 to 1.13)
Model 2: Exchangeable regression coefficients	18.5	18.5	133.2	$\beta_2 = 0.018$ (-0.036 to 0.067) $\beta_3 = 0.011$ (-0.033 to 0.051) $\beta_4 = 0.016$ (-0.027 to 0.056) $\beta_5 = 0.009$ (-0.091 to 0.070) $\beta_6 = 0.014$ (-0.10 to 0.13) $\sigma_B = 0.016$ (0.0007 to 0.22)	0.31 (0.011 to 1.55)
Model 3: Common regression coefficient	18.4	17.6	132.2	β =0.014 (-0.019 to 0.046)	0.25 (0.009 to 1.38)
Housing tenure (percentage in rented accommodation) 8 studies, 5 interventions, mean=50.25%					
No covariates	14.0	13.7	108.4		0.28 (0.017 to 1.22)
Model 2*: Exchangeable regression coefficients	14.1	14.7	109.5	$\beta_2 = 0.011$ (-0.022 to 0.044) $\beta_3 = 0.006$ (-0.028 to 0.040) $\beta_4 = 0.003$ (-1.36 to 0.84) $\beta_5 = 0.009$ (-1.44 to 1.80) $\sigma_B = 0.10$ (0.002 to 1.65)	0.34 (0.015 to 1.68)
Model 3: Common regression coefficient	13.6	14.1	108.2	β =0.009 (-0.013 to 0.031)	0.25 (0.010 to 1.76)

Table 5-16 (continued) NMA for the possession of safety gates adjusting for covariates

	Residual deviance $\bar{D}$	pD	DIC	Beta (95%CrI)[#] σ_B(95%CrI)	Between study sd τ (95% CrI)
Unemployed (percentage at least one parent unemployed) 7 studies, 6 interventions, mean=36.4%					
No covariates	12.7	13.0	94.1		0.43 (0.014 to 1.77)
Model 2*: Exchangeable regression coefficients	13.5	14.0	95.9	$\beta_2 = 0.006$ (-1.53 to 2.30) $\beta_3 = 0.023$ (-0.38 to 1.65) $\beta_4 = 0.028$(-0.36 to 1.64) $\beta_5 = -0.032$ (-1.04 to 0.23) $\beta_6 = 0.002$ (-1.18 to 1.98) $\sigma_B = 0.210$ (0.003 to 1.81)	0.94 (0.051 to 1.94)
Model 3: Common regression coefficient	13.4	13.8	95.6	β =-0.005 (-0.16 to 0.15)	1.13 (0.045 to 4.65)
Age (mean/median age of child in years) 12 studies, 7 interventions, mean=1.48 years					
No covariates	22.5	20.1	159.8		0.23 (0.015 to 0.87)
Model 2: Exchangeable regression coefficients	22.8	22.0	162.6	$\beta_2 = -0.006$ (-1.16 to 0.83) $\beta_3 = 0.051$ (-0.97 to 0.83) $\beta_4 = -0.066$ (-2.02 to 1.62) $\beta_5 = 0.112$ (-1.46 to 1.43) $\beta_6 = 0.010$ (-2.11 to 1.77) $\beta_7 = 0.004$ (-2.05 to 1.87) $\sigma_B = 0.349$ (-0.012 to 1.69)	0.33 (0.012 to 1.29)
Model 3: Common regression coefficient	22.7	21.2	161.6	β =0.012 (-0.76 to 0.71)	0.26 (0.010 to 1.10)

pD effective number of parameters, DIC deviance information criteria

* poor mixing and autocorrelation

[#] shrunken estimates, covariates centred on the mean

5.13 Network meta-analysis using IPD and aggregate data

Of the 12 studies include in the possession of a fitted safety gate network meta-analysis, IPD was available for 10 of the studies. The interventions in these studies has been displayed in the network diagram given in Figure 5-19. The model described in section 4.13 was fitted to the studies using MCMC in WinBUGS (the code is given in Appendix M).

The results (Table 5-17), including IPD, are similar to the aggregate data results, showing that only the most intensive intervention (7) increases possession of fitted safety gates compared to usual care (OR 8.00, 95% CrI: 3.32 to 19.8, $\tau = 0.19$ (95%CrI: 0.01 to 0.78)) and compared to the other interventions. The results should be equivalent as the IPD does not add extra information when the overall mean effects are of interest and covariates are not considered. The small differences observed are probably due to how the clustering adjustment is done in the two models and possibly influence of the prior distributions.
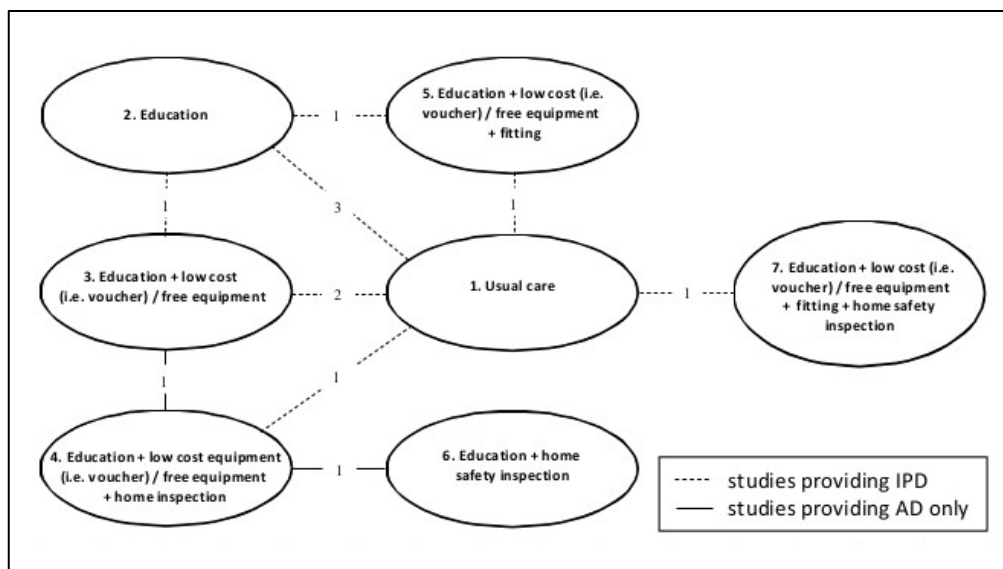
Figure 5-19 Network diagram showing IPD and AD of studies reporting possession of safety gate

Table 5-17 NMA for the possession of a fitted safety gate using IPD and aggregate

Possession of a fitted safety gates NMA*							
Interventions	**Usual care (1)**	**Education (2)**	**Education + Equipment (3)**	**Education + Equipment + Home inspection (4)**	**Education + Equipment + Fitting (5)**	**Education + Home inspection (6)**	**Education + Equipment + Fitting + Home inspection (7)**
Usual care (1)		1.47 (0.94 , 2.55)	1.58 (0.90 , 2.84)	1.30 (0.73 , 2.63)	1.46 (0.79 , 2.95)	1.45 (0.60 , 4.04)	8.00[c] (3.32 , 19.8)
Education (2)	1.43 (0.90 , 2.49)		1.07 (0.52 , 2.02)	0.88 (0.40 , 1.94)	1.00 (0.44 , 2.16)	0.99 (0.34 , 2.86)	5.43[c] (1.88 , 14.6)
Education + Equipment (3)	1.63 (0.93 , 3.03)	1.14 (0.56 , 2.23)		0.82 (0.41 , 1.79)	0.93 (0.41 , 2.24)	0.92 (0.35 , 2.69)	5.05[c] (1.77 , 14.7)
Education + Equipment + Home inspection (4)	1.28 (0.69 , 2.79)	0.90 (0.41 , 2.07)	0.78 (0.38 , 1.77)		1.13 (0.44 , 2.73)	1.12 (0.55 , 2.30)	6.16[c] (1.93 , 17.7)
Education + Equipment + Fitting (5)	1.52 (0.84 , 3.38)	1.07 (0.51 , 2.41)	0.94 (0.42 , 2.41)	1.20 (0.45 , 3.25)		0.99 (0.32 , 3.22)	5.45[c] (1.72 , 16.0)
Education + Home inspection (6)	1.43 (0.56 , 4.42)	1.01 (0.33 , 3.25)	0.88 (0.32 , 2.80)	1.12 (0.52 , 2.49)	0.94 (0.27 , 3.28)		5.52[c] (1.40 , 18.9)
Education + Equipment + Fitting + Home inspection (7)	7.80[c] (3.08 , 21.3)	5.46[c] (1.75 , 16.1)	4.77[c] (1.56 , 15.2)	6.13[c] (1.75 , 18.7)	5.07[c] (1.47 , 15.9)	5.48[c] (1.23 , 20.7)	

NMA – network meta-analysis

*OR(95%CrI) below the leading diagonal are estimated from the NMA of aggregate data only and above are estimated from the NMA using IPD and aggregate data.

[c] CrI does not contain 1

5.14 Network meta-analysis using IPD and aggregate data including a covariate

This section summarises the results of the network meta-analysis model combining IPD and aggregate data including a covariate as described in section 4.14. This analysis combines the analyses presented in section 5.12 (NMA including a covariate) and 5.13 (NMA combining IPD and aggregate data), so the studies included in this analysis have already been described. There were some problems when undertaking this analysis, these included:

- The coding given in the paper by Saramago et al did not run so after contacting the author and getting his original code the coding was changed for the IPD variable representing intervention (treat)
- Poor mixing of the beta parameters (particularly the between study beta) for the covariates representing gender and unemployment.
- The model for the covariate representing family type would not run and the reason why could not be determined
- The Phelan study recruited all households on the birth of the baby and the intervention was applied to the intervention group at varying ages with a mean of 6.3 months. The IPD data for age was not available so it was decided to exclude this study from this analysis (Phelan, Khoury et al. 2010).

The results for the ethnic group and housing tenure covariates are presented in Table 5-18 as examples of the results. As previously shown in section 5.12 when the aggregate data only was used, there is little evidence of covariate effects. There is no statistically significant evidence of ecological bias as the 95% credible intervals for the difference in the between and within regression coefficients all contain zero so the between- and within-study estimates could be combined. Table 5-18 shows the regression coefficients when between- and within-study estimates are combined for the housing tenure covariate. There is some evidence that when education is combined with equipment and a home safety inspection (β = 0.603 95%CrI: (0.053 to 1.33)) or fitting (β = 0.563 95%CrI: (0.165 to 0.971)) that the intervention is more effective in rented accommodation. Further investigation is required.

Table 5-18 NMA for the possession of a safety gate combining IPD and aggregate data including a covariate

	Intervention	Log OR Estimate(95%CrI)
Ethnic group (black or minority ethnicity) (9 IPD, 1 aggregate)		
Regression coefficients for within study association, β^W, compared to usual care	Educ (2)	0.352 (-0.493 to 1.47)
	Educ+Equip (3)	0.194 (-1.15 to 1.37)
	Educ +Equip+HIS (4)	0.422 (-0.425 to 1.82)
	Educ +Equip+Fit (5)	0.225 (-0.336 to 0.801)
	Educ +Equip+Fit+HSI (7)	0.029 (-1.18 to 0.820)
Regression coefficients for between study association, β^B, compared to usual care	Educ (2)	0.525 (-0.698 to 1.83)
	Educ+Equip (3)	0.379 (-0.894 to 1.78)
	Educ +Equip+HIS (4)	0.670 (-0.799 to 2.18)
	Educ +Equip+Fit (5)	0.508 (-1.36 to 2.51)
	Educ +Equip+Fit+HSI (7)	0.501 (-1.42 to 2.40)
Difference, $\beta^{diff} = \beta^B - \beta^W$, compared to usual care	Educ (2)	0.156 (-1.45 to 1.67)
	Educ+Equip (3)	0.200 (-1.54 to 2.06)
	Educ +Equip+HIS (4)	0.199 (-1.71 to 1.94)
	Educ +Equip+Fit (5)	0.289 (-1.64 to 2.34)
	Educ +Equip+Fit+HSI (7)	0.473 (-1.45 to 2.80)
Between study variance τ^2		0.062 (0.0002 to 1.45)
DIC		4678
Housing tenure (rented accommodation) (7 IPD, 1 aggregate)		
Regression coefficients for within study association β^W	Educ (2)	0.0003 (-0.863 to 0.688)
	Educ+Equip (3)	0.285 (-0.745 to 1.05)
	Educ +Equip+HIS (4)	0.603 (-0.002 to 1.36)
	Educ +Equip+Fit (5)	0.554 (0.14 to 0.975)
Regression coefficients for between study association β^B	Educ (2)	0.870 (-0.844 to 2.97)
	Educ+Equip (3)	0.724 (-1.13 to 3.02)
	Educ +Equip+HIS (4)	0.826 (-1.55 to 3.69)
	Educ +Equip+Fit (5)	0.856 (-1.55 to 3.67)
Difference $\beta^{diff} = \beta^B - \beta^W$	Educ (2)	0.885 (-0.965 to 3.19)
	Educ+Equip (3)	0.47 (-1.52 to 2.97)
	Educ +Equip+HIS (4)	0.222 (-2.30 to 3.09)
	Educ +Equip+Fit (5)	0.312 (-2.14 to 3.13)
Between study variance τ^2		0.377 (0.008 to 2.15)
DIC		4309
Between and within study variance combined		
Regression coefficients β, compared to usual care	Educ (2)	0.222 (-0. 578 to 0.768)
	Educ+Equip (3)	0.394 (-0.459 to 1.04)
	Educ +Equip+HIS (4)	0.603 (0.053 to 1.33)
	Educ +Equip+Fit (5)	0.563 (0.165 to 0.971)
Between study variance τ^2		0.044 (<0.001 to 1.33)
DIC		4308

5.15 Chapter summary

This chapter has presented the results of the evidence synthesis of the data from studies reporting interventions to increase the possession and use of safety equipment to prevent falls and increase falls prevention behaviour. The chapter started with a pairwise meta-analysis investigating if an enhanced intervention arm was more effective than a control/usual care intervention arm. Sub-group analyses were used to look at the results by different types of study design characteristics and meta-regression was used to consider the effect of participant characteristics. Because there was heterogeneity in the interventions reported, they were made up of various components, a network meta-analysis was used to investigate which of intervention combinations was most effective. Seven interventions were identified.

From the pairwise meta-analysis, households in the enhanced intervention arm were more likely to possess a fitted safety gate (12 studies, OR 1.61 (95% CI: 1.19, 2.17)) and less likely to possess or use a baby walker (9 studies, OR for no possession 1.57 (95%CI: 1.18, 2.09)), than families in the usual care arm. There was little evidence that the enhanced intervention increased the possession of window locks (6 studies, OR=1.17 (95%CI: 0.87, 1.57)), increased the use of bath mats or decals (4 studies, OR=1.10 (95%CI: 0.68, 1.78)) or decreased leaving a child unattended on a high surface (3 studies, OR=1.20 (95%CI: 0.84, 1.72)).

An increasing proportion of male children in a study reduces the effectiveness of the enhanced intervention to increase possession of a fitted safety gate when only aggregate data was used but is not statistically significant when IPD is included. An enhanced intervention is significantly more effective in increasing possession of a fitted safety gate amongst households in rented accommodation compared to households in the owner-occupied group. These results should be treated with caution due to the high number of tests conducted.

The NMA found that education plus home safety inspection plus providing and fitting low-cost/free equipment was the most effective intervention (OR 7.80, 95% CrI 3.18 to 21.3; p(best) = 0.97, SUCRA=99%) for increasing possession of a fitted safety gate compared to usual care and it was also statistically significantly more effective than all

of the other six interventions. It would be expected that a higher level of intensity would increase the effectiveness. Education was the most effective intervention for reducing the number of households that do not possess or use a baby walker compared to usual care (OR=0.48, 95% CrI: 0.31 to 0.84, p(best)=0.65). Equipment is not relevant in this case so this result is as would be expected. The NMAs of interventions to increase the use of window locks and reduce leaving a child unattended on a high surface showed very little difference between interventions. A NMA could not be conducted of interventions to increase the possession of bath mats as the network was not connected. The results were robust when studies that were not RCTs were excluded.

Where IPD was available it was combined with the remaining aggregate data and incorporated into the NMA. The effects of covariates were also investigated in both the NMA of aggregate study data and the NMA incorporating IPD. There was little evidence of any covariate effects but the number of studies was decreased for the most complex analysis. There were a few problems with the IPD and aggregate data NMA that need to be investigated further, possibly by simplifying the model and not splitting the between-study and within-study variability and exploring the IPD further.

Sensitivity analyses on the choice of prior distributions, particularly for the between-study variances, and on the choice of initial values were conducted. The results were not sensitive.

The results from the NMA (without IPD and covariates), because they are more relevant than the pairwise meta-analysis results to policy makers and the providers of the interventions, will be incorporated into the decision modelling described in Chapter 6. The results are presented in chapter 7.

6   Decision Modelling for Economic Evaluation

An economic evaluation using a comprehensive decision modelling approach, which will be used to explore the cost-effectiveness of the different interventions used to prevent falls down stairs in Chapter 7, is described in this chapter. This approach has four stages (Cooper, Sutton et al. 2004): (i) a systematic review and meta-analyses (described in Chapters 3 and 4), (ii) estimation of model inputs (including effectiveness, transition probabilities and costs) described in sections 6.1 and 6.2, (iii) sensitivity analysis for data and model specifications described in section 6.4, and (iv) evaluation of the model described in section 6.3. The approach is described in this chapter in which stages (i)-(iv) can be evaluated simultaneously within a single Bayesian model using MCMC simulation and the decision model is presented in Chapter 7.

6.1   Economic evaluations

6.1.1   Introduction

Economic evaluations are increasingly being applied in health care, including the assessment of prevention programmes, and are used to inform policy and decision making using limited resources efficiently (Husereau, Drummond et al. 2013). Decision analytical modelling compares the expected costs and consequences (utilities) of decision options by synthesising information from multiple sources and applying mathematical techniques usually with computer software. The aim is to provide decision makers with the best available evidence to reach a decision accounting for any variability, uncertainty and heterogeneity associated with possible decisions (Petrou, Gray 2011, Briggs, Claxton et al. 2006). Variability is the variation between individual subjects, with the same underlying characteristics, that occurs by chance. Heterogeneity relates to differences between the individual subjects that can partly be explained, often because characteristics, such as age of child, socioeconomic background and ethnicity, have been recorded in the evidence base and can be included in the modelling process. Decision models aim to capture uncertainty in the

expected costs and benefits for each intervention by considering the uncertainty surrounding the inputs into the model (Briggs, Claxton et al. 2006). There are two types of uncertainty: parameter uncertainty and structural uncertainty and they are described further in section 6.1.4.

Most decision models tend to use a cohort model to characterise the experience of the "average" subject from a population sharing the same characteristics and focussing on expected costs and health effects rather than explicitly considering the outcome for the individual subject (Briggs, Claxton et al. 2006). In this thesis decision trees and Markov models have been used to specify a cohort model and are described below. A subject level model can be used to estimate the mean costs and health benefits by considering the costs and health benefits of each individual, hence allowing for variability in patient outcomes (Davis, Stevenson et al. 2014) but is not used in this thesis.

6.1.2   Decision trees

A decision model considers the possible consequences that arise from a set of at least two alternative health care interventions being evaluated and decision trees are the simplest form (Briggs, Claxton et al. 2006). Probabilities are estimated for the consequences and, in the case of an economic decision model, each of the consequences has an outcome in terms of a measure of the benefit (utility) and cost. The expected costs and utilities for each of the interventions being considered can be estimated by weighting the costs and utilities by the estimated probabilities in a cost-utility analysis. An example of a decision tree for a decision question is given in Figure 6-1.

Each branch in Figure 6-1 reflects possible mutually exclusive routes through the decision tree from left to right. The decision is the choice of two interventions, A and B. On intervention A the subject could have no injury (probability p_1) or an injury (probability $1-p_1$) from which they could recover (probability p_2) or not recover (probability $1-p_2$). Similarly for intervention B. The decision tree is "averaged out" and "rolled back" using the probabilities allowing the expected costs and utilities, such as life-years or quality adjusted life years (QALYs), of each intervention to be calculated at

131

the terminal nodes (Gray 2011, Petrou, Gray 2011). Measuring utilities and costs is described in section 6.2.

Decision trees can only move forward, left to right, and it is difficult to incorporate recurring outcomes without ending up with a large number of long and complex pathways (Briggs, Claxton et al. 2006, Petrou, Gray 2011). A Markov model is an alternative to a decision tree or can be used in combination with it; it is called a recursive decision tree as it allows recurring outcomes. Because of the flexibility of the approach, a combined decision tree and Markov model has been used for this application. Markov models are described in section 6.1.3.

Pathway Intervention A	Probability	Cost	Expected cost	Utility	Expected Utility
1	p_1	C_1	p_1*C_1	U_1	p_1*U_1
2	$(1-p_1)*p_2$	C_2	$((1-p_1)*p_2)*C_2$	U_2	$((1-p_1)*p_2)*U_2$
3	$(1-p_1)*(1-p_2)$	C_3	$((1-p_1)*(1-p_2))*C_3$	U_3	$((1-p_1)*(1-p_2))*U_3$
Total	1		total expected costs for intervention A		total expected utilities for intervention A
Intervention B	Probability	Cost	Expected cost	Utility	Expected Utility
4	p_3	C_4	p_3*C_4	U_4	p_3*U_4
5	$(1-p_3)*p_4$	C_5	$((1-p_3)*p_4)*C_5$	U_5	$((1-p_3)*p_4)*U_5$
6	$(1-p_3)*(1-p_4)$	C_6	$((1-p_3)*(1-p_4))*C_6$	U_6	$((1-p_3)*(1-p_4))*U_6$
Total	1		total expected costs for intervention B		total expected utilities for intervention B

Figure 6-1 Example of a decision tree

6.1.3 Markov models

A Markov model can be used to capture the transition between various health states. Patients are assumed to be in one of a finite number of health states at any point in time and make transitions between those health states over a series of discrete time intervals or cycles (Spiegelhalter, Abrams et al. 2004, Cooper, Sutton et al. 2004, Briggs, Claxton et al. 2006, Drummond 2005). Transition probabilities of staying in a state or moving to another in a cycle need to be determined and a termination condition must be set which can be a specified number of cycles, e.g. years a subject is followed up for, or moving into an absorbing state that cannot be left, e.g. dead.

A hypothetical cohort of individuals is usually followed through a Markov model over time so expected costs and utilities can be estimated. A simple example of a Markov model is given in Figure 6-2. During each cycle of the model, e.g. a year, subjects can remain in states A, B or C, move between these states (with restricted movement, e.g. they cannot move to A from B or C), or move into the absorbing state, death(D), from which they cannot move. The transition probabilities (tp) for moving between states are given in the transition probability matrix in Figure 6-2. Figure 6-2 also shows the costs (C) and utilities (U) attached to each health state ($A - D$). The costs and utilities can be combined with the transition probabilities to calculate expected costs and utilities for each cycle and summed over cycles.

A conventional two-stage approach can be used in which the effectiveness parameter estimates and their uncertainty are firstly estimated using a meta-analysis and then secondly these estimates are assigned distributions, input into the decision model and evaluated using MCMC, for example in an Excel spreadsheet or a statistical software package. When using NMA the effect estimates for the different intervention comparisons are estimated jointly which, in most cases, induces correlations. This correlation structure needs to be maintained when specifying a distribution for the absolute intervention effects for the decision model. WinBUGS can be used to fit the NMA and the coda output, for each iteration of the sampler, can be extracted and used as the empirical distribution in the decision model (Dias, Sutton et al. 2013c).

A better approach is to integrate the two stages described above so the evidence synthesis results, for example from the network meta-analysis presented in section 5.9, can be integrated into the probabilistic decision model as a single process. This means that the joint posterior distribution of the absolute effects of the interventions are fed into the model and the uncertainty and correlation propagated through the model. This is referred to as "Comprehensive Decision Analysis" (Cooper, Sutton et al. 2004, Dias, Sutton et al. 2013c).

A potential disadvantage of Markov model is that it is memoryless, i.e. it has no memory of the previous state or the time of the transition (Briggs, Claxton et al. 2006). During each cycle of the model the transition probabilities depend only on the present state and not on any history of the subject.

Transition probability (tp) matrix

		To			
	State	A	B	C	Death (D)
From	A	tp^{AA}	tp^{AB}	tp^{AC}	tp^{AD}
	B	0	tp^{BB}	0	tp^{BD}
	C	0	tp^{CB}	tp^{CC}	tp^{CD}
	Death (D)	0	0	0	$tp^{DD} = 1$

where tp^{XY} is the transition probability of moving from state X to state Y
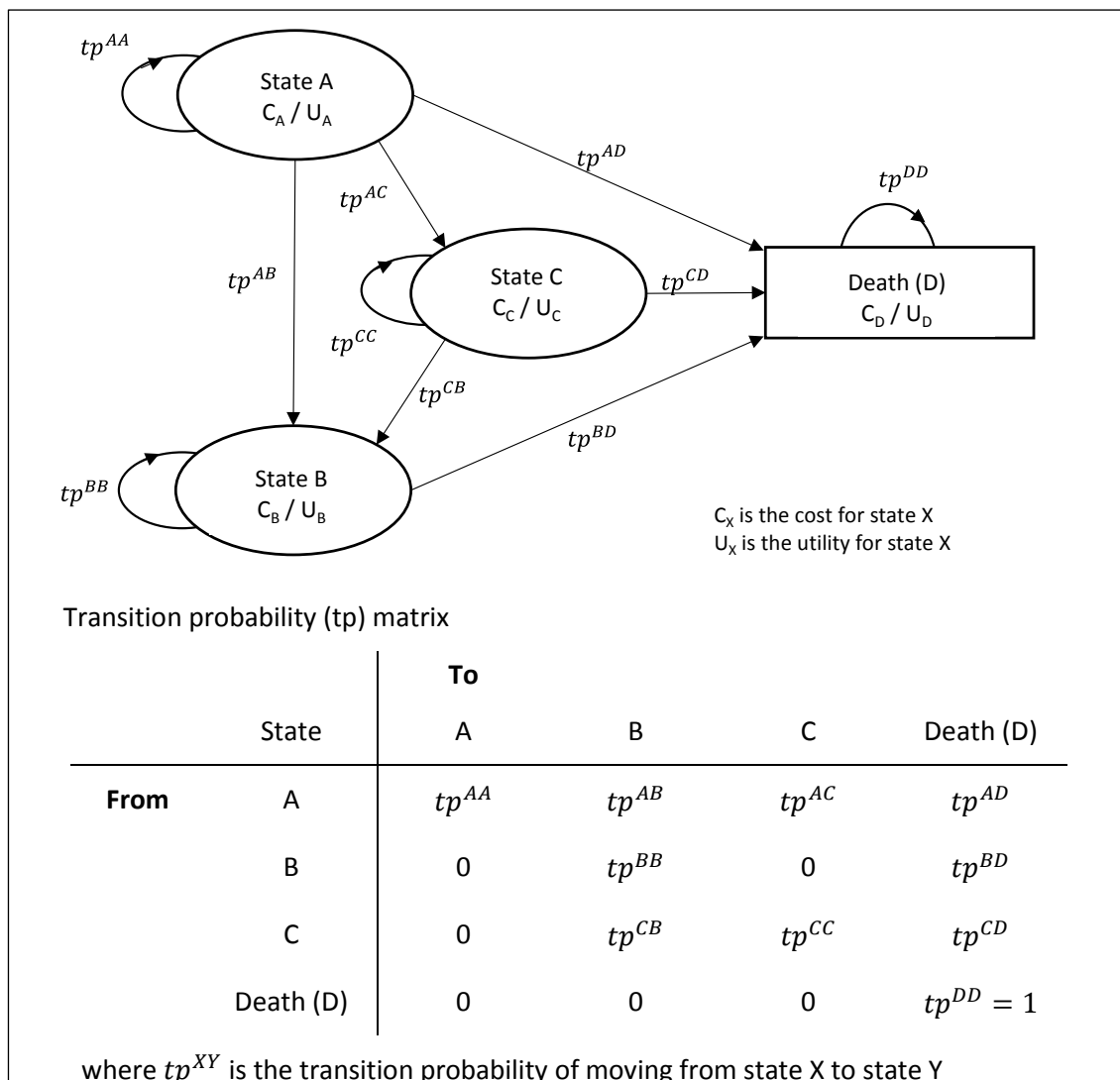
Figure 6-2 Example of a Markov model and transition probability matrix

6.1.4 Uncertainty in a decision model

To determine the expected costs and benefits accurately, it is necessary to consider the uncertainty surrounding the inputs to the model. The examination and reporting of uncertainty is good decision modelling practice and there are two different types of uncertainty (Briggs, Weinstein et al. 2012):

- Parameter uncertainty – in the estimation of the parameter of interest
- Model/Structural uncertainty – inherent in the form chosen for the decision model.

Parameter uncertainty may be represented by including parameters as probability distributions, often referred to as a probabilistic sensitivity analysis (PSA). PSA provides a form of sensitivity analysis which allows investigators to easily see the joint impact of the uncertainty in multiple parameters on the expected costs, benefits and on decision uncertainty (Dias, Sutton et al. 2013c). This method is recommended by NICE as the preferred method to explore the uncertainty arising from imprecision in model parameters and providing the best estimates of mean costs and outcomes (NICE 2012).

A Markov model, using Markov chain Monte Carlo (MCMC) simulation techniques implemented in WinBUGS, is used in Chapter 7 to allow the uncertainty in model input parameter values, such as the transition probabilities, costs and health utilities, to be incorporated as probability distributions (Cooper, Spiegelhalter et al. 2013, NICE 2014).

Structural uncertainty should be set out in the choice of decision model described in section 6.2 as different assumptions made can impact on the estimated uncertainty.

6.2 Developing the decision model

6.2.1 Describing the base case

Guidelines have been published that attempt to consolidate reporting economic evaluations of health care interventions (Husereau, Drummond et al. 2013) and NICE

have a similar summary in their manual for developing NICE guidelines (NICE 2014). The description below is based on these sources.

A clear explicit statement should be made of the study question and its relevance for health policy or practice decisions. The methodological assumptions for the base case of a decision model given in Table 6-1 need to be clearly described.

6.2.2   QALYs and utilities

The QALY is a generic measure of health outcome used to make comparisons across different healthcare interventions. It incorporates the impact of an intervention on a subject's length of life and on their health-related quality of life (HRQoL or utility score) (Whitehead, Ali 2010, Brazier 2007).

QALYs are calculated by multiplying health state preference scores, or utility weights, by years of life.

> QALYs = number of years lived x utility

Utilities are often measured on a scale of 0 to 1 where 0 indicates death and 1 represents perfect health but some scales can give negative utilities indicating a state worse than death. So for example, a subject allocated a utility of 0.8 and followed up for ten years will have 8 QALYs. Utilities can be based on a variety of measures including health related questionnaire and visual analogue scale scores such as the EQ-5D. NICE recommend EQ-5D as the preferred measure of health-related quality of life in adults and in clinical trials it is usually collected from subjects alongside the clinical outcome measure (NICE 2014). When it is not available from clinical studies, data can be sourced from the literature. Kind et al in 1999 produced a series of tables of age/sex population norms in adults aged 18 and over for EQ-5D that can be used as baseline utility values (Kind, Dolan et al. 1998). Similar values for children are not available, although there are specific validated measures for measuring HRQol in adolescents (aged 11-17) such as KIDSCREEN-10 (Kidscreen), CHU9D (Stevens, Ratcliffe 2012) and EQ-5D-Y (Ravens-Sieberer, Wille et al. 2010). There are challenges in how to elicit the health state preferences (self- or proxy- reported) and cross country differences so these instruments do not always give comparable results (Chen, Flynn et al. 2015). The

adult population norms can be extrapolated back to earlier ages and these values adjusted for injury specific utilities found in the literature. More on these adjustments is discussed in chapter 7.

Table 6-1 Summarising the base case of the decision model
(Husereau, Drummond et al. 2013, NICE 2014)

Element of Assessment	
Type of economic evaluation	A cost-utility analysis is recommended for interventions with health outcomes in NHS settings.
Perspective on costs	Whether only NHS and public sector settings (PSS) costs are considered or wider societal costs are included.
Perspective on outcome	All direct health effects on individuals but could include non-health benefits.
Prevention strategy	Interventions being compared.
Evidence on outcomes	Systematic review and meta-analysis
Measure of health effect (utilities)	Quality Adjusted Life Years (QALYs) (see section 6.2.2)
Main source of data for measurement of health related quality of life (HRQL)	Reported directly by patients
Source of preference data for valuation of changes in HRQL	Representative sample of the UK population.
Base year for calculating costs/prices	All costs should be based on the same base year.
Discount rate	Discounting is applied to generate the present value of expected costs and outcomes. Same annual discount rate should be applied to all costs and health effects.
Target cohort	Base case population to be simulated and followed through the decision model.
Time horizon	The start and end points (in time) over which the costs and consequences of a health intervention will be measured and valued. Long enough for all important differences in costs or outcomes between the interventions being compared to have an impact.

It can be difficult to find a probability distribution for the expected utility, it has an upper bound of 1 but many utility scales can be negative. If the expected utility is close to 1 and the variance is small a beta distribution can be used (equation (6.1)). The beta distribution is constrained to lie between 0 and 1 so is not appropriate for states close to death that may be negative. The utility is often subtracted from 1 to give a utility decrement and a lognormal or gamma distribution can be applied (Briggs, Claxton et al. 2006).

$x \sim beta(a, b)$ represents a distribution with the properties

$$p(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad \text{for } 0 < x < 1, a, b > 0$$

(6.1)

$$E(X|a,b) = \frac{a}{a+b}$$
$$Var(X|a,b) = \frac{ab}{(a+b)^2(a+b+1)}$$

QALYs that occur in the future are discounted to current values, to incorporate the idea that people prefer to receive health benefits now rather than in the future (i.e. positive time preference). Section 6.2.4 discusses discounting further.

6.2.3 Costs

There are 3 stages involved in the process of costing health care interventions (Drummond 2001):

1. Identification of costs. The costs can be:
   - direct as a cost of and result of intervention both in terms of materials, equipment, overheads, medical costs, doctor time
   - indirect costs, the opportunity costs of patient and care givers losing time whilst being sick or providing unpaid care
   - patient costs such as transport and out of pocket expenses
   - future costs directly and indirectly related to the intervention
   - intangible costs, distress, anxiety and impact on QOL resulting from poor health and treatment. Very difficult to measure.
2. Measurement of costs. There can be many sources of cost data. For example in this thesis PSSRU provides estimates of national costs for a wide range of health and social care costs of health visitor time, emergency treatment and stays in hospital (Curtis 2012) and costs of installing intervention equipment are obtained by personal communication with the company installing the equipment.
3. Translation into a monetary amount. All costs should be identified but they are not all measurable.

Costs are often modelled as Gamma distributions (equation (6.2)). They are constrained to be non-negative and are made up of counts of resource use weighted by unit costs (Briggs, Claxton et al. 2006). Alternatively, log normal distributions can be used as they also constrain values to be non-negative.

X ~ Gamma[a,b] represents a distribution with the properties

$$p(x|a,b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \quad \text{for } x > 0, \; a, b > 0$$
$$E(X|a,b) = \frac{a}{b}$$
$$Var(X|a,b) = \frac{a}{b^2}$$

(6.2)

6.2.4 Adjustments to costs, discounting and inflation

Resource use and costs may have been collected from different periods of time, different countries and in different currencies. Adjustments need to be made to make these relate to the same time and units. The key adjustments are discounting, inflationary adjustments and currency conversion. (Gray 2011)

Discounting

Costs (and health benefits) occurring at different times should not be given the same weighting and should reflect when they are incurred (and realised). There is a preference to delay cost as long as possible and receive the health benefits as soon as possible. Costs (and benefits) occurring today are valued more highly than those that will occur in the future (Drummond 2005, Gray 2011, Phillips 2005). Discounting costs reflects individual preference for costs to be experienced in the future rather than the present. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future (NICE 2012).

Future costs and benefits are discounted using equation (6.3).

$$Cp = \sum_t \frac{Cf_t}{(1+R)^t}$$

(6.3)

where Cp is the present value of costs, Cf_t is the future cost at year t, and R is the rate of discount.

The discount rate of 3.5% in the UK is set by the Treasury in The Green Book (HM Treasury 2011) and is recommended by NICE when appraising healthcare technologies and public health interventions (NICE 2012, NICE April 2013).

Inflation

It is also important to ensure that all intervention costs are placed on a common base year. Costs could be determined for different years and should be adjusted to the base year to eliminate the effects of inflation. The Hospital and Community Health Services (HCHS) pay and price index in England is a weighted average of two separate inflation indices: the pay cost index (PCI) and health service cost index (HSCI). 1987/88 is the base year and the pay and price index for 2011/12 is 285.7. To convert a cost, e.g. £500, from 2009/10 prices (HCHS index=268.6) to 2011/12 prices use equation (6.4) (Curtis 2012).

$$2011/12 \text{ prices} = \frac{HCHS \text{ } index \text{ } 2011/12}{HCHS \text{ } index \text{ } 2009/10} \times \text{price } 2009/10 = \frac{285.7}{268.6} \times £500 \quad (6.4)$$
$$= £531.65$$

The methodology for the pay cost index was revised in 2011/12 so slightly different values are now used (282.5 for 2011/12) (Curtis 2017). The model described in chapter 7 uses the indices published in 2012 (Curtis 2012).

6.3   Assessing cost-effectiveness

6.3.1   Cost-effectiveness plane

Economic evaluations compare the costs and effects of several different interventions. There are four possible outcomes when comparing a pair of interventions A and B represented by the four quadrants in Figure 6-3. Intervention A can be more costly and lead to lower health gains than B (NW quadrant). In that case A is said to be dominated by B. Vice versa, if A is less expensive, but leads to better outcomes, it is said to

dominate B (SE quadrant). Interventions that dominate another tend to be recommended normally. In the NE quadrant A is more effective and more expensive than B and decisions need to be made whether paying more for better outcomes is 'worth it'. Similarly, in quadrant SW, intervention A is less effective but also less costly than B, and decision are needed on whether the cost savings are worth the health losses. Decisions are usually made based on the incremental cost-effectiveness ratio (ICER) described in section 6.3.2.

The outputs from a probabilistic decision model will give the distribution over the difference in costs, difference in effects and the joint cost-effect distribution for the interventions.

6.3.2 Incremental cost-effectiveness ratios (ICERs)

Costs and effects are usually reported as ICERs. It is the measure of cost effectiveness of a health intervention compared with an alternative, defined as the difference in costs (incremental cost) divided by the difference in health effects (incremental effect). Equation (6.5) gives the ICER for a cost-utility analysis comparing two interventions, A and B, using QALYs as the measure of health effect.

$$\text{ICER} = \frac{(\text{cost of intervention B}) - (\text{cost of intervention A})}{(\text{QALY B} - \text{QALY A})} = \frac{\Delta C}{\Delta E} < K \quad (6.5)$$

where
ΔC is the difference in costs and ΔE is the difference in QALYs
K is the acceptable threshold ratio

The acceptable threshold ratio (K in Equation (6.5) and Figure 6-3) is the maximum amount a decision maker may be willing to pay for a unit gain in health (QALY). The ICER (or ICUR as it is sometimes known for a cost-utility analysis) at the acceptable threshold ratio value, (K in Figure 6-3), can be plotted on the cost-effectiveness plane as a straight line that passes through the origin and the coordinate$(\Delta E, \Delta C)$. If the ICER is lower than this acceptable threshold ratio of the decision maker and in the NE quadrant in Figure 6-3 then the intervention should be recommended (Briggs, Claxton et al. 2006). This threshold ratio is often unknown in practice. NICE has not identified an ICER above which interventions should not be recommended, however, based on the decisions they have made, interventions with an ICER of less than £20,000 per

QALY gained are considered to be cost-effective (NICE 2013). ICERs above £20,000 per QALY gained are judged on the degree of certainty around the ICER, the presence of strong reasons indicating that intervention is innovative and adds health gains that may not have been adequately captured (NICE 2014). Above an ICER of £30,000 per QALY gained, a very strong case has to be made (NICE 2013).
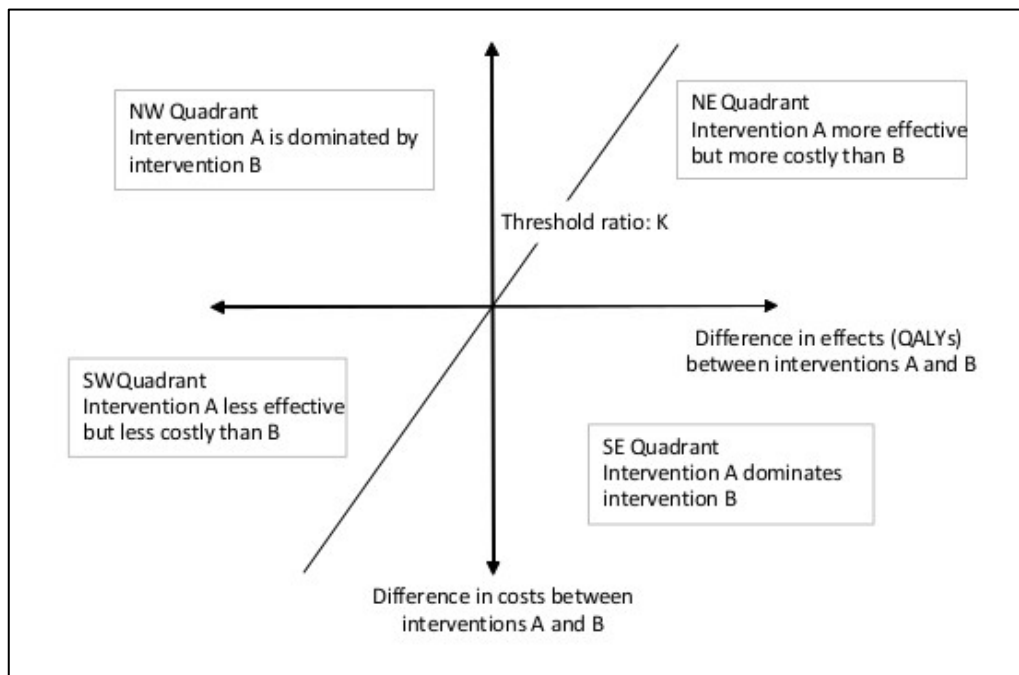
Figure 6-3 Cost-effectiveness plane

Interval estimates for the ICER can be found using the simulation results but there are problems with the ICER particularly when the simulation results cross the axes. Ratios of the same sign but from different quadrants are not strictly comparable. Negative ICERs in the NW quadrant of Figure 6-3 have a different interpretation to negative ICERs in the SE quadrant but will be grouped together if the ICERs are ranked (Briggs, Claxton et al. 2006). To overcome this issue the incremental net benefit is calculated.

6.3.3 Incremental net (monetary) benefit (INB) and cost-effectiveness acceptability curves (CEAC)

The incremental net benefit (Stinnett, Mullahy et al. 2013) is found by rearranging the ICER decision rule, $\Delta C/\Delta E < K$, to give equation (6.6).

$$INB(K) = K\Delta E - \Delta C > 0 \tag{6.6}$$

where

K is the threshold ratio

This avoids the calculation of ambiguous ratios.

The simulated INB values can be used to calculate the probability that an intervention is cost-effective for a given threshold K (Welton 2012) (equation (6.7)).

$$ProbCE(K) = P(INB(K) > 0) = \frac{\text{No. of simulations } INB > 0}{\text{Total no. of simulations}} \tag{6.7}$$

A cost-effectiveness acceptability curve (CEAC) shows the probability intervention A (versus intervention B) is cost-effective for different values a decision maker is willing to pay (K). Several interventions can be compared to B and plotted on the same graph.

6.3.4   Dominance and extended dominance

When several alternative interventions (e.g. A1, A2, A3, A4, A5) are compared to an intervention (B) then the principles of dominance and extended dominance can be applied. Dominance and extended dominance are illustrated in Figure 6-4.

All interventions A1-A5 are more effective and more costly than intervention B. The aim is for maximum effect and minimum cost and so the line that joins the origin with A2, A3 and A5 is the cost-effectiveness frontier as these interventions are the ones that are closest to meeting this aim. Intervention A1 is more costly and less effective than intervention A2 so is said to be dominated by A2 and hence can be eliminated from comparisons. Intervention A4 is not dominated by either A3 or A5, it is more effective but more costly than A3 and less costly but less effective than A5 but, along the cost-effectiveness frontier between a and b there are points that dominate A4 and hence a combination of A3 and A5 may be preferable to A4 and A4 is said to be extendedly dominated by A3 and A5 (Gray 2011).
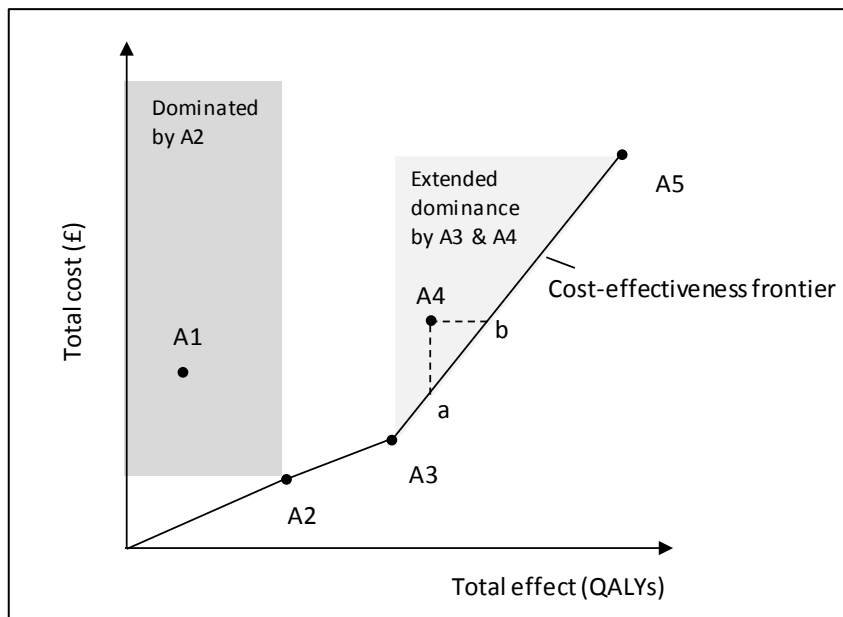
Figure 6-4 Dominance and extended dominance (Gray 2011)

6.4 Sensitivity analysis and subgroup analysis

Probabilistic sensitivity analysis (PSA) is used by running the model each time varying the parameter values across specified distributions. Structural or model uncertainty reflects the uncertainty surrounding the structure of the model and the assumptions underpinning it. It is usually examined using sensitivity analysis, re-running the model with different model structural assumptions. Heterogeneity reflects the difference between subgroups of patients, and any sub-group analyses should be pre-defined and justified in terms of their relevance.

6.5 Standards for developing and reporting

In developing and reporting a decision analytic model the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (Husereau, Drummond et al. 2013), principles for good modelling practice and design in Philips et al (Philips, Bojke et al. 2006), together with the methods for the development of NICE Public Health guidance (NICE 2012) should be used.

6.6   Summary

This chapter has presented the methods for undertaking a cost-utility analysis using a comprehensive probabilistic decision modelling framework so the effectiveness estimates, with estimates of uncertainty, from the meta-analyses can be incorporated into a single decision model. Decision trees, Markov models and the model inputs have been described and how they are used to estimate expected costs and utilities with uncertainty accounted for using probabilistic sensitivity analysis (sections 6.1and 6.2). To evaluate the cost-effectiveness of alternative interventions ICERs, INB and CEACs are described (section 6.3). Sensitivity analyses around structural and model uncertainty are recommended (section 6.4).

7   Decision Model Development and Evaluation

This chapter presents the development of a comprehensive decision model to investigate the cost-effectiveness, using a cost-utility analysis as described in Chapter 6, of interventions aiming to increase the use of safety gates to prevent falls down stairs. The model development, including discussion of some of the assumptions made, is described in section 7.1 - 7.3. The model, broken down into three stages, is described in sections 7.4, 7.5 and 7.6. Utilities, implementation and sensitivity analyses are discussed in sections 7.7, 7.8 and 7.9 respectively. The results are presented in section 7.10. The results are reported in the NIHR report for the Keeping Children Safe at Home Programme (Kendrick, Ablewhite et al. 2017), https://www.journalslibrary.nihr.ac.uk/pgfar/pgfar05140/#/abstract. My contribution included: network meta-analysis for falls prevention interventions and decision model evaluating cost-effectiveness of interventions to increase possession of a fitted safety gate.

7.1   Economic evaluation

A cost-utility analysis, using a comprehensive decision model (section 6.2) implemented using Bayesian MCMC simulation approach, integrating the evidence synthesis (network meta-analysis section 5.9) and decision model in a combined decision tree and Markov model, is developed (Cooper, Sutton et al. 2004, Welton 2012). The model compares a range of different intervention strategies aimed at increasing possession of a fitted safety gate to reduce falls in children under three in the home; children under three have the highest risk of falling down stairs and a safety gate is recommended until the child is at least two years of age (NHS Choices 2016). The interventions are identified in the published systematic review (Young, Wynn et al. 2013) described in chapter 3 and the network meta-analysis (Hubbard, Cooper et al. 2015), detailed in chapter 5, as: (1) Usual care (UC); (2) Education (E); (3) Education + free or low cost equipment (E + FE); (4) Education + free or low cost equipment +

home safety inspection (E + FE + HI); (5) Education + free or low cost equipment + fitting (E + FE + F); (6) Education + home safety inspection (E + HI); (7) Education + free or low cost equipment + fitting + home safety inspection (E + FE + F + HI).

7.2 Model structure

The model follows a hypothetical population of new-borns over their first three years of life to investigate the impact the possession of fitted safety gates on the stairs in their household would have on the overall lifetime (100 years in total to account for most of the population being dead by this time) costs and quality of life. The first three years are used because the use of a safety gates beyond the age of 3 years is not recommended (Hayes, Kendrick et al. 2014). 100,000 households are simulated from a general UK population and the model assumes a single child in a household who would benefit from fitted safety gates on the stairs. These assumptions will be discussed further and the impact of some assessed in sensitivity analyses (section 7.10.2).

Findings are expressed in terms of incremental cost-effectiveness ratios (ICERs) and probabilities of interventions being cost-effective at different decision-makers' cost per additional QALY thresholds (section 6.3.2). Interventions are determined to be dominant if they have lower costs and are more effective than an alternative intervention and extendedly dominant if they have lower costs and are less effective but the ICER is higher (section 6.3.4). The probability that an intervention is cost-effective is presented for an ICER of £30,000 and also £50,000 (section 6.3.2).

A three stage comprehensive decision model (Cooper 2004, Welton 2012) is developed based on models used to investigate the cost-effectiveness of smoke alarm give away schemes on health outcomes in children (Saramago, Cooper et al. 2014, Pitt, Anderson et al. 2009). Figure 7-1 presents a schematic diagram of the model structure with more detailed descriptions of the three stages in sections 7.4, 7.5 and 7.6. The model is used to estimate lifetime QALYs and costs of the interventions are estimated from a public sector perspective. Costs include National Health Service (NHS) and Personal Social Services (Curtis 2012) costs, discounted at the standard annual rate of 3.5% for both costs and health effects (NICE 2014). The first stage of the model is the intervention stage in which a decision tree format is used to evaluate the costs and effectiveness of

Stage 1: Intervention stage	Stage 2: Markov model for child aged 0-3	Stage 3: Long term Markov model (age 3-100)
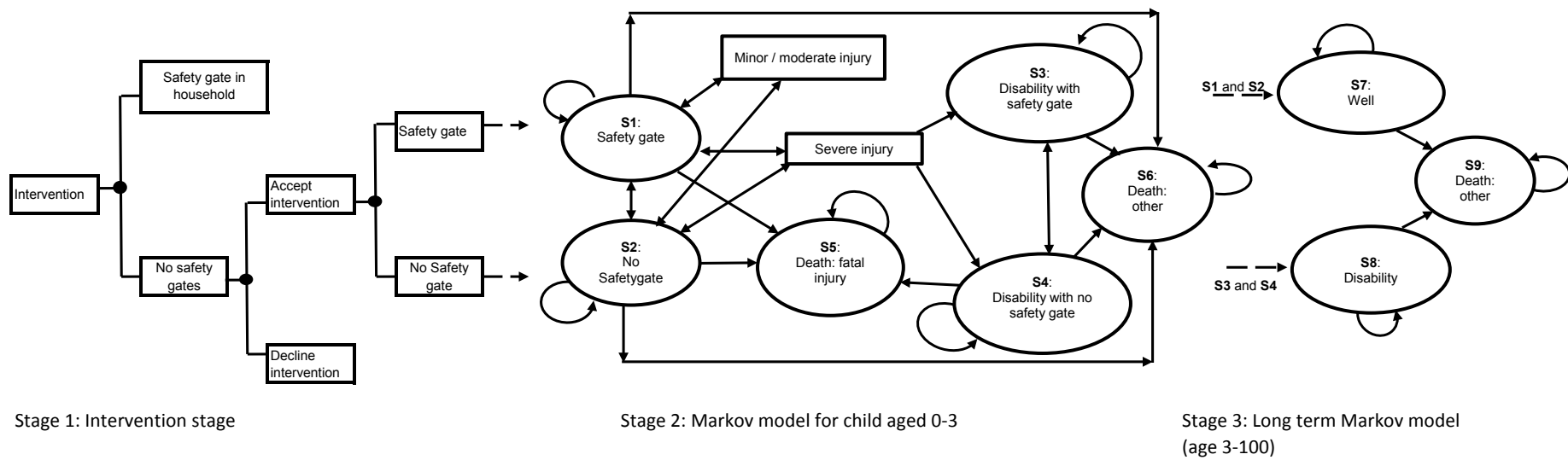
Figure 7-1 A schematic diagram of the decision model structure

(arrows indicate direction of households/individuals through the model)

the seven interventions (section 7.4). At the end of this stage households will either possess or not possess fitted safety gates. Stage 2 uses a Markov structure to estimate the costs and QALYs associated with the interventions over the first three years of the child's life and has six distinct states (S1-S6) based on safety gate possession and health (section 7.5). At the end of stage 2 the child is in one of four states; two absorbing death states (S5 and S6), from which the child cannot move, and states from which the child can move into the next stage, well (S1 and S2) and disabled (S3 and S4). Stage 3 uses a Markov structure with three states for the child/individual from age three to 100, well (S7), disabled (S8) and dead from other causes (S9) (section 7.6). There are costs and utilities attached to each of the states and these are described in sections 7.4.2, 7.5.2, 7.6.1 and 7.7.

7.3   Model assumptions

A summary of the base case methodological assumptions is outlined in Table 7-1 using the structure outlined in Table 6-1.

Other assumptions in the modelling include:

- The possession of fitted safety gates in the household is a surrogate/intermediate outcome linked to a reduction in risk of injury/death due to a stairway fall.
- Probability of a household accepting an intervention is assumed the same across all interventions due to a lack of information on the acceptance of the different programmes and the wide range of interventions.
- Benefit of a household possessing fitted safety gates is for a single child aged 0 to 3 years of age. It ignores potential (positive or negative) effects on sibling(s) and/or parent(s) living in the same household, e.g. an older child may climb over the safety gate and therefore have an increased risk of injury. The number of children is increased to 1.8 in sensitivity analysis 7 to reflect the average number of children in a UK household (Table 7-10).

Table 7-1 Summary of the decision model base case

Element of Assessment	Base case
Type of economic evaluation	Cost-utility analysis
Perspective on costs	Public sector UK, NHS and Personal Social Services (PSS) (Curtis 2012)
Outcome event	Accidental fall down stairs
Perspective on outcome	All health effects on individuals
Prevention strategy	Two fitted safety gates, top and bottom of stairway as recommended by NHS Choices (NHS Choices 2016)
Effectiveness evidence on outcomes	Network meta-analysis to simultaneously synthesise evidence from seven interventions to increase the possession of a fitted safety gate (Hubbard, Cooper et al. 2015)
Measure of health effect (utilities)	Quality Adjusted-Life Years (QALYs)
Main source of data for measurement of health related quality of life (HRQL)	Reported directly by patients (HALO report) (Nicholl, Turner et al. 2009)
Source of preference data for valuation of changes in HRQL	Representative sample of the public (UK Population norms) (Kind, Hardman et al. 1999)
Base year for calculating costs/prices	2012, costs/prices prior to 2012 are inflated (section 0) using the Hospital & Community Health Services (HCHS) index (Curtis 2012) (Sensitivity analysis 1 (Figure 7-5)
Discount rate	3.5% annual rate for both costs and utilities (NICE 2012) (section 0)
Simulated cohort	100,000 UK households with a single child aged 0-3
Number of intervention strategies	7 (Hubbard, Cooper et al. 2015)
Reference (comparator) intervention	Usual care
Time horizon	100 years in 1 year cycles

- Probability of a future stairway fall injury is assumed to be independent of previous stairway fall injuries, and remains constant throughout the relevant model timeframe (i.e. 3 years for part 2 of the model). Evidence does indicate that some children are more likely to have repeated falls injuries (Towner, Dowswell et al. 2005)

- Only one stairway fall injury allowed in a single one-year cycle.

- Intervention offered when the child is born

- The child may down the stairs if a safety gate is fitted. A child could have a fall with a fitted safety gate if not used appropriately or child can climb. No evidence found.

Other assumptions are described when the transition probabilities and costs are presented in sections 7.4 - 7.6 and some assumptions will be assessed in a sensitivity analysis in section 7.10.2.

7.4   Stage 1: Intervention stage

A decision tree structure is used to estimate the costs and outcome, in terms of increasing possession of a fitted safety gate to prevent a stairway fall, associated with the seven interventions being compared. This is referred to as the intervention model (Figure 7-2) and accounts for baseline prevalence of possession of a fitted safety gate.

7.4.1   Stage 1: Transition probabilities

Results from the network meta-analysis (section 5.9) that estimated the effectiveness of the interventions in increasing the possession of a fitted safety gate, are used to inform stage 1 of the model. These results are integrated into the decision model so the posterior distribution of intervention effects with the between-study precision is input directly into the model (Cooper, Abrams et al. 2003).

An estimate of the population probability of possession of fitted safety gates on the stairs prior to intervention is determined using a pairwise meta-analysis of studies giving either baseline or usual care arm estimates identified in the systematic review informing the network meta-analysis (Young, Wynn et al. 2013). This analysis is also integrated into the decision model so the posterior distribution of the probability is used to inform stage 1. A list of the studies in this meta-analyses is given in Appendix O and the results are presented in Table 7-2, showing an estimated probability of 0.56 (95%CrI: 0.43, 0.68) for possessing fitted safety gates in households with a child aged

0-3 years. These studies are heterogeneous (I^2=96.2%, p<0.001), representing different socio-economic groups and countries and the usual care arm is very varied.

Similarly, the probability of a household accepting the intervention if they did not already possess fitted safety gates is estimated using a meta-analysis of studies reporting the proportion of households accepting interventions in the systematic review (10 studies) (Young, Wynn et al. 2013). This probability is assumed the same for all interventions as no information is available to set different probabilities and is integrated into the decision model so the posterior distribution is used to inform stage 1. A list of the studies in this meta-analysis is given in Appendix P and the results are presented in Table 7-2, showing an estimated probability of 0.77 (95%CrI: 0.53, 0.90) for accepting an intervention in households with a child aged 0-3 years.
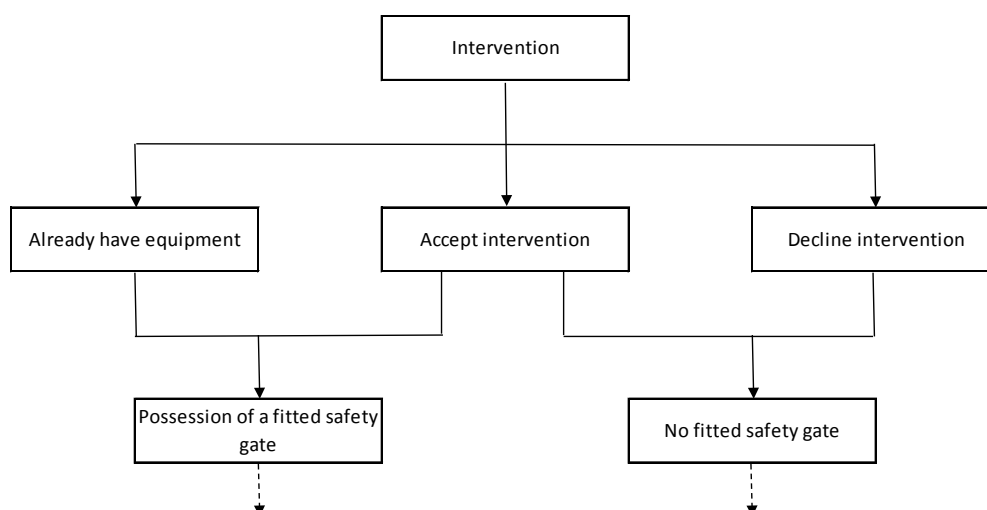
Figure 7-2 Stage 1 Intervention Model

7.4.2   Stage 1: Costs

Costs of interventions are estimated from available UK data and expert opinion (Curtis 2012). Administration costs of an intervention programme are estimated using the costs of a smoke alarm giveaway program and used in the other published cost-utility analyses in this area and are given in Table 7-3 (DiGuiseppi, Slater et al. 1999, Saramago, Cooper et al. 2014).

Table 7-2 General base-case model inputs - transition probabilities stage 1

	Point estimate (Standard Error or 95% Credibility Interval)	Parameter distribution	Source of information
STAGE 1: INTERVENTION MODEL			
Probabilities of possessing a fitted safety gate following each intervention:			
(1) Usual care	0.64 (0.60 to 0.68)		
(2) Education	0.73 (0.56 to 0.86)		
(3) Education + free or low cost equipment	0.75 (0.56 to 0.88)		
(4) Education + free or low cost equipment + home safety inspection	0.72(0.46 to 0.89)	Posterior distribution is inputted directly from NMA analysis	Network meta-analysis to evaluate the effectiveness of interventions to prevent falls in children under age 5 years (Hubbard, Cooper et al. 2015)
(5) Education + free or low cost equipment + fitting	0.75 (0.49 to 0.91)		
(6) Education + home safety inspection	0.74 (0.33 to 0.95)		
(7) Education + free or low cost equipment + fitting + home safety inspection	0.93 (0.75 to 0.98)		
Safety gate			
Baseline probability that a household has a fitted safety gate	0.56 (0.43 to 0.68)	Normal on logit scale	M-A of baseline data and control groups from NMA studies with usual care in control arm.
Probability of accepting intervention	0.76 (0.53 to 0.90)	Normal on logit scale	M-A of participation rates recorded in NMA studies.
No of children per household	1	Assumption	

Total costs per household of the seven interventions informing the decision model are given in Table 7-3 and Table 7-4. The interventions varied considerably between the different studies in the NMA making the costs difficult to estimate, for example education could be face to face or just a leaflet, equipment could be one or two safety gates, free or low cost, and often interventions are aimed at preventing more than one type of home accident not just safety gates for preventing falls down stairs. Sensitivity analyses are used to investigate the effects of changing the costs of the intervention (Sensitivity analyses 1-3 outlined in Table 7-10) and results discussed in section 7.10.2.

Costs are calculated for both the household level and for the cost of running an intervention programme.

The cost of a home safety education programme for a household is based on 5 minutes of a health visitor's time during a visit to the clinic or as part of a routine home visit given by PSSRU Unit Costs of Health and Social Care 2012 (Curtis 2012) as £44 per hour. In the base case it is assumed that a household had two safety gates (top and bottom of stairs) as this is what is recommended by NHS Choices (NHS Choices 2016) although many studies gave a discount to the household to purchase low cost equipment or may have only included a single gate (these are considered in sensitivity analyses). The cost of a safety gate is taken from the NICE PH30 costing tool as £18 per safety gate (NICE 2010a) and expert advice is sought on the cost of installation from Groundwork Creswell who quoted 18 minutes of a fitter's time at £24.93 per hour in September 2014. Home safety inspections are costed at 5 minutes of a health visitor's time given by PSSRU 2012 (Curtis 2012) as £44 per hour assuming that the inspection is part of a routine visit. Administration costs of an intervention programme are estimated using the costs of a smoke alarm giveaway program and used in the other published cost-utility analyses in this area and are given in Table 7-3 (DiGuiseppi, Slater et al. 1999, Saramago, Cooper et al. 2014).

Total costs per household of the seven interventions considered in the decision model are given in Table 7-3 and Table 7-4.

7.5   Stage 2: Child aged 0-3

Stage 2 is a Markov state-transition model that estimates the costs and QALYs after the intervention strategies, aimed at increasing possession of fitted safety gates over the first three years of life (child aged 0-3), have been implemented and uses the output from the intervention model as the primary input to determine whether the child enters stage 2 as being in a household that possesses fitted safety gates or not (S1 and S2 in Figure 7-3).

Table 7-3 Base-case model inputs for costs for stage 1 (updated to 2012 prices)

	Point estimate (Standard Error or 95% Credibility Interval)	Parameter distribution	Source of information
STAGE 1: INTERVENTION MODEL			
Intervention costs per household determined as part of an intervention programme			
Cost of home safety inspection based on cost of health visitor for 5 minutes of their time during a routine visit	£44/hour, thus 5min = £3.67	Fixed	PSSRU 2012 (Curtis 2012)
Cost of safety equipment (safety gates x2)	£38.30	Fixed	NICE PH30 Costing template 2010 (£18 per safety gate) updated to 2012 prices (NICE 2010a)
Cost of installation	18 minutes to fit a safety gate at a cost of £24.93 per hour=£7.48	Fixed	Personal communication with Gary Smith, from Groundwork Cresswell, Ashfield and Mansfield Limited on 29/09/14)
Cost of providing education programme per household accepting intervention - based on cost of health visitor for 5 minutes of their time during a routine visit	Assuming £44/hour, thus 5min = £3.67	Fixed	Assumption (based on PSSRU 2012) (Curtis 2012)
Cost of travel (time and travel) when intervention is provided in the home	£5	Fixed	Nottingham home safety scheme hourly rate including on costs and vehicle costs is £25 (estimated through personal communication with Gary Smith from Groundwork Cresswell, Ashfield and Mansfield Limited on 29/09/14) to install 5 items of safety equipment. 20% of hourly rate is allocated to safety gates.
Fixed cost of an intervention scheme – programme coordination	Considering a simulated cohort of 100,000 households: £79,529	Fixed	(DiGuiseppi, Slater et al. 1999) – updated to 2012 prices
Additional cost administrative incurred for each household that accept intervention	Distribution costs divided by the number of households in the cohort and updated to 2012 prices = £0.46	Fixed	(DiGuiseppi, Slater et al. 1999) – updated to 2012 prices

Table 7-4 Costs of the interventions per household, base case

Intervention	Costs base case
(1) Usual care	£0.00
(2) Education (face to face)	£3.67
(3) Education (face to face) + 2 free safety gates	£39.67
(4) Education (face to face) + 2 free safety gates + home safety inspection	£43.34
(5) Education (face to face) + 2 free safety gates + fitting	£47.15
(6) Education (face to face) + home safety inspection	£7.34
(7) Education (face to face) + 2 free safety gates + fitting + home safety inspection	£50.82

At the end of a cycle (one year) the child will be in one of six different states S1-S6 with or without possession of a fitted safety gate (Figure 7-3). During the year cycle they may have had a minor, moderate or severe fall injury and may have recovered to states S1 and S2, may have been left with a long-term permanent injury/disability, states S3 and S4, or may have died from a fatal fall injury, state S5, an absorbing state from which the child cannot move. The child may also have died from another cause, state S6, also an absorbing state. The transition probability matrix for stage 2 is shown in Table 7-5.
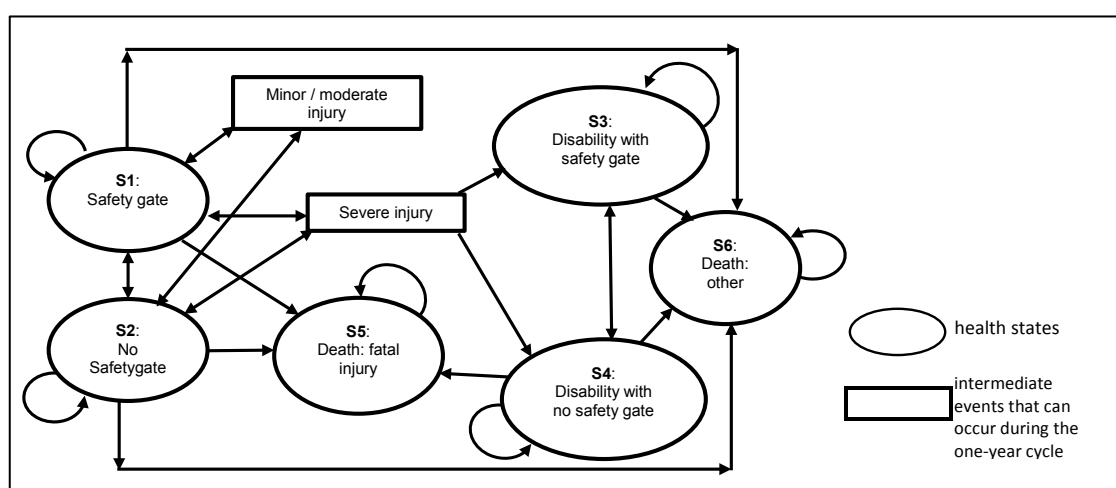
Figure 7-3 Stage 2 Markov model for child aged 0-3 years

Table 7-5 Transition probability matrix for Stage 2 of the decision model

		To					
	State	S1	S2	S3	S4	S5	S6
From	S1	✓	✓	✓	✓	✓	✓
	S2	✓	✓	✓	✓	✓	✓
	S3	✗	✗	✓	✓	✗	✓
	S4	✗	✗	✓	✓	✗	✓
	S5	✗	✗	✗	✗	✓	✗
	S6	✗	✗	✗	✗	✗	✓

7.5.1 Stage 2 transition probabilities

Figure 7-4 shows the decision tree demonstrating the transition to each Markov state in Stage 2 of the model. Within a single cycle, transition probabilities are based on published UK based evidence where available and, where evidence is not available, on expert advice and opinion (Table 7-6). Many assumptions are made and are discussed below and in the discussion section. The sensitivity of some of the assumptions made is tested in sensitivity analyses.

A child can only have one fall injury per year but can have falls in all the three years with the same probability. After having a fall the household can remain in the same state in terms of possessing or not possessing fitted safety gates or return to a different safety gates state, i.e. safety gates may be installed after a stairway fall in a household with no safety gates or the safety gates may be removed after a stairway fall in a household with safety gates. The probability of a fall injury in the whole population of 0-3 year olds is estimated using the number of falls injuries (and confidence interval) in children aged 0-4 in the UK from the 2002 Home and Accident Surveillance System (HASS) (through personal communication with Helen Shaw, RoSPA, 1 May 2014) and using the 2001 census UK population of children aged 0-4 (Office for National Statistics). Numbers were not available for 0-3 year olds. The probability is adjusted by whether the household possessed fitted safety gates or not using the results from the KCS case-control study that investigated interventions effective in protecting against stairway falls (Kendrick, Zou et al. 2015). In the KCS

case-control study, cases of stairway falls in children aged 0-4 are matched to community controls and the use of safety gates on the stairs recorded in the month prior to the stairway fall injury or completing the control group questionnaire. Compared to controls in the adjusted analysis, parents of cases are significantly more likely to have no stair safety gates (OR= 2.50, 95%CI 1.90, 3.29) (Table 7-7) than to have a closed stair safety gate.
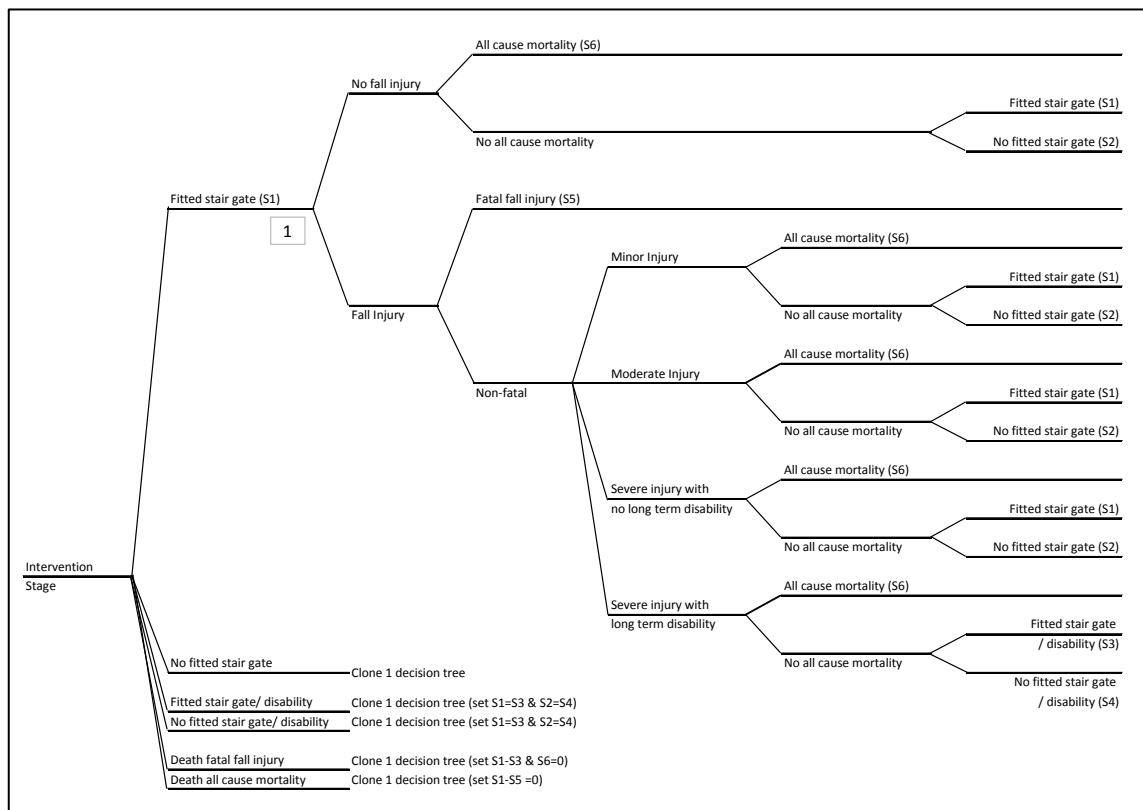
Figure 7-4 Decision tree underlying the Markov stage 2 model

Table 7-6 General base-case model inputs - transition probabilities stage 2 and 3

	Point estimate (Standard Error or 95% Credibility Interval)	Parameter distribution	Source of information
STAGES 2 and 3: Age 0-2 and LONG-TERM (Age 3+) MODEL			
Number of falls in children aged 0-4	mean=41,246 (se=84.28)	Normal	HASS(2002) (extracts from the Department of Trade and Industry's Home and Leisure Accident Surveillance System (HASS/LASS), Helen Shaw, RoSPA, 1 May 2014, personal communication). Stairway falls UK 2002 in child aged 0-4. Lower limit 41,081 & upper limit 41,411 for number of falls
Probability of fall	Number of falls/3,486,469		2001 census population for child aged 0-4 UK (Office for National Statistics)
Relative risk of fall down stairs when safety gate is in use vs no safety gate	Ln(OR) = 0.916 se(Ln(OR))=0.14 (equation (7.1))	Normal	Stairway falls case-control study 2014: cases vs community controlled adjusted analysis OR (for "Did not use safety gate" vs "closed safety gate")=2.50 (95% CI 1.90 to 3.29).
Probability of using emergency ambulance	0.242	Fixed	Hospital Episode Statistics (Health and Social Care Information Centre 2013) 24.2% of all cases arrived by emergency transfer (ambulance/helicopter). Used for all severities of injuries.
Probability of mild falls injury (attends ED but not admitted)	2604/2724=0.9560		HASS 2002 (extracts from the Department of Trade and Industry's Home and Leisure Accident Surveillance System (HASS/LASS), Helen Shaw, RoSPA, 1 May 2014, personal communication)
Probability of moderate falls injury (attends ED & admitted <2 days)	88/2724=0.0323	Multinomial for all severity of injuries	
Probability of severe falls injury (attends ED & admitted ≥2 days) but not long-term disability	(32 – 2)/2724 =0.0110		Severe injuries with estimated number with long term disability subtracted, i.e. 0.000652* 2724 = 1.78 ≈ 2
Probability of severe falls injury (attends ED & admitted ≥2 days) and has a long-term disability	0.000652		The Economic Burden of Injury in Canada 2004 (SMARTRISK 2009)
Probability of fatal falls injury	0.000000163		England and Wales mortality statistics: four stairway deaths in those aged 0–4 years in 2002–12, average of 0.57 per year ; n = 3,496,750 children aged 0–4 year olds in 2011 census (Office for National Statistics 2013a, Office for National Statistics 2013b)
Probability the household keeps the safety gate already in place after fall	0.95	Uniform (0.9,1)	Assumption
Probability the household remains in the no safety gate arm after fall	0.56	Uniform (0.5, 0.62)	Based on (Morrongiello, Howard et al. 2009)

Table 7-7 Results from the Case-Control Study to investigate interventions effective in protecting against stairway falls (Kendrick, Zou et al. 2015)

Exposures	Cases	Controls	Adjusted OR (95% CI)	Confounders adjusted for
Safety gate closed	174(29.7)	1245(51.1)	1 [Ref]	IMD, distance from hospital, HADS (hospital anxiety and depression scale), PDH (parenting daily hassles scale), first child, stairs safety, hours out-of-home child care
Safety gate left open	210(35.9)	555(22.8)	3.09(2.39,4.00)	
No stair safety gate	201(34.4)	636(26.1)	2.50 (1.90, 3.29)	

Hence the probability of a stairway fall ($fall$) in the group with fitted safety gates (sg), p($fall|sg$), is calculated using equation (7.1).

$$p(fall) = p(fall|sg) \times p(sg) + p(fall|no\ sg) \times p(no\ sg)$$

$$= p(fall|sg) \times p(sg) + p(fall|sg) \times OR \times p(no\ sg)$$

and hence $\quad p(fall|sg) = p(fall) \Big/ (p(sg) + OR \times p(no\ sg))$ $\qquad$ (7.1)

$$=0.01183/0.56+2.50 \text{ x } 0.44 = 0.007$$

where

$p(fall) =$ probability of a fall down stairs for child $0-3$

$$= \text{HASS UK falls } 2002 \Big/ 2001 \text{ census population for UK}$$
$$= {41,246} \Big/ {3,486,469} = 0.01183$$

$OR = 2.50$ is the odds ratio from the case-control study and used to estimate the relative risk of a fall in the community controls compared to the falls cases (Kendrick, Zou et al. 2015)

$p(sg) = 0.56$ is the estimate from the meta-analysis of studies giving the baseline or usual care possession of fitted safety gates (section 7.4.1).

$p(no\ sg) = 1 - p(sg) = 0.44$

Uncertainty around these estimates is incorporated in the model by expressing the number of falls and the odds ratio (on the log scale) as normal distributions (Table 7-6).

Falls injuries are defined as: mild with a reported injury but outpatient only; moderate with a reported injury requiring admittance to hospital for observation and minor

160

treatment; severe requiring admittance to hospital for two or more nights; severe leading to a permanent disability; fatal (Figure 7-4). The estimates of the probabilities are given in Table 7-6. The probability of a fatal injury is determined by taking the average yearly number of stairway deaths in children aged 0-4 from the England and Wales mortality statistics 2002-2012 out of the population estimate of 0-4 year olds from the 2011 census estimate for England and Wales; there are a very low number of stairway falls deaths per year (in some years no stairway falls deaths) (Office for National Statistics 2013b). The probabilities for the other types of injuries are estimated from the HASS 2002 figures for 0-4 year old victims of home accidents involving a fall on stairs/steps by the length of inpatient stay (through personal communication with Helen Shaw, RoSPA, 1 May 2014). The probability of an injury resulting in a permanent disability is taken from the Canadian SMARTRISK report "The Economic Burden of Injury in Canada" (SMARTRISK 2009) that gave rates of 71.6 male children aged 0-4 per 100,000 and 58.5 female children aged 0-4 per 100,000 partial or total permanent disabilities in 2004.

Uncertainty around the estimates for fall severity (not fatal) is incorporated in the model by using a multinomial distribution for the number of each type of fall severity out of the total recorded number of falls in the HASS data for 2002 with a Dirichlet prior (non-informative with alpha=1) for the multinomial probabilities (equation (7.2)).

$$numberinjury_i \sim multinomial(p(injuryseverity)_i, \ totalinjuries)$$

$$p(injuryseverity)_i \sim Dirichlet(\alpha = 1)$$

(7.2)

where

$i =$ mild, moderate, severe recovers & severe permanent disability injury
$numberinjury_i =$ number with injury severity i
$p(injuryseverity) =$ probability of injury type i
$totalinjuries =$ total number with the four severities of injury

At the end of each cycle of the stage 2 model, the household can change from possession of fitted safety gates to no possession if there is a fall injury, and vice versa. Most households are assumed to retain possession of fitted safety gates and the probability that a household moves to possession from no possession after a fall injury

is based on Morrongiello 2008 who suggested that injury to children can evoke positive changes in parental beliefs about injuries and their preventability (Morrongiello, Matheis 2007). Probabilities were estimated (no evidence available) and uncertainty around the estimates incorporated into the model using a uniform distribution as follows:

P(moving from no fitted safety gates to possession of fitted safety gates state)~unif(0.9,1)
P(moving from possession of fitted safety gates to no fitted safety gates state)~unif(0.5,0.62)

7.5.2   Stage 2 costs

Costs have to be estimated for each of the severities of injuries, these are summarised in Table 7-8. Hospital Episode Statistics for England, Accident and Emergency (A&E) show that ambulances are estimated to attend 24.2% of the fall injuries (Health and Social Care Information Centre 2013), regardless of severity (this assumptions is investigated in a sensitivity analysis in section 7.10.2), and the ambulance and treatment costs are estimated using the median and interquartile range (£263 (IQR: £248, £277) from the PSSRU Unit Costs of Health and Social Care 2012 (Curtis 2012). The interquartile range is converted to an approximate standard deviation (sd=21.48). For a minor stairway fall injury the cost is estimated as the cost of emergency department treatment not leading to a hospital inpatient stay (£112 (sd=27.46)), for a moderate injury the cost of emergency department treatment leading to hospital inpatient stay (£146 (sd=42.22)) plus the cost of a non-elective short(<2 days) admission and for a severe injury the cost of emergency department treatment leading to hospital inpatient stay (£146 (sd=42.22)) plus the cost of a non-elective long (>2 days) admission. Uncertainty around these costs is incorporated in the model using gamma distributions (section 6.2.3). For a fatal fall the cost (£205.50) is taken from the estimated cost reported for smoke alarm fatalities (Ginnelly, Sculpher et al. 2005) which includes the coroner's cost and autopsy.

Table 7-8 Base-case model inputs for costs in stages 2 and 3 (updated to 2012 prices)

	Point estimate (Standard Error or 95% Credibility Interval)	Parameter distribution	Source of information
STAGES 2 and 3: PRE-SCHOOL and LONG-TERM MODEL			
Cost of emergency transfers included for 25.4% of all falls injuries	£263 (£21.48)	Log Normal	PSSRU (Curtis 2012)
Cost of emergency department treatment of cases not leading to hospital inpatient stay (minor injury)	£112 (£27.46)	Log Normal	PSSRU (Curtis 2012)
Cost of emergency department treatment for cases leading to hospital inpatient stay (moderate or severe injury)	£146 (£42.22)	Log Normal	PSSRU (Curtis 2012)
Cost of a non-elective short (<2 days) inpatient admission	£586 (£223.70)	Log Normal	PSSRU (Curtis 2012)
Cost of a non-elective long (≥2days) inpatient admission	£2461 (£810.37)	Log Normal	PSSRU (Curtis 2012)
Annual cost of chronic ill-health	£380.30 (£98.44)	Gamma	HALO study report from the Medical Care Research Unit (J Nicholl, personal communication)
Cost of fatal injury	£205.50	Fixed	Ginelly (2005) reported in functional smoke alarm model (Saramago *et al.* 2014)

7.6 Stage 3: Long-term model, child aged 3+

Stage 3 of the model follows the child from aged 3 through adulthood until death or aged 100 years (Figure 7-5). The model therefore accounts for any lifetime effects of stairway falls injuries. Children enter the model in one of four states: well; disabled; dead from a fatal fall injury; or dead from other causes. Stairway falls are likely to take place beyond the age of three but the possession of a safety gates is less likely to prevent the fall so the intervention is no longer assumed to have an effect and the intervention groups are assumed equal.
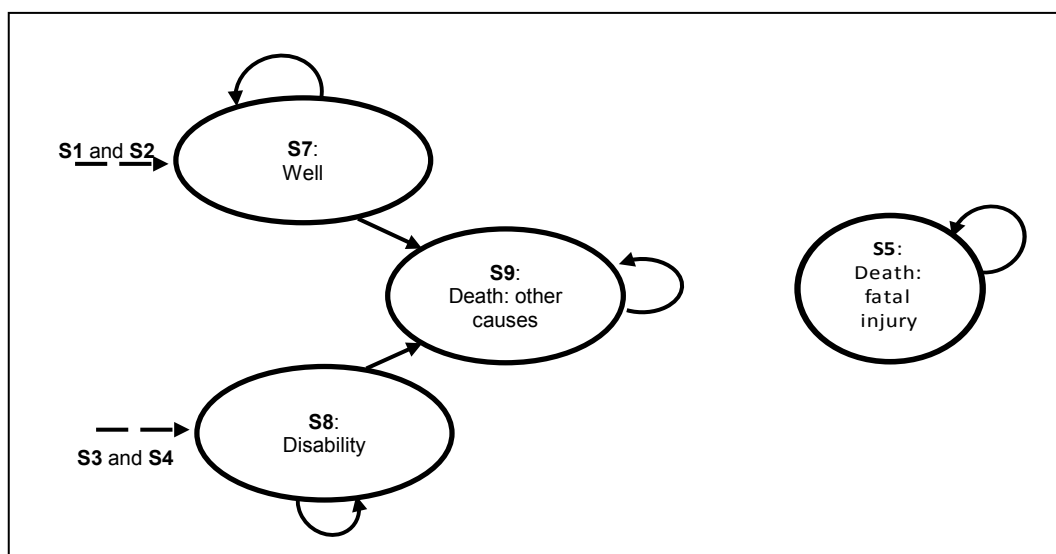
Figure 7-5 Long term Markov model age 3+

7.6.1 Stage 3 costs

Permanent disabilities suffered due to a stairway fall aged 0-3 are assumed to affect the child throughout their life and hence incur costs throughout. To determine the long-term costs the mean yearly follow-up costs to the NHS for the 580 survey responders (2009) is reported to be £342 (95% CI: £192 to £539) in 2007/08 prices in the HALO study report from the Medical Care Research Unit (J Nicholl, personal communication). Bootstrapped confidence intervals are given and, as they are not symmetrical, an average of the upper and lower estimates of the standard error is calculated. The estimates are used in a gamma distribution to incorporate the uncertainty. Costs are updated to 2012 prices and are summarised in Table 7-8.

7.7 Utilities

Base case utilities are taken from a nationally representative survey of 3395 UK population of men and women aged 18 and over (Kind, Hardman et al. 1999) and as no baseline information is available for under 18's it is assumed that the utilities are the same as for the 18-25 year olds. These values, with other utilities used, are summarised in Table 7-9 and will be used to represent the population with no fall

injury. It is difficult to find values for the utility decrements for each severity of injury. Miller et al 2000 gave a utility decrement of 0.10 (QALY's lost per case) for all falls injuries in children aged 0-19 (Miller, Romano et al. 2000) and Brussoni et al 2013 investigated the reliability of the EQ-5D-3L among a paediatric injury population of all injuries aged 0-16 (Brussoni, Kruse et al. 2013). Using this information the utility deficits for a moderate injury is fixed at 0.10 based on the Miller 2000 figure. Uniform distributions are used for minor and severe injuries as follows:

Minor injury utility deficit ~ Uniform(0, 0.1)

Severe injury utility deficit ~ Uniform (0.1, 0.3)

The impact of these assumptions is investigated through sensitivity analyses 6 and 7 (Table 7-10).

The mean utility deficit for disability, 0.16 (s.e.=0.025), is taken from the HALO study report from the Medical Care Research Unit (J Nicholl, personal communication) and uncertainty is incorporated using a beta distribution (equation (6.1)).

7.8   Model implementation

The model is implemented in WinBUGS 1.4.3 (Spiegelhalter, Thomas et al. 2003) and parameters estimated using Markov Chain Monte Carlo (MCMC) simulation (WinBUGS code is given in Appendix Q). A burn-in period of 10,000 iterations is followed by 10,000 iterations using three chains with different starting values to give estimates and check model convergence (section 4.1). Parameter uncertainty is accounted for in the model by defining a probability distribution for each parameter where possible. Mean costs and mean QALYs are calculated by averaging across the 10,000 MCMC simulations.

Table 7-9 Base-case model inputs for quality of life weights

	Point estimate (Standard Error or 95% Credibility Interval)	Parameter distribution	Source of information
Utility parameters per cycle			
Utility deficit for minor injury	0.05	Uniform(0,0.1)	Assumption – half moderate utility deficit
Utility deficit for moderate injury	0.10	Fixed	Utility decrement 0.10 for falls injury ages 0-4 (Miller, Romano et al. 2000, Brussoni, Kruse et al. 2013) looked at all injuries aged 0-16, one month change in EQ-5D-3L.
Utility deficit for severe injury	0.20	Uniform(0.1,0.3)	Assumption – double moderate and long-term disability
Utility deficit associated with disability per year	0.10 (SE=0.025)	Beta	Medical Care Research Unit (J Nicholl, personal communication) – updated to 2012 prices
General background utilities for non-injured population	< 25yrs 0.94 (sd=0.12) 25-34yrs 0.93 (sd=0.15) 35-44yrs 0.91 (sd=0.16) 45-54yrs 0.85 (sd=0.25) 55-64yrs 0.80 (sd=0.26) 65-74yrs 0.78 (sd=0.26) >75yrs 0.73 (sd=0.27)	Normal	UK Population Norms (Kind, Hardman et al. 1999)

7.9   Sensitivity analysis

There are many model assumptions and different data sources so sensitivity analyses are performed to ensure that the results are robust and also to determine if any interventions are more cost-effective for different participant groups. The sensitivity analyses conducted are described in Table 7-10. A burn-in period of 10,000 iterations is followed by 10,000 iterations, the same as the original analysis.

7.10 Results

7.10.1 Base-case analysis

The findings from the base case of the decision analysis evaluating the cost effectiveness of different interventions to increase possession of fitted safety gates to prevent stairway falls are described below.

Table 7-10 Decision model sensitivity analyses

SA1	Reduce the number of safety gates from two (costing £38.30) to one costing £19.15.
SA2	Cost of education changed from £6.66 (based on 20 minutes of a local authority workers time) to £0.56 (cost of home safety information pack per family reported in the Safe At Home Project report, 2011).
SA3	Reduce the cost of safety gate from £19.14 to £14.14 under the low cost equipment giveaway (voucher for a £5 safety gate).
SA4	Fixed costs of intervention reduced to £40,000
SA5	Change baseline possession of a safety gate to 0.44 from 0.56 reflecting the Watson et al study aimed at households in a deprived area (Watson, Kendrick et al. 2005)
SA6	Changing utility deficits to 0.07, 0.19 and 0.34 for mild, moderate and severe injuries respectively to reflect Brussoni et al
SA7	Removing uncertainty in utility deficits.
SA8	Increase the number of children in a household from 1 to 1.8.

In the base-case analysis seven interventions are evaluated (Table 7-11), of which education (E) had the lowest estimated ICER when compared to usual care (UC) with £284,068 per QALY gained. Four of the seven interventions had higher costs or higher ICERs than more effective interventions, namely education + free or low cost safety equipment (E+FE), education + free or low cost safety equipment + home safety inspection (E+FE+HSI), education + free or low cost safety equipment + fitting of equipment (E+FE+F) and education + home safety inspection (E+HSI).

Figure 7-6 shows the probability of the alternative interventions being cost effective for a range of willingness to pay thresholds. At a threshold value of £30,000 per QALY gained, the NICE threshold, usual care has the highest probability of being cost effective (0.999). In fact, for all thresholds up to £100,000 none of the other interventions are cost-effective. For each of the interventions, 5,000 simulated samples of the incremental costs and incremental QALYs compared to usual care were plotted on a cost-effectiveness plane (Figure 7-7). The ICERs lie predominantly in the north-east quadrant which suggests that all of the interventions compared to usual care are more costly but also more effective than usual care.

It can be seen from Table 7-11 and Figure 7-7 that the main driver of the of cost-effectiveness was the cost of providing the interventions; the interventions all produced similar gains in QALYs but differed in terms of the incremental costs with the most intensive interventions being the most costly.

The number of households in each state at the end of the first two stages in the decision model (Figure 7-1) are shown in Appendix Q. There are very low numbers of severe accidents causing permanent disability and deaths from falls down stairs.
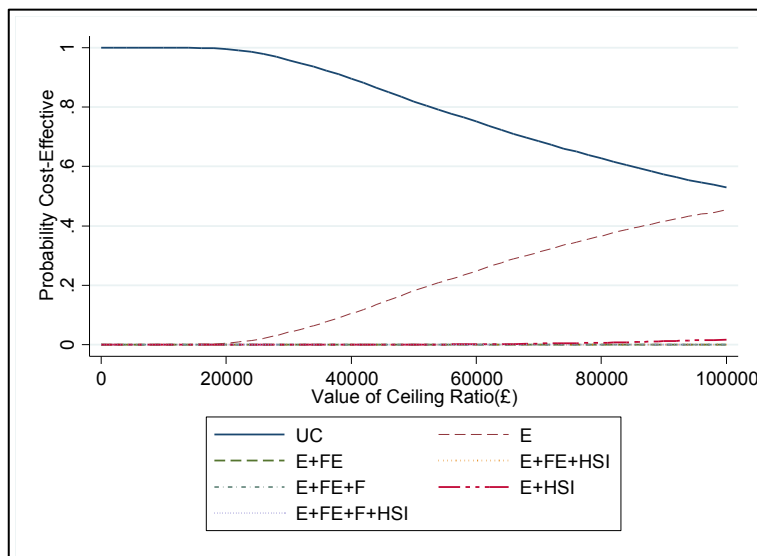
Figure 7-6 Cost-effectiveness acceptability curves for the base case analysis indicating the probability that each intervention is the most cost-effective for a range of willingness-to-pay ratios.

UC- usual care, E- education, FE – free/low cost equipment, F- fitting, HSI – home safety inspection
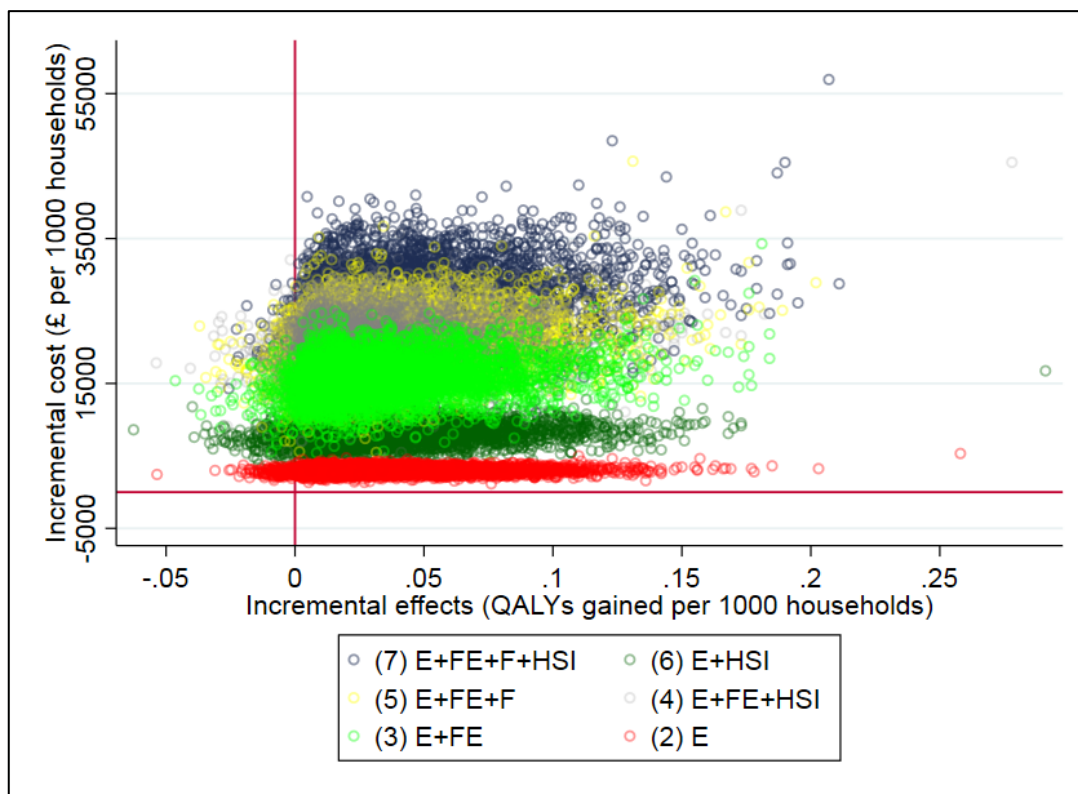
Figure 7-7 Cost-effectiveness plane showing simulated ICERs for each intervention compared to usual care

E- education, FE – free/low cost equipment, F- fitting, HSI – home safety inspection

7.10.2 Sensitivity analysis

A range of sensitivity analyses varying the base-case assumptions and inputs are implemented (Table 7-10) and the results are presented in Table 7-12. All assessed the probability of interventions being cost effective at a threshold of £30,000. The results are not very sensitive to the changes although the cost of the most intensive intervention (7) is reduced in most sensitivity analyses.

Table 7-11 Base case cost-effectiveness results for safety gates to prevent stairway falls (per 1,000 households)

Intervention	Expected QALYs	Expected Costs (£s)	Incremental QALYs	Incremental Costs (£s)	ICER (£s per QALY)	Probability CE (£30,000)	Probability CE (£50,000)
(1) UC	25,056.326 (25039.202 to 25073.452)	3431 (2446 to 4826)				0.999	0.999
(2) E	25,056.334 (25039.207 to 25073.460)	5529 (4543 to 6859)	0.007	2089	284,068	<0.001	0.001
(3) E + FE	25,056.334 (25039.209 to 25073.462)	18,358 (13,338 to 23,472)			Extendedly dominated	<0.001	<0.001
(4) E + FE + HSI	25,056.334 (25039.211 to 25073.458)	21,252 (15,203 to 27,432)			Dominated	<0.001	<0.001
(5) E + FE + F	25,056.334 (25039.207 to 25073.462)	25,017 (17,621 to 32,589)			Dominated	<0.001	<0.001
(6) E + HSI	25,056.334 (25039.209 to 25073.458)	8454 (6803 to 10240)			Dominated	<0.001	<0.001
(7) E + FE + F + HSI	25,056.335 (25039.212 to 25073.462)	26,227 (18,409 to 34,246)	0.009	22,745	2,405,800	<0.001	<0.001

Data are expected QALY (95% credibility interval) and expected costs (95% credibility interval) per 1,000 households. (1) UC = usual care; (2) E = education; (3) E + FE = education + low cost/free equipment; (4) E + FE + F = education + low cost/free equipment + fitting; (5) E+ FE + F = education + low cost/free equipment + Fitting; (6) E + HSI = education + home safety inspection; (7) E + FE + F + HSI = education + low cost/free equipment + home safety inspection + Fitting. Probability CE = probability that intervention is cost effective at a £30,000/£50,000 threshold value. QALYs = quality-adjusted life years.

Dominated = costs more but delivers less QALYs. Extended dominance = ICER greater than that of a more effective intervention

Table 7-12 Sensitivity analyses (SA) results for safety gates to prevent stairway falls (per 1,000 households)

	Expected QALYs	Expected Costs (£s)	Incremental QALYs	Incremental Costs (£s)	ICER (£s per QALY)	Probability CE (£30,000)	Probability CE (£50,000)
SA1: number of safety gates reduced from two to one							
(1) UC	25,056 (25039 to 25073)	3,428 (2446 to 4847)				0.999	0.999
(2) E	25,056 (25039 to 25073)	5,529 (4543 to 6883)	0.007	2090	283,228	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	17,361 (12683 to 22083)	0.009	13,860	1,466,433	<0.001	<0.001
SA2: reducing the cost of education by using the cost of providing a leaflet only							
(1) UC	25,056 (25040 to 25073)	3,428 (2446 to 4847)				0.996	0.961
(2) E	25,056 (25040 to 25073)	4,482 3537 to 5854)	0.007	1053	143,846	0.004	0.039
(7) E + FE + F + HSI	25,056 (25040 to 25073)	25,217 (17712 to 32842)	0.009	21,714	2,296,038	<0.001	<0.001
SA3: providing low cost (£5 voucher) rather than free safety gates							
(1) UC	25,056 (25040 to 25073)	3,428 (2446 to 4847)				0.999	0.999
(2) E	25,056 (25039 to 25073)	5,529 (4543 to 6883)	0.007	2090	283,228	<0.001	<0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	22,919 (16233 to 29678)	0.009	19,411	2,053,078	<0.001	<0.001

Table 7-12 (continued) Sensitivity analyses (SA) results for safety gates to prevent stairway falls (per 1,000 households)

	Expected QALYs	Expected Costs (£s)	Incremental QALYs	Incremental Costs (£s)	ICER (£s per QALY)	Probability CE (£30,000)	Probability CE (£50,000)
SA4: fixed costs of intervention reduced to £40,000							
(1) UC	25,056 (25039 to 25073)	3428 (2446 to 4847)				0.999	0.999
(2) E	25,056 (25039 to 25073)	5529 (4543 to 6884)	0.007	2090	157,348	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	26,252 (18372 to 34271)	0.009	22,752	1,336,429	<0.001	<0.001
SA5: changing baseline possession of safety gate to 0.44 from 0.56 to reflect deprived households (Watson et al, 2005)							
(1) UC	25,056 (25039 to 25073)	3,141 (2258 to 4428)				1	0.999
(2) E	25,056 (25039 to 25073)	5,569 (4592 to 6866)	0.008	2,436	291,812	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	31,690 (23318 to 36884)	0.011	28,522	2,612,847	<0.001	<0.001
SA6: changing utility deficits to 0.07, 0.19 and 0.34 for mild, moderate and severe injuries respectively and using a beta distribution to reflect Brussoni et al (2013)							
(1) UC	25,056 (25040 to 25073)	3,141 (2258 to 4428)				1	0.999
(2) E	25,056 (25039 to 25073)	5,569 (4591 to 6866)	0.008	2086	267,482	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	31,690 (23,318 to 36,884)	0.010	22,686	2,257,270	<0.001	<0.001

Table 7-12 (continued) Sensitivity analyses (SA) results for safety gates to prevent stairway falls (per 1,000 households)

	Expected QALYs	Expected Costs (£s)	Incremental QALYs	Incremental Costs (£s)	ICER (£s per QALY)	Probability CE (£30,000)	Probability CE (£50,000)
SA7: Removed uncertainty in utility deficits							
(1) UC	25,056 (25040 to 25073)	3,429 (2446 to 4838)				1	0.999
(2) E	25,056 (25039 to 25073)	5,22 (4546 to 6872)	0.007	2,089	285,292	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	26,218 (18320 to 34159)	0.009	22,753	2,414,228	<0.001	<0.001
SA8: number of children in household increased from 1 to 1.8							
(1) UC	25,056 (25039 to 25073)	3236 (2229 to 4685)				0.999	0.999
(2) E	25,056 (25039 to 25073)	5572 (4582 to 6866)	0.008	2,319	292,258	<0.001	0.001
(7) E + FE + F + HSI	25,056 (25039 to 25073)	29,867 (18,141 to 41,807)	0.010	26,566	2,585,853	<0.001	<0.001

Data are expected QALY (95% credibility interval) and expected costs (95% credibility interval) per 1,000 households. (1) UC = usual care; (2) E = education; (3) E + FE = education + low cost/free equipment; (4) E + FE + F = education + low cost/free equipment + fitting; (5) E+ FE + F = education + low cost/free equipment + Fitting; (6) E + HSI = education + home safety inspection; (7) E + FE + F + HSI = education + low cost/free equipment + home safety inspection + Fitting. Probability CE = probability that intervention is cost effective at a £30,000/£50,000 threshold value. QALYs = quality-adjusted life years.

7.11 Chapter summary

This chapter has covered the development, implementation and findings from a comprehensive decision model to assess the cost-effectiveness of seven interventions to increase the possession of a fitted safety gate to prevent falls down stairs in a hypothetical population of 100,000 households with one child. The model follows the child from birth through to 100 years (sections 7.2 and 7.4-7.8). The network meta-analysis, described in section 5.10, was used to inform the decision model on the effectiveness of the seven interventions. The NMA showed that the most intensive intervention that included education, free equipment, a home safety inspection and fitting was the most effective in increasing possession of a fitted safety gate.

At a threshold value of £30,000 per QALY gained, usual care has the highest probability of being cost effective (0.999). When compared to usual care, the education (E) intervention had the lowest estimated ICER when compared to usual care (UC) with £284,068 per QALY gained. Four of the seven interventions had higher costs or higher ICERs than more effective interventions so were classed as dominated. Although the most intensive intervention is most effective, in terms of the NICE guidelines, it is not cost effective with an ICER of £2,405,800 compared to usual care.

In developing this model many assumptions were made due to the lack of evidence or multiple sources of evidence (section 7.3). Some of these assumptions are assessed using the sensitivity analyses (section 7.9) but not all are fully addressed and will be discussed further in Chapter 8. The sensitivity analyses showed that the results were slightly sensitive to reducing the cost of the education intervention but did not make the intervention cost-effective. The cost of the interventions was the main driver of the cost-effectiveness analysis. There was very little difference in the QALYs between the interventions but the cost differences for 1,000 households varied considerably.

8 Discussion, recommendations and conclusions

This final chapter gives an overall summary of the findings (section 8.1), describes the strengths and limitations of the work (section 8.2) and recommends extensions (section 8.3). An overall conclusion is given in section 8.4.

8.1 Summary of findings

The findings of the research presented in this thesis will be discussed in this section and split into the intervention effectiveness and cost-effectiveness results and the methodological challenges.

8.1.1 Intervention effectiveness and cost-effectiveness for falls prevention

Sixteen studies are identified in a systematic review that reported fall prevention interventions. Chapter 5 presents the results of a pairwise random effects meta-analysis of home safety interventions for the prevention of falls injuries as part of an update of a Cochrane systematic review and meta-analysis (Kendrick, Wynn et al. 2013). There is evidence that enhanced interventions compared to a reduced/usual care intervention increase the possession of a fitted safety gate (12 studies, OR 1.61 (95% CI: 1.19, 2.17)) and decrease the possession and use of a baby walker (9 studies, OR for no possession 1.57 (95%CI: 1.18, 2.09)). There is little evidence that the enhanced intervention increased the possession of window locks, increased the use of bath mats or decals or decreased leaving a child unattended on a high surface. One study, Phelan 2010, is identified as having a high effect estimate compared to the other studies. When this study is excluded there is still a difference between the enhanced intervention and usual care arms for the possession of a fitted safety gate. There is heterogeneity between the studies in the safety gate and baby walker analyses so sub-group analyses and meta-regression are used to explore the heterogeneity. Sub-group analyses are used to explore differences in the study design; studies that provided free or low cost equipment, are administered in a clinical setting, are RCTs with adequate concealment, not blinded and less than 80% follow-up

provided evidence that the enhanced intervention is effective but there are low numbers of studies in some sub-groups. Meta-regression showed the effectiveness of the enhanced intervention for increasing possession of a fitted safety gate decreases with an increasing percentage of boys in the study. There are no other statistically significant covariate effects. There is little evidence that the baseline risk (possession of a fitted safety gate) had any effect on the intervention effect.

For 12 of the 16 studies IPD is available so the meta-regression is repeated by combining IPD and aggregate data. There is little evidence of ecological bias and the only statistically significant result is that an enhanced intervention is more effective for increasing possession of a fitted safety gate among households in rented accommodation compared to owner-occupied households (OR=1.62 (95%CrI: 1.18, 2.24)). The gender effect observed in the aggregate data analysis is no longer statistically significant. There are a high number of tests undertaken so the results need to be treated with caution.

Seven different interventions are identified combining intervention components for a NMA. The NMA found that the most intensive intervention, education plus home safety inspection plus providing and fitting low-cost/free equipment, is the most effective intervention (OR 7.80, 95% CrI 3.18 to 21.3; p(best) = 0.97, SUCRA=99%) for increasing possession of a fitted safety gate compared to usual care and it is also statistically significantly more effective than all of the other five interventions. This result is as would be expected as this intervention is a combination of all intervention components. IPD, where available, is incorporated into the NMA and the uncertainty around the estimates is reduced, the most intensive intervention is the most effective compared to usual care (OR 8.00, 95% CrI 3.32 to 19.8).

Education is the most effective intervention for reducing the number of households that do not possess or use a baby walker compared to usual care (OR=0.48, 95% CrI: 0.31 to 0.84, p(best)=0.65). Equipment is not relevant for this outcome so more intensive interventions are unlikely to increase efficacy. There is little evidence of a difference between the interventions for the use of window locks or never leaving a

child on a high surface and the network is not connected for the use of bath mat outcome.

In Chapter 6 the NMA results are used to inform a cost-utility model to estimate the mean costs and quality adjusted life years (QALYS) associated with the seven interventions for increasing possession of a fitted stair safety gate to prevent falls down stairs. A simulated cohort of 100,000 UK households with a new-born are followed through the intervention, for the first three years of life (aged 0-3) when a safety gate is recommended, and then long-term to 100 years. Costs are from a public sector/NHS perspective.

At a threshold value of £30,000 per QALY gained usual care is the only cost-effective intervention. Usual care (*p* best = 0.999) has the highest probability of being cost-effective (at £30,000 per QALY) and education has the lowest ICER (£284,068 per QALY) of all of the interventions. The most intensive intervention, which is the most effective, has the highest costs leading to it being not cost-effective (ICER £688,772). The main driver of cost-effectiveness is the cost of providing the interventions; they all produced similar gains in QALYs but differ in terms of the incremental costs with the most intensive interventions being the most costly. Sensitivity analyses changing some of the parameters modelled did not produce any cost-effective interventions compared to usual care.

8.1.2 Methodological challenges

In Chapter 2 evaluations of the effectiveness and cost-effectiveness of public health interventions are discussed. NICE and the Cochrane Public Health Review Group both recommend a systematic review with meta-analyses, of the evidence base from primary research and previous reviews, to assess effectiveness and a health economic evaluation using QALYs to assess the cost-effectiveness of the interventions ((NICE 2012, The Cochrane Collaboration 2018)). Several methodological challenges are identified that need to be addressed, including: limited evidence on the question of interest; multiple component interventions; lack of RCT evidence; use of cluster randomised trials where clustering is not adjusted for; widely scattered evidence base across disciplines; interest in differential outcomes for different socio-economic

groups; interventions interact with the context in which they are implemented; long-term benefits need to be considered. These are all issues in the context of the prevention of falls in the home in children. A review of NICE public health evaluations concluded that more advanced methods of evidence synthesis as conducted in this thesis should be used ((Achana, Hubbard et al. 2014). One of the reasons given for not using meta-analysis is the heterogeneity of interventions and this is identified as a problem in the motivating example in this thesis. The problem with heterogeneous interventions is addressed using network meta-analysis, covariates are included to address heterogeneity between studies.

The enhanced intervention arms in the pairwise meta-analysis are heterogeneous and include various combinations of education, home safety inspection, provision of free or low-cost safety equipment and fitting of equipment. The control/usual care arms also varies across studies; usual care is the most common control intervention but some control arms receive generic safety advice or elements of the intervention, for example home safety inspection but not home safety equipment. One study, Phelan et al (2010), for which the enhanced intervention effect is much higher than the other studies, has the most intensive enhanced intervention; education, free equipment, fitting and a home safety inspection, which is much more effective than the control arm, usual care (Phelan, Khoury et al. 2010). The pairwise meta-analysis does not distinguish between the different interventions in either the enhanced arm or the usual care arm so the interventions are split into seven combinations of the components education, free/low cost equipment, home safety inspection and fitting. Even with seven interventions there is still heterogeneity and components are lumped together. Extending to nine interventions still demonstrates that the most intensive interventions is the most effective.

Network meta-analysis is used to compare more than two interventions in a network of evidence and can provide effect estimates on intervention comparisons not evidenced in the primary studies in the network. This provides more useful information for decision makers who want to know which intervention is best and assist in making decisions when there is missing comparative evidence. Interventions

can be ranked in order of effectiveness and the probability that an intervention is "best" estimated.

A Bayesian MCMC simulation approach is taken to estimate the effectiveness of interventions in the meta-analysis. Using this approach parameters are treated as random variables. Prior distributions have been specified as vague so the data dominates. Credible intervals are reported for parameter estimates and they are easier to interpret than confidence intervals. This approach can be used to fit more complex advanced meta-analysis models such as the NMA with IPD and covariates and the results can be integrated in the decision model. However, for the most complex analyses there are problems due to the low numbers of studies.

The analyses focus on the interventions to increase possession of a fitted safety gate because more studies are identified. The baby walker analysis NMA is described because it included a three-arm study that needed to be accounted for in the analysis. Other falls prevention outcomes had fewer studies which limited the analyses.

The authors of the systematic review were successful in getting the individual participant data for thirteen of the sixteen studies reporting falls prevention interventions. Meta-regression combining IPD and aggregate data is undertaken for both the pairwise model and NMA model. Getting the IPD data in the correct form took time and rules had to be determined for when to use random effects, fixed effect and whether to split the between- and within-study variability because of low numbers of studies and poor convergence.

The cost-effectiveness of interventions is determined by integrating the posterior distributions of the NMA effectiveness estimates in a comprehensive probabilistic decision model. Uncertainty around the parameters for the probabilities, costs and utilities is represented by including parameters as probability distributions. The ICER is used to determine if interventions are more effective than usual care with a threshold of £30,000 and £50,000. There is no fixed threshold value and public health interventions, such as the childhood home accident prevention interventions, are difficult to compare to clinical interventions where NICE recommend a threshold of £30,000 per QALY to determine cost-effectiveness ((NICE 2014).

In developing the decision model the main challenge is to find up-to-date and relevant evidence to inform the structure and parameters of the model. Many assumptions are made and assessed through sensitivity analyses. The three stage model is used with an intervention stage, a stage when the child is aged 0-3 when the child is at highest risk of a fall and the intervention is recommended, and a final stage when there may be costs and effects arising from any long-term falls injuries. Transition probabilities, costs and utilities are determined for each stage of the model from the best available information which is fairly limited or there are often multiple sources.

8.2   Strengths and limitations

The systematic review that identified the data used in the analyses, is undertaken just prior to the meta-analyses so included all relevant primary studies and reviews and limited bias. The quality of the included studies is variable and there is a limited number of available studies. Studies showed wide variation in terms of the content of the intervention, population size, socioeconomic background, delivery method of the intervention and follow-up period. Most interventions are implemented in high income countries and hence not generalisable to middle or low income countries. Many studies had small sample sizes and limited power. Not all studies are RCTs but sensitivity analyses showed that the results are robust when only RCT evidence is analysed. This heterogeneity could be seen as a strength in that the effect of some study and participant characteristics could be explored and the results are generalisable to a wider population.

Across the Keeping Children Safe at Home programme similar systematic reviews were undertaken to identify evidence on interventions to prevent other home accidents such as scalds and poisonings. Some studies reported on interventions to prevent multiple types of accidents, for example equipment may be provided such as cupboard locks to prevent poisonings as well as safety gates to prevent falls or include home inspections to check smoke alarms as well as stair safety. For multifaceted interventions it is not possible to determine which components are responsible for the observed effects and the injury types are analysed separately. For the falls prevention

interventions the analyses assumed that the intervention is solely to increase one specific falls prevention behaviour, e.g. possession of a fitted safety gate.

The KCS Programme systematic reviews are the first published reviews in the field of child home injury prevention to obtain and use individual participant data. IPD meta-analysis is seen as the "gold standard" particularly when covariate effects are explored. IPD meta-regression produces a more powerful and reliable analysis and analysis methods can be standardised if they differ in the reporting of the primary study results (section 4.9). Meta-regression, used to examine the impact of covariates on the intervention effectiveness estimate, can help to account for heterogeneity and be used to investigate if any interventions are more effective for specific participant groups. Decision makers are keen to target interventions to populations where they may be more effective to reduce costs. During the KCS Programme, methods were developed to include IPD in a NMA model to compare interventions to increase the use of smoke alarms along with covariates (Saramago, Sutton et al. 2012).

The results given in sections (5.9 - 5.14) are the first NMAs of interventions to prevent falls at home in childhood. The findings of NMAs are useful for policy makers, service commissioners and providers when choosing between interventions as the interventions can be ranked in order of effectiveness and comparisons between interventions that have not been evidenced can be estimated. In the pairwise meta-analysis that informed the Cochrane review update, the interventions are lumped together into the enhanced intervention and control arms. In the NMA there is still some lumping of interventions, for example some studies provided free safety equipment and others gave vouchers for low cost equipment and in some the equipment is not relevant. This is investigated in a sensitivity analysis splitting relevant and not relevant equipment and the results showed little difference between these interventions. With complex public health interventions such as this example, it is very difficult to get clear distinct interventions.

There is very little previous evidence of economic evaluations of interventions to prevent falls, or any other unintentional injuries, in the home in children under five; Pitt et al 2009 implemented a cost-utility analysis of home safety interventions aimed

to reduce unintentional injuries in children under 15 years of age (Pitt, Anderson et al. 2009). The decision analyses, undertaken in this thesis and reported in the KCS Programme report (Kendrick, Ablewhite et al. 2017), for interventions to prevent falls are the first studies of this type to evaluate home safety interventions for the prevention of these injuries in the UK. The model developed is fully probabilistic; accounting for uncertainty in the parameters for probabilities, costs and utilities.

Decision analyses are not undertaken for interventions to reduce baby walker use as more complex analyses are required to take account of the potential protective effect of walkers on some types of falls, changes in risk of walker-related falls from changes to EU standards for baby walkers, strong warnings issued to discourage the use of baby walkers in countries including the UK and USA and, in Canada in 2004, a ban on the importation and advertisement of baby walkers including modified and second hand baby walkers.

There are many limitations linked to the decision model. Difficulties in defining interventions and usual care, as described above, also apply to the decision analyses. Many assumptions are made and, although sensitivity analyses are used to assess the impact of varying these assumptions, not all assumptions are able to be investigated. There is a lack of data on accident rates; home accident statistics, HASS data, were only collected until 2002 (reference). Some of the data on utilities and costs are obtained from other countries because UK data are not available so may not be generalisable. Suitable utilities for children under 18 were not found and it has to be assumed that they had the same utility as an 18 year-old.

8.3   Extensions

8.3.1   Evidence base

The systematic review that informs the meta-analyses in this thesis, identifies studies reporting interventions to reduce childhood falls in the home prior to the end of 2010. Since this time new studies will have been conducted so the review could be updated.

Not all studies identified in the review are randomised controlled trials; there are three non-RCT and one controlled before and after study. Study quality is assessed in terms of allocation concealment, blinding, prevalence of confounders and follow-up. This information is considered in a sensitivity analysis by excluding each study and examining the effect on the effect estimates and used in a sub-group analysis of the pairwise meta-analysis. There are approaches available to combine evidence from studies demonstrating bias or from observational data that account for bias. Ibrahim et al (2000) proposed using an informative power prior distribution for the effect estimate derived from the observational data (Ibrahim, Chen 2000). An alternative is to model the potential bias in the observational studies using an extra variance component representing the bias and using the bias-adjusted estimate from the observational studies to specify a prior distribution for the intervention effect estimate in a new meta-analysis. For this method the variance is based on what is reported in empirical evidence (Welton, Ades et al. 2009, Turner, Spiegelhalter et al. 2009). RCTs can also be adjusted for risk of bias using a probability of bias model (Dias, Welton et al. 2010b). Covariate values for the studies where only aggregate data is available is recorded across the arms. The effect of imbalances between the arms on the evidence synthesis methods could be investigated by developing the above methods (Turner, Spiegelhalter et al. 2009).

There are further outputs from the KCS Programme that could be considered in the analyses, including the case-control studies for falls down stairs, falls off furniture and falls on the flat (Kendrick, Ablewhite et al. 2017).

The systematic review of NICE public health evaluations considered evaluations from 2006 to September 2012 (Achana, Hubbard et al. 2014) and found that only 9(29%) of the 39 evaluations were informed by a pairwise meta-analysis. Since the review was conducted another 33 appraisals have been published (up to June 2017); six replacing earlier appraisals in the review and seven of the appraisals have been updates. It would be interesting to see if there has been any increase in more advanced methods of evidence synthesis in use.

8.3.2 Evidence synthesis

For the pairwise meta-analysis, a covariate, baseline possession of a fitted safety gate, is included to try to reduce heterogeneity in the model. Similarly, this method can be extended to include baseline risk in a NMA. This method is more complex as many of the studies do not have a usual care/control arm, both arms have an active intervention. To account for the lack of usual care/control arm these studies are given missing values for the baseline risk and the values are assumed missing at random. Using a Bayesian approach in WinBUGS, the NMA model code can be modified to include the baseline risk and under the exchangeability assumption the missing baseline risks are estimated (Achana, Cooper et al. 2012). Models can be fitted with separate or exchangeable regression coefficients as described in section 5.12.

Many of the studies reporting interventions to prevent falls injuries reported on multiple falls injury preventions (Table 5-1), for example Phelan et al reported on possession of safety gates, reducing babywalker use, possession of window locks and use of bath mats/decals (Phelan, Khoury et al. 2010). Multiple outcomes can be modelled in a NMA framework with borrowing of strength across networks. The models for each outcome individually could be developed to extrapolate across the evidence networks to allow information sharing on the effectiveness of interventions in promoting other safety practices for the prevention of falls. This could also be extended to across different types of unintentional home injury prevention in children, for example interventions often include providing cupboard locks to prevent poisonings, smoke alarms to prevent fire injuries and thermostatic mixing valves (TMV) on hot water taps to prevent scald injuries. These are all identified in the systematic reviews conducted as part of the Keeping Children Safe at Home Programme (Kendrick, Ablewhite et al. 2017). Achana et al (2014) developed a two-stage approach: in stage 1 information is borrowed across outcomes as well as across studies through modelling the within-study and between-study correlation structure; and in stage 2, assuming the intervention effects are exchangeable between outcomes, predict effect estimates for all outcomes (Achana, Cooper et al. 2014). For outcomes where evidence is sparse or the intervention is not evidenced in a primary study effect

estimates can still be found. Poisoning prevention outcomes from the KCS Programme are used as the motivating example for this paper.

The interventions identified in the systematic review are component based; the most intensive intervention consisted of education, free safety equipment, fitting and a home safety inspection components (Phelan, Khoury et al. 2010). The analyses performed in this report does not consider separating the effects of the different components or explicitly estimate any possible interactions between components. For decision makers it may be more useful for the analysis to address the question "which intervention component(s) has the highest probability of being most effective" (Welton, Caldwell et al. 2009, Caldwell, Welton 2016). Welton et al 2009 proposed three components-based meta-regression extensions to the NMA model. The first model, an additive effects model, assumed that the effects of each component adds. The second, a two-way interaction model, allows pairs of components to have a larger or smaller effect than if they are added, and the third model, a full-interaction model, which allows for interaction between more than two components. Applying these methods will give further information on which components and combinations of components in the interventions are most effective including combinations for which there is no evidence.

All three of the above models could be extended and combined to include the IPD but may be limited by the amount of data available in this application. Further work to explore the benefit of adding IPD could be undertaken by using simulation studies

There is limited available evidence of falls injuries but during the KCS programme further data was collected which may be useful to incorporate in the analyses. Surrogate endpoint models could be developed if there is direct evidence between safety practices and injury data. These models will simultaneously model both uptake and injury rate analyses to allow the latter to borrow strength from the former (because injury rates is what is primarily of interest - but the majority of the data is on uptake and is used as a surrogate for injury rates) (Daniels, Hughes 1997, Bujkiewicz, Thompson et al. 2017).

WinBUGS is used in most analyses to perform MCMC simulation. A Bayesian approach to NMA can also be implemented in R using the GeMTC command and calling the JAGS (Just another Gibbs sampler) software from R (Valkenhoef, Lu et al. 2012). Frequentist (classical) approaches have recently been developed for the NMA methods described, and so the analyses could be replicated using R (netmeta (Rücker, Schwarzer 2016)) and Stata (mvmeta (White 2009)). The development of these methods means that NMA methods are more accessible to those who do not have specialist knowledge of WinBUGS but there is still limited assessment of model fit using the frequentist methods.

8.3.3   Decision modelling

Many assumptions are made when designing the structure of the decision model and setting the parameter estimates and distributions in the decision model. The sensitivity analyses could be extended to investigate these in more detail however the cost-effectiveness is highly dependent on the costs so the impact of changing the costs would be of most interest. The accident rates might be slightly underestimated because the cycle in Stage 2 is a year and a child could only have one fall accident per year; it has been shown in the literature that some children are likely to be admitted to hospital on multiple occasions so the cycle could be reduced to a shorter time period (Sellar, Ferguson et al. 1991). The costs and accident estimates are based on 2012 values where possible; this could be updated further but many of the estimates of accident numbers are based on the 2002 HASS data that has not been updated. All costs and benefits are estimated from a NHS and PSS perspective as recommended by NICE (NICE 2012), so other perspectives could be taken into account such as also including out-of-pocket expenses such as time taken off work by parents and its effect on the household and wider economy. Cooper et al (2016) published the short-term costs of falls, poisonings and scalds occurring in under 5 year olds in England results from a multicentre longitudinal study (Cooper, Kendrick et al. 2016). These costs include NHS costs in addition to the first hospital visit, including GP and Health Visitor appointments, outpatient visits, prescribed medication and non-NHS costs such as time off work, travel costs, childcare and purchased aids. The utility data could be

explored further in terms of availability and the distributions used for the utility parameters.

The results from the network meta-analysis including IPD and covariates could have been incorporated into the decision model. The effect estimates from the IPD analyses have reduced uncertainty. For the decision model the benefit of adding IPD could be explored. By including the results split by covariate value, the cost-effectiveness could be evaluated further by sub-groups in the population, for example it may be more effective or cost-effective to target an intervention to households with low socioeconomic status. The covariates had little effect on the effect of the intervention effectiveness so this may not provide any further information for this example.

The decision model assumes that interventions are aimed at preventing only one type of injury. Many of the interventions evidenced in the studies are aimed at preventing a range of injuries, for example some home safety equipment schemes fitted equipment to reduce the risk of falls, poisonings and scalds; travel costs and the cost of safety equipment fitters' time will be counted in all the separate decision models and overestimate costs. More complex decision analyses could be developed to incorporate costs and benefits across multiple interventions and injury types.

Given the uncertainty in any cost-effectiveness evaluation there is always a chance that the wrong decision will be made. This will have costs associated with it in terms of health benefit and resources used. Therefore a further decision is whether, in order to reduce the decision uncertainty, more research should be undertaken. The decision makers/providers of the interventions have to make decisions based on what they expect to happen given the best available evidence. A hypothetical perfectly informed decision maker would always choose the intervention that provides the greatest net benefit, but in the real world they risk making the wrong decision, referred to as the risk of decision uncertainty. The methods of quantifying decision uncertainty and evaluating research according to its impact on decision uncertainty are referred to as value of information (VOI) analyses (Claxton, Neumann et al. 2001, Briggs, Claxton et al. 2006, Sculpher, Claxton 2005, Welton 2012).

VOI estimates the value of collecting additional data to reduce decision uncertainty and provides an indication of the optimal design for additional research to obtain these data by combining the probability and monetary consequences of an incorrect decision (Fenwick, Claxton et al. 2001). The expected value of perfect information (EVPI) can be calculated to assess the value of reducing all decision uncertainty, it is the difference between the expected net benefit (NB) given perfect information (the intervention with the higher NB in each simulation is selected and then the mean of these values calculated) and the expected NB of the current information (Edlin, McCabe et al. 2015).

Given the lack of good quality data that is often available to inform evaluations of public health interventions, value of information analysis, following the decision model, can help answer questions on whether it would be cost-effective to collect more data, which parameters should be considered for further data collection and how much data should be collected (Briggs, Claxton et al. 2006, Sculpher, Claxton 2005, Welton 2012). In this case study it would be useful to explore the value of additional information in terms of utilities.

As part of the KCS Programme, Saramago et al, 2014, included a VOI in the cost-effectiveness analysis of interventions to increase the possession of a smoke alarm and assessed the value of conducting further research on reducing decision uncertainty associated with whether to recommend a smoke alarm giveaway scheme for households with children (Saramago, Cooper et al. 2014). For the motivating example there is poor quality of information and it is unlikely that all uncertainty has been accounted for in the decision model. For this reason, along with the clear lack of cost-effectiveness of any of the interventions compared to usual care, a VOI analysis is not conducted but could be considered in future analyses.

8.4   Overall conclusions

The work in this thesis represents analyses undertaken as part of a five-year multi-centre collaborative research programme funded by the National Institute for Health

Research (NIHR). The aim is to increase the evidence-base for interventions to prevent thermal injury, falls and poisoning in the home for the under-fives. The focus of the thesis is on the effectiveness and cost-effectiveness of interventions to increase the possession of fitted safety equipment or promote good safety practices to prevent falls.

This work has illustrated that a pairwise meta-analysis that lumps together multiple intervention components does not provide the information that decision makers require regarding which specific intervention(s) to recommend. A network meta-analysis can provide more informative results by splitting 'all interventions' into specific interventions (e.g. education, education and equipment, etc.) enabling decision makers to identify the 'most' effective intervention(s). Incorporating IPD reduces the uncertainty in the effectiveness estimates and by including covariates in the models provides information on whether some interventions are more effective and cost-effective for subgroups within the population and thus should be targeted to specific populations.

Appraisals of public health interventions are usually complex and rarely informed by analyses beyond a narrative review and/or pairwise meta-analysis, despite the current NICE guidance recommendation, for both clinical and public health appraisals, that network meta-analysis, to compare trials where more than two interventions are evidenced, should be considered (NICE 2014). There is often a perceived lack of high quality evidence but, providing reviewers quality assess non-RCTs to identify well-conducted studies, meta-analyses can be conducted as demonstrated in this thesis. Heterogeneity in study designs, including interventions, outcome measures and scope, is often cited as a reason for not performing a meta-analysis but exploring the heterogeneity and attempting to account for it should be part of the analysis; although there may be circumstances where meta-analyses may not be advisable. Another reason cited is a lack of knowledge on and ability to apply more complex methods of evidence synthesis but analyses, such as network meta-analysis, are being introduced into standard statistical software packages and the NICE technical support documents referred to in this thesis (http://nicedsu.org.uk/technical-support-

documents/evidence-synthesis-tsd-series/) and a new text book (Welton 2012) provide clear advice and examples on how to use these methods.

The results from the network meta-analysis can be used to inform an economic evaluation so that the benefits and cost effectiveness of the recommendations can be evaluated as required by NICE. This ensures that model parameters are evidence based and decision makers are informed about specific interventions packages that are most effective.

By categorising interventions more finely the analyses presented have extended the existing NICE guidance on home accident prevention (Pitt, Anderson et al. 2009) to provide more informative recommendations on the most effective intervention packages rather than a combined intervention group vs a control. The results showed that the most intensive intervention is most effective in increasing the possession of fitted safety gates to prevent falls down stairs. However, not surprisingly, it is also the most costly intervention and was found not to be cost-effective at the willingness to pay threshold of £30,000. For the other interventions aimed at increasing the possession of home safety equipment and good safety practices to prevent falls there is less evidence, the equipment provided and fitted is not usually relevant and so cost-effectiveness analyses are not conducted.

The findings from this evaluation may be used, together with other analyses conducted in the KCS Programme, to identify packages of interventions to prevent multiple home accidents, such as falls, poisonings, fires, and target groups in the community where there may be health inequalities. By addressing multiple home accidents it should reduce costs and improve cost-effectiveness. A VOI analysis can also be used to evaluate if further research in this area would be worthwhile.

The findings from the NMA and decision model in this thesis have been used to inform Injury Prevention Briefings (IPB) as part of the KCS programme (Kendrick, Ablewhite et al. 2017). The evidence on the effectiveness and cost-effectiveness of home safety interventions is combined with best practice obtained from those running injury prevention programmes. The first IPB covered the prevention of fire-related injury. The second IPB was produced at the end of the KCS programme of work and covered fire-

related injuries, scalds, falls and poisonings. The final report for the KCS Programme, is promoted as an NIHR Signal and informs NICE guidance on strategies to prevent unintentional injuries among children. Local providers of home accident prevention interventions can use this information and that from suggested future research to prioritise allocation of their very limited resources.

Appendices

Appendix A WinBUGS code for random effects model and meta-regression for possession of a fitted safety gate

```
#random effects meta-analysis
model{
for(i in 1:ns){                          # ns number of studies
   delta[i,1] <- 0                        # treat effect is zero for control group
   mu[i] ~ dnorm(0,.0001)                 # vague priors for all trial baselines
   for (k in 1:2) {
           r[i,k] ~ dbin(p[i,k],n[i,k])          # binomial likelihood
           logit(p[i,k]) <- mu[i] + delta[i,k]   # model for linear predictor
           rhat[i,k] <- p[i,k] * n[i,k]          # expected value of the numerators
           #Deviance contribution
           dev[i,k] <- 2*(r[i,k]*(log(r[i,k])-log(rhat[i,k]))+(n[i,k]-r[i,k])*
                       (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
           }
   # residual deviance contribution for trial i
   resdev[i] <- sum(dev[i,])
   delta[i,2] ~ dnorm(d[2],prec)       # trial-specific LOR distributions
 }
totresdev <- sum(resdev[])            #Total Residual Deviance

d[1]<- 0                   # intervention effect is zero for control
d[2] ~ dnorm(0,.0001)      # vague prior for intervention effect logOR
or<-exp(d[2])
tau ~ dunif(0,10)          # vague prior for between-trial SD
prec <- pow(tau,-2)        # between-trial precision = (1/between-trial variance)

delta.new~dnorm(d[2],prec)                     #estimate a predictive distribution for
OR.new<-exp(delta.new)                         #underlying effect in a new study
delta[ns+2,2]<-d[2]                            #store to plot on caterpillar plot
delta[ns+3,2]<-delta.new

#Sensitivity analyses on the prior for tau
   #tau.sq~dunif(0.001,4)

   #ltau.sq~dunif(-10,1.386)
   #tau.sq<-exp(ltau.sq)

   #prec~dgamma(0.001,0.001)
   #tau.sq<-1/prec[4]

# Cross-validation for the Phelan study
# For plotting put mean and predicitive distn in elements (ns+2)& (ns+3) respectively
   for (i in 1:(ns-1)){deltaplot[i] <- delta[i,2]}
   deltaplot[(ns+2)] <- d[2]                    # RE mean
   deltaplot[(ns+3)] <- delta[ns,2]             # predictive distribution for "new" trial ns

   p.base ~ dbeta(a,b)         # draw baseline (control group) effect
   a <- r[ns+1,1]                      # no events in control group
   b <- n[ns+1,1]-r[ns+1,1]    # no of non-events in control group

# predictive prob of event in intervention group
logit(p.new) <- logit(p.base) + delta[ns,2]

# draw predicted number of events in intervention group
r.new ~ dbin(p.new, n[ns+1,2])

# Bayesian p-value: probability of obtaining a value as extreme as the
# value observed (r[ns+1,2], study 13 Phelan(2010)), given the model and the remaining data
# extreme value larger
p.cross <- step(r.new - r[ns+1,2]) - 0.5*equals(r.new,r[ns+1,2])
}
```

```
#data with new study cross validation
list(ns=12)
r[,2]    n[,2]    r[,1]    n[,1]
52       64       50       69
28       49       25       47
158      482      166      469
76       85       70       89
44       47       45       50
223.1508          323.6131          214.2603          323.6131          #cluster trial adjusted
12.84775          47.43786          10.87118          47.43786
310.9266          376.77969         348.4379          436.79774
23       54       10       41
408      742      328      718
60       69       29       38
NA       1        NA       1        #new trial
131      146      78       147      #Phelan study
END
```

#meta-regression – gender covariate
```
model{
for(i in 1:ns){
    rc[i]~dbin(pc[i],nc[i])                         # binomial likelihoods
    rt[i]~dbin(pt[i],nt[i])                         #different to previous model expression
    logit(pc[i])<- mu[i]                            # model for control group
    logit(pt[i])<- mu[i] + delta[i] + beta*(cov[i]-mn.cov)    # model for treat, covariate centred
    mu[i]~dnorm(0.0,1.0E-6)                         # vague priors for all study baselines
    delta[i]~dnorm(d,prec)
                   }
d~dnorm(0.0,1.0E-6)                                 # vague priors for intervention effect
prec<-1/(tau.sq)                                    #between-trial precision = 1/between-study variance
tau~dunif(0,10)                                     # vague priors for between study s.d.
tau.sq<-tau*tau

OR<-exp(d.uncent)
beta~dnorm(0.0,1.0E-06)

d.uncent<-d-(beta*mn.cov)
eq.point<-d.uncent/(-beta)

d.50<-d.uncent+beta*50
d.new.50~dnorm(d.50,prec)
OR.new.50<-exp(d.new.50)
  }
```

```
#Covariate %male
list(ns=8, mn.cov=51.63)
rt[]     nt[]     rc[]     nc[]     cov[]
28       49       25       47       57
158      482      166      469      59
76       85       70       89       48
223.1508          323.6131          214.2603          323.6131          52        #cluster trial adjusted
23       54       10       41       48
408      742      328      718      51
60       69       29       38       52
131      146      78       147      46
END
```

Appendix B Using the arcsine transformation for assessing publication bias in a pairwise meta-analysis

*** Using the arcsine transformation method by Thompson in Stata

*** (Rucker, Schwarzer, Carpenter 2008)

. gen part1 = asin(sqrt(adjstairgateni/(adjstairgateni + adjnostairgateni)))
. gen part2 = asin(sqrt(adjstairgatenc/(adjstairgatenc + adjnostairgatenc)))
. gen asindiff = part1- part2

. gen var_asindiff = 1/(4* (adjstairgateni+ adjnostairgateni)) + 1/(4*(adjstairgatenc+ adjnostairgatenc))
. gen se_asindiff = sqrt(var_asindiff)

. metareg asindiff se_asindiff, wsse(se_asindiff) graph mm

** In Stata the Knapp-Hartung modification is used in order to use the t-distribution

```
Meta-regression                                       Number of obs =      12
Method of moments estimate of between-study variance  tau2          = .01012
% residual variation due to heterogeneity             I-squared_res = 78.98%
Proportion of between-study variance explained        Adj R-squared =  1.02%
With Knapp-Hartung modification
-----------------------------------------------------------------------------
   asindiff |     Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------+----------------------------------------------------------------
se_asindiff |   .788278    1.20483     0.65   0.528    -1.896252    3.472808
      _cons |  .0535972   .0853681     0.63   0.544    -.1366148    .2438091
-----------------------------------------------------------------------------
```
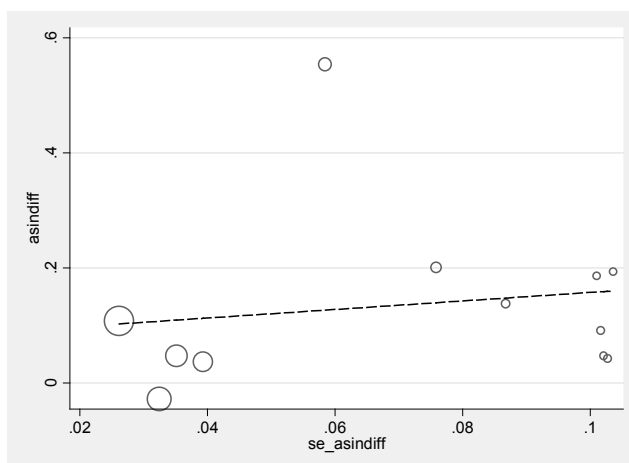
Appendix C Number of studies available for each outcome and covariate combination for the meta-regression

Covariate	Safety gates	Baby walkers	Bathmats	Window locks	High surfaces
Gender					
Cluster IPD	1	0	0	1	0
Non-Cluster IPD	5	3	2	3	1
Cluster AD	0	0	0	0	0
Non-Cluster AD	2	3	0	1	2
Ethnic group					
Cluster IPD	2	1	0	1	0
Non-Cluster IPD	7	2	2	3	2
Cluster AD	1	0	0	0	0
Non-Cluster AD	0	2	0	1	0
Family type					
Cluster IPD	2	1	0	1	0
Non-Cluster IPD	7	4	2	4	1
Cluster AD	1	0	1	0	0
Non-Cluster AD	0	1	0	0	1
Housing tenure					
Cluster IPD	2	1	0	1	0
Non-Cluster IPD	5	2	1	2	1
Cluster AD	0	0	0	0	0
Non-Cluster AD	1	3	0	0	2
Parental unemployment					
Cluster IPD	1	0	0	1	0
Non-Cluster IPD	4	3	3	4	1
Cluster AD	1	0	0	0	0
Non-Cluster AD	0	0	0	0	0
Age					
Cluster IPD	1	0	0	1	0
Non-Cluster IPD	7	4	2	2	2
Cluster AD	2	1	1	0	0
Non-Cluster AD	2	3	0	1	1

Appendix D Meta-regression combining IPD non-cluster, IPD cluster, aggregate non-cluster, aggregate cluster

Model to combine IPD non-cluster, IPD cluster, aggregate non-cluster, aggregate cluster
using random effects with covariates
Splitting between and within variance
Safety gate - gender

model
```
        {

### Model for non cluster ipd trial data ###

  for( i in 1:n.non.cluster.subjects) {
    outcome[i] ~ dbern(p[i])
    # covariate model

    #not splitting the study variance
    #logit(p[i]) <- mu[study[i]] + delta[study[i]]*treat[i] + beta0[study[i]]*cov[i] + beta*(cov[i]*treat[i])

    #splitting between and within study variance
    logit(p[i]) <- mu[study[i]] + delta[study[i]]*treat[i] + beta0[study[i]]*(cov[i] - mcov) +
        beta.w*((cov[i] - mcov) - (meancov[i] - mcov))*treat[i] + beta*(meancov[i] - mcov)*treat[i]

    temp1[i] <- id1[i]
    tempuniq1[i]<- uniqueid[i]
                }

  for(i in 1:n.ipd.non.cluster.trials){
                    mu[i]~dnorm(0, 1.0E-6)
                    beta0[i]~dnorm(0, 1.0E-6)
                    }

### Model for cluster ipd cluster trial data  ###

  for( i in 1:n.cluster.subjects) {
    c.outcome[i] ~ dbern(c.p[i])
    # cluster covariate model

    #not splitting the study variance
    #logit(c.p[i]) <- c.mu[c.study[i],c.cluster[i]] + delta[c.study[i]+ n.ipd.non.cluster.trials]*c.treat[i] +
        #c.beta0[c.study[i]]*c.cov[i] + beta*(c.cov[i]*c.treat[i])

    #splitting between and within study variance
    logit(c.p[i]) <- c.mu[c.study[i],c.cluster[i]] + delta[c.study[i] + n.ipd.non.cluster.trials]*c.treat[i] +
        c.beta0[c.study[i]]*(c.cov[i] - mcov) + beta.w*((c.cov[i]-mcov) - (c.meancov[i] -
        mcov))*c.treat[i] + beta*(c.meancov[i] - mcov)*c.treat[i]

    temp2[i] <- id2[i]
    tempuniq2[i] <- c.uniqueid[i]
                    }

  for(i in 1:n.ipd.cluster.trials){
    for(j in 1:n.cluster.max) {
      c.mu[i,j] ~ dnorm(mu.mean[i], inv.tau.sq.mu[i])
                    }

    mu.mean[i] ~ dnorm(0.0,1.0E-6)
    inv.tau.sq.mu[i]<-1/(sigma.mu[i]*sigma.mu[i])
    sigma.mu[i]~dunif(0,10)
    tau.sq.mu[i] <- sigma.mu[i]*sigma.mu[i]
    c.beta0[i] ~ dnorm(0.0,1.0E-6)
                    }
```

Model for aggregated non-cluster trial data

```
for( i in 1:n.agg.non.cluster.trials) {
  temp3[i] <- id3[i]
  rc[i] ~ dbin(pc[i], nc[i])
  rt[i] ~ dbin(pt[i], nt[i])
  logit(pc[i]) <- mu.a[i]
  logit(pt[i]) <- mu.a[i] + delta[I + n.ipd.non.cluster.trials + n.ipd.cluster.trials] + beta*a.cov[i]

  mu.a[i] ~ dnorm(0.0,1.0E-6)
              }
```

Model for aggregated cluster trial data – assuming not adjusted prior to this###

```
for( i in 1:n.agg.cluster.trials) {
  temp4[i] <- id4[i]
  design.effect[i] <- 1+ (ave.cluster.size[i] -1)*icc[i]
  outcome.var.corrected[i] <- outcome.var[i]*design.effect[i]
  outcome.prec[i] <- 1/outcome.var.corrected[i]
  outcome.logor[i]~dnorm(delt[i+n.ipd.non.cluster.trials+n.ipd.c#luster.trials+
        n.agg.non.cluster.trials], outcome.prec[i])
  delt[i+n.ipd.non.cluster.trials+n.ipd.cluster.trials+n.agg.non.cluster.trials] <-
        delta[i+n.ipd.non.cluster.trials+n.ipd.cluster.trials+n.agg.non.cluster.trials] + beta*ca.cov[i]

  icc[i] ~ dnorm(0.0079,1967)I(0,) #cluster adjustment
  #}
```

Model for combining all estimates of intervention effect from 4 data sources

```
for(j in 1:n.ipd.non.cluster.trials+n.ipd.cluster.trials+n.agg.non.cluster.trials+n.agg.cluster.trials){
  delta[j] ~ dnorm(d, inv.tau.sq)
              }

  d ~ dnorm(0.0, 1.0E-6)
  d.uncen<-d-mcov         #uncentring using the mean covariate
  or <- exp(d.uncen)
  orcov <- exp(d.uncen+beta)
  orcov.w <- exp(d.uncen+beta.w)
  p.value<-1-step(beta)

  beta ~ dnorm(0.0, 1.0E-6)
  beta.w ~ dnorm(0.0, 1.0E-6)

  inv.tau.sq<-1/(sigma*sigma)
  sigma~dunif(0,10)
  tau.sq <- sigma*sigma

  beta.diff<-beta.w - beta
  ratio.OR<-exp(beta)
  ratio.OR.w<-exp(beta.w)
        }
```

Data - example

list(n.non.cluster.subjects=2135, n.ipd.non.cluster.trials=5, n.agg.non.cluster.trials=2, n.agg.cluster.trials=0, n.ipd.cluster.trials=1, n.cluster.subjects=723, n.cluster.max=37, mcov=0.5119)

#non cluster IPD

id1[]	treat[]	cov[]	study[]	outcome[]		uniqueid[]	meancov[]
12	1	0	1	1	1	0.510416667	
12	0	0	1	1	3	0.510416667	
12	1	0	1	1	5	0.510416667	
12	0	0	1	1	6	0.510416667	
:							

#cluster IPD

id2[]	c.cov[]	c.cluster[]		c.treat[]	c.outcome[]		c.study[]	c.uniqueid[]	c.meancov[]
49	0	1	1	0	1	9	0.521438451		
49	0	2	1	1	1	22	0.521438451		
49	0	2	1	0	1	23	0.521438451		
49	0	2	1	1	1	24	0.521438451		
:									

#non cluster summary

id3[]	rt[]	nt[]	rc[]	nc[]	a.cov[]
24	158	482	166	469	0.59
9019	60	69	29	38	0.52
END					

#cluster summary – no studies for this covariate so blank out cluster summary model
#id4[]   outcome.logor[]   outcome.var[]   ave.cluster.size[]   ca.cov[]

Appendix E   Forest plots for the pairwise meta-analysis

Possession and use of a baby walker and possession of window locks)

Possession and use of a baby walker

Author(year)	Intervention events/total	Control events/total	OR (95% CI)	% Weight
King J (2001)	29/482	33/469	1.18 (0.71, 1.98)	13.50
Nansel (2002)	19/85	30/89	1.77 (0.90, 3.46)	10.30
Snazjder (2003)	19/47	14/50	0.57 (0.25, 1.34)	7.70
Tan (2004)	143/228	393/480	2.69 (1.88, 3.83)	17.58
Posner (2004)	4/7	4/8	0.75 (0.10, 5.77)	1.81
Kendrick (2005)*	60/212	105/248	1.86 (1.26, 2.75)	16.62
Babul (2007)**	48/335	31/148	1.58 (0.96, 2.61)	13.86
Nansel (2008)	13/69	12/38	1.99 (0.80, 4.95)	6.96
Phelan (2010)	24/140	29/138	1.29 (0.71, 2.34)	11.66
Overall (I-squared = 51.3%, p = 0.037)			1.57 (1.18, 2.09)	100.00

NOTE: Weights are from random effects analysis

Possession of windows locks

Author(year)	Intervention events/total	Control events/total	OR (95% CI)	% Weight
Clamp (1998)	80/83	72/82	3.70 (0.98, 13.99)	4.47
Kendrick (1999)*	323/362	339/366	0.66 (0.39, 1.10)	18.56
King J (2001)	299/482	285/469	1.05 (0.81, 1.37)	31.51
Watson (2004)	550/767	493/741	1.27 (1.02, 1.59)	33.96
Hendrickson (2004)	24/34	21/39	2.06 (0.78, 5.43)	7.67
Phelan (2010)	146/149	145/150	1.68 (0.39, 7.15)	3.82
Overall (I-squared = 51.8%, p = 0.065)			1.17 (0.87, 1.57)	100.00

NOTE: Weights are from random effects analysis

* cluster randomised trial, numbers adjusted for clustering prior to conducting the meta-analysis
** three arm study, enhanced intervention arm

Possession of bath mats and never leaving a child unattended on a high surface

Possession non-slip bath mats or decals

Author(year)	Intervention events/total	Control events/total	OR (95% CI)	% Weight
Petridou (1997)*	49/98	50/100	0.98 (0.56, 1.72)	31.26
Snazjder (2003)	31/48	37/49	0.59 (0.25, 1.43)	19.30
Posner (2004)	44/49	34/47	3.36 (1.09, 10.36)	13.74
Phelan (2010)	59/149	56/150	1.10 (0.69, 1.75)	35.71
Overall (I-squared = 48.6%, p = 0.120)			1.10 (0.68, 1.78)	100.00

NOTE: Weights are from random effects analysis

Does not leave child unattended on high surface

Author(year)	Intervention events/total	Control events/total	OR (95% CI)	% Weight
Posner (2004)	5/49	6/47	1.29 (0.37, 4.54)	8.13
Babul (2007)**	173/331	69/148	0.80 (0.54, 1.18)	85.67
Nansel (2008)	55/62	21/24	0.89 (0.21, 3.77)	6.21
Overall (I-squared = 0.0%, p = 0.773)			0.84 (0.58, 1.20)	100.00

NOTE: Weights are from random effects analysis

* cluster randomised trial, numbers adjusted for clustering prior to conducting the meta-analysis
** three arm study, enhanced intervention arm

Appendix F Meta-regression combining IPD and aggregate data ORs (95% credible interval) for falls prevention practices by child age, gender, and social group

Covariate	Safety gates	Baby walkers	Bathmats	Window locks	High surfaces
Between & within variance combined					
Gender			*fixed*		*zero cells in only IPD*
Boys	1.64(0.85 to 3.31)	0.67(0.32 to 1.37)	1.34(0.70 to 2.52)	1.45(0.80 to 2.92)	
Girls	1.92(0.99 to 3.85)	1.04(0.49 to 2.18)	1.27(0.71 to 2.31)	0.85(0.46 to 1.70)	
Interaction term	0.86(0.62 to 1.18)	0.64(0.26 to 1.59)	1.05(0.43 to 2.53)	1.72(1.16 to 2.57)	
Ethnic group			*fixed*		*fixed*
Black & minority ethnic groups	1.98(1.17 to 3.34)	0.77(0.29 to 2.49)	1.63(0.80 to 3.73)	1.58(0.58 to 5.11)	0.85(0.26 to 3.01)
White	1.65(1.01 to 2.76)	1.03(0.30 to 2.59)	1.17(0.68 to 2.00)	1.36(0.57 to 3.43)	0.59(0.05 to 13.32)
Interaction term	1.19(0.77 to 1.85)	0.79(0.33 to 2.02)	1.40(0.58 to 3.42)	1.13(0.62 to 2.05)	1.43(0.04 to 31.78)
Family type			*fixed*		*fixed*
Single-parent family	2.03(1.16 to 3.62)	0.89(0.32 to 2.46)	0.60(0.16 to 1.99)	0.98(0.37 to 3.19)	4588(1.18 to 3.14×10^{9})
Two-parent family	1.82(1.12 to 3.02)	0.92(0.41 to 1.87)	1.00(0.69 to 1.44)	1.51(0.63 to 4.76)	0.46(0.09 to 1.34)
Interaction term	1.11(0.75 to 1.65)	0.99(0.44 to 2.24)	0.60(0.15 to 2.14)	0.65(0.40 to 1.05)	9933(0.98 to 3.2×10^{10})
Housing tenure			*N/A – 1 IPD*	*fixed*	*fixed*
Non-owner occupied	1.98(1.48 to 2.66)	1.22(0.48 to 2.93)		1.13(0.03 to 54.7)	0.44(0.04 to 3.65)
Owner occupied	1.22(0.96 to 1.61)	1.36(0.53 to 3.34)		1.48(0.04 to 75.5)	2.51(0.58 to 13.06)
Interaction term	1.62(1.18 to 2.24)	0.90(0.54 to 1.47)		0.76(0.50 to 1.17)	0.18(0.003 to 5.76)
Unemployed		*fixed*	*fixed*		*N/A – 1 IPD*
1 or more parents unemployed	2.08(0.77 to 5.86)	0.39(0.14 to 1.04)	2.07(0.91 to 4.78)	1.40(0.58 to 4.23)	
Parents employed	1.82(0.67 to 5.01)	0.87(0.49 to 1.51)	0.91(0.59 to 1.42)	1.40(0.63 to 4.49)	
Interaction term	1.15(0.77 to 1.71)	0.45(0.14 to 1.40)	2.28(0.88 to 5.86)	0.98(0.62 to 1.55)	
Age		*N/A due to age of walker use*	*fixed*		*N/A due to age for leaving child on high surface*
OR at age 0	1.40(1.02 to 2.06)		1.16(0.80 to 1.71)	1.00(0.30 to 4.87)	
OR at age 4	1.26(0.81 to 2.02)		1.08(0.78 to 1.50)	1.27(0.35 to 5.84)	
Interaction term	0.97(0.84 to 1.13)		0.98(0.90 to 1.06)	1.06(0.90 to 1.23)	

Appendix G Meta-regression combining IPD and aggregate data splitting the between and within study variance

ORs (95% credible interval) for falls prevention practices by child age, gender, and social group*

Covariate	Safety gates	Baby walkers	Window locks
Between & within variance split			
Gender	*centred*	*centred*	*centred*
Within			
Boys	1.04(0.47 to 2.40)	0.42(0.11 to 1.44)	1.18(0.40 to 3.63)
Girls	1.22(0.58 to 2.66)	0.65(0.24 to 1.53)	0.69(0.25 to 2.00)
Interaction term	0.86(0.62 to 1.18)	0.65(0.27 to 1.58)	1.72(1.16 to 2.54)
Between			
Boys	0.42(0.04 to 4.25)	0.28(0.06 to 1.84)	0.57(0.03 to 17.31)
Girls	1.22(0.58 to 2.66)	0.65(0.24 to 1.53)	0.69(0.25 to 2.00)
Interaction term	0.34(0.03 to 4.40)	0.44(0.05 to 5.36)	0.83(0.04 to 32.82)
Difference	0.92(-1.64 to 3.44)	0.40(-2.22 to 2.61)	0.73(-2.95 to 3.90)
Ethnic group			
Within			
Black & minority ethnic groups	2.04(0.86 to 5.07)	1.45(0.18 to 6.65)	1.38(0.28 to 16.38)
White	1.70(0.83 to 3.64)	1.49(0.20 to 5.01)	1.25(0.29 to 14.7)
Interaction term	1.20(0.73 to 1.96)	1.02(0.41 to 2.53)	1.08(0.58 to 1.99)
Between			
Black & minority ethnic groups	1.91(0.79 to 4.33)	0.49(0.14 to 3.54)	2.21(0.02 to 46.99)
White	1.70(0.83 to 3.64)	1.49(0.20 to 5.01)	1.25(0.29 to 14.7)
Interaction term	1.12(0.32 to 3.68)	0.33(0.06 to 7.72)	1.77(0.003 to 61.72)
Difference	0.07(-1.22 to 1.45)	1.11(-2.13 to 2.99)	-0.48(-4.08 to 5.97)
Family type			
Within			
Single-parent family	2.25(1.00 to 3.17)	1.56(0.28 to 8.21)	0.70(0.03 to 18.15)
Two-parent family	1.99(0.98 to 4.17)	1.52(0.32 to 6.70)	1.09(0.06 to 26.92)
Interaction term	0.78(0.10 to 5.68)	1.03(0.46 to 2.34)	0.64(0.39 to 1.03)
Between			
Single-parent family	1.57(0.32 to 7.55)	0.04(0.00 to 105.7)	4.63 (0 to 3.9×10^6)
Two-parent family	1.99(0.98 to 4.17)	1.52(0.32 to 6.70)	1.09(0.06 to 26.92)
Interaction term	1.13(0.76 to 1.69)	0.03(0.00 to 253.0)	4.26(0 to 4×10^7)
Difference	0.36(-1.68 to 2.46)	3.62(-5.53 to 12.38)	-1.90(-17.99 to 13.27)
Housing tenure			*N/A*
Within			
Non-owner occupied	1.73(0.98 to 3.07)	0.89(0.04 to 20.79)	
Owner occupied	1.10(0.71 to 1.74)	1.01(0.04 to 22.64)	
Interaction term	1.58(1.13 to 2.22)	0.89(0.53 to 1.48)	
Between			
Non-owner occupied	2.36(1.19 to 4.67)	1.54(0.10 to 21.6)	
Owner occupied	1.10(0.71 to 1.74)	1.01(0.04 to 22.64)	
Interaction term	2.13(0.75 to 6.04)	1.55(0.01 to 227.3)	
Difference	-0.31(-1.40 to 0.78)	-0.56(-5.56 to 4.74)	

Appendix G (continued)

Covariate	Safety gates	Baby walkers	Window locks
Unemployed			
Within		N/A	
1 or more parents unemployed	2.94(0.36 to 15.49)		0.83(0.12 to 27.23)
Parents employed	2.53(0.32 to 12.41)		0.90(0.14 to 29.21)
Interaction term	1.17(0.78 to 1.75)		0.90(0.567 to 1.43)
Between			
1 or more parents unemployed	1.30(0.16 to 20.13)		2.33(0.05 to 36.86)
Parents employed	2.53(0.32 to 12.41)		0.90(0.14 to 29.21)
Interaction term	0.49(0.03 35.38)		2.63(0.003 to 127.8)
Difference	0.85(-3.43 to 3.82)		-1.07(-4.99 to 5.82)
Age			
Within			
OR at age 0	2.36(1.16 to 5.02)	0.60(0.09 to 5.41)	
OR at age 4	1.91(0.72 to 5.30)	18.9(0.05 to 10630)	
Interaction term	0.95(0.79 to 1.13)	2.35(0.58 to 10.28)	
Between			
OR at age 0	2.36(1.16 to 5.02)	0.60(0.09 to 5.41)	
OR at age 4	0.85(0.20 to 3.40)	3.68(0.01 to 788.5)	
Interaction term	0.78(0.47 to 1.24)	1.58(0.18 to 9.17)	
Difference	0.20(-0.29 to 0.72)	0.41(-1.80 to 2.94)	

* there is insufficient data to split the variance for the bathmats, high surfaces and window locks prevention practices

Appendix H WinBUGS code for NMA

```
# Network meta-analysis

# Binomial likelihood, logit link, random effect model, multi-arm trials

model{
for(i in 1:ns){                          # LOOP THROUGH STUDIES
        temp1[i] <- id[i]
 w[i,1] <- 0                             # adjustment for multi-arm trials is zero for control arm
 delta[i,1] <- 0                         # intervention effect is zero for control arm
 mu[i] ~ dnorm(0,.0001)                  # vague priors for all study baselines

 for (k in 1:na[i]) {                    # LOOP THROUGH ARMS
   r[i,k] ~ dbin(p[i,k],n[i,k])          # binomial likelihood
   logit(p[i,k]) <- mu[i] + delta[i,k]   # model for linear predictor
   rhat[i,k] <- p[i,k] * n[i,k]          # expected value of the numerators
    #Deviance contribution
       dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k])) + (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k] –
                  rhat[i,k])))
           }
 resdev[i] <- sum(dev[i,1:na[i]])        # summed res deviance contribution for study

 for (k in 2:na[i]) {                            # LOOP THROUGH ARMS
   delta[i,k] ~ dnorm(md[i,k],taud[i,k])         # trial-specific LOR distributions
   md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]    # mean of LOR dist(with multi-arm correction)
   taud[i,k] <- tau *2*(k-1)/k                   # precision of LOR dists(with multi-arm correction)
   w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]]) # adjustment for multi-arm RCTs
   sw[i,k] <- sum(w[i,1:k-1])/(k-1)              # cumulative adjustment for multi-arm trials
          }
}

totresdev <- sum(resdev[])               #Total Residual Deviance
d[1]<- 0                                 # intervention effect is zero for reference intervention
for (k in 2:nt)  { d[k] ~ dnorm(0,.0001)}    # vague priors for intervention effects
sd ~ dunif(0,2)
tau <- pow(sd,-2)


# Intervention A baseline, based on  average of the 8 studies including it.
for (i in 1:ns) { mu1[i] <- mu[i] * equals(t[i,1],1) }
for (k in 1:nt)  { logit(T[k])<- sum(mu1[])/8 +d[k] }

mn.mu1<-sum(mu1[])/8

#Rank the intervention effects (with 1=best) & record the best intervention
for (k in 1:nt) { rk[k]<-nt+1 - rank(T[],k)
          best[k]<-equals(rk[k],1)}

#All pairwise log odds ratios and odds ratios
for (c in 1:(nt-1))
        { for (k in (c+1):nt)
                { lor[c,k] <- d[k] - d[c]
                  log(or[c,k]) <- lor[c,k]
                  RR[c,k]<-or[c,k]/(1-mn.mu1+mn.mu1*or[c,k])
                }
        }
}
```

Network Meta-Analysis Summary Forest Plot Table for Possession of safety gates

Comparators		H-H Trials	Odds Ratio (95% Crl)	Summary Forest Plot (log Scale)
E+Eq+Fitting+HI Median Rank	E+Equipment	0	0.21 (0.07 to 0.64) NA	
	E+Eq+Fitting	0	0.20 (0.06 to 0.67) NA	
	E+HI	0	0.19 (0.05 to 0.85) NA	
	Education	0	0.19 (0.06 to 0.56) NA	
	E+Eq+HI	0	0.17 (0.05 to 0.58) NA	
	Usual Care	1	0.13 (0.05 to 0.34) 0.13 (0.07 to 0.24)	
E+Equipment	E+Eq+Fitting	0	0.94 (0.40 to 2.45) NA	
	E+HI	0	0.88 (0.30 to 2.96) NA	
	Education	1	0.88 (0.45 to 1.77) 0.85 (0.38 to 1.91)	
	E+Eq+HI	1	0.78 (0.37 to 1.82) 1.25 (0.49 to 3.17)	
	Usual Care	2	0.62 (0.33 to 1.09) 0.52 (0.11 to 2.36)	
E+Eq+Fitting	E+HI	0	0.95 (0.26 to 3.56) NA	
	Education	1	0.93 (0.41 to 2.10) 0.61 (0.14 to 2.72)	
	E+Eq+HI	0	0.84 (0.31 to 2.31) NA	
	Usual Care	1	0.66 (0.29 to 1.26) 0.69 (0.56 to 0.85)	
E+HI	Education	0	0.99 (0.29 to 3.16) NA	
	E+Eq+HI	1	0.88 (0.39 to 2.06) 0.89 (0.68 to 1.16)	
	Usual Care	0	0.70 (0.20 to 1.93) NA	
Education	E+Eq+HI	0	0.90 (0.40 to 2.17) NA	
	Usual Care	3	0.70 (0.39 to 1.11) 0.64 (0.19 to 1.73)	
E+Eq+HI	Usual Care	1	0.79 (0.34 to 1.47) 0.88 (0.63 to 1.23)	

Heterogeneity: between-study variance = 0.05; 95% Crl (0.000 to 0.835)

Key: NMA results in black; Pairwise MA results in grey. 95% Crl and PI presented as error bars. Interventions are displayed sorted by median rank.

Odds Ratio with 95% Crl & 95% PI (log scale)

1/16   1/4   1   4

Appendix J NMA results for possession of window locks and not leaving child on high surface

Results of a NMA model of the possession of window locks expressed as Odds Ratios (95% CrI)

	Usual care (1)	Education + Equipment (3)	Education + Equipment + Home inspection (4)	Education + Equipment + Fitting (5)	Education + Home inspection (6)	Education + Equipment + Fitting + Home inspection (7)
Usual care (1)		4.09 (0.27 , 67.87)	1.05 (0.19 , 6.89)	1.28 (0.11 , 14.2)	1.10 (0.057 , 25.2)	1.74 (0.11 , 30.5)
Education + Equipment (3)	0.27 (0.46 , 1.11)		0.26 (0.010 , 7.29)	0.31 (0.007 , 11.4)	0.27 (0.004 , 17.1)	0.42 (0.008 , 21.0)
Education + Equipment + Home inspection (4)	0.93 (0.31 , 2.80)			1.24 (0.054 , 22.1)	1.06 (0.092 , 12.6)	1.65 (0.055 , 44.1)
Education + Equipment + Fitting (5)	0.78 (0.63, 0.98)				0.85 (0.020 , 44.8)	1.38 (0.034 , 57.2)
Education + Home inspection (6)			0.95 (0.72 , 1.24)			1.56 (0.024 , 89.8)
Education + Equipment + Fitting + Home inspection (7)	0.60 (0.091 , 3.13)					

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio.
[a] values above the stepped line are results from the NMA; those below the line are direct estimates from a trial or where more than one were available, a meta-analysis. Blank cells indicate that no direct evidence on specific pairwise comparisons was available.
[b] Column and row headings signify intervention or comparison (intervention number)
[c] significant at the 5% level

Assessment of best intervention for the possession of window locks

	Possession of window locks	
Intervention	Probability intervention is best	Median intervention rank (95% CrI)
Usual care (1)	0.15	3(1, 5)
Education + Equipment (3)	0.052	6(1, 6)
Education + Equipment + Home inspection (4)	0.16	3(1, 6)
Education + Equipment + Fitting (5)	0.19	4(1, 6)
Education + Home inspection (6)	0.26	3(1, 6)
Education + Equipment +Fitting + Home inspection (7)	0.18	4(1, 6)

Results of a NMA model of not leaving a child alone on high surface expressed as Odds Ratios (95% CrI)

	Usual care (1)	Education (2)	Education + Equipment (3)	Education + Equipment + Home inspection (4)
Usual care (1)		1.94 (0.20 , 14.37)	1.13 (0.12 , 6.75)	1.01 (0.089 , 9.18)
Education (2)	0.89 (0.18 , 5.84)		0.56 (0.064 , 4.65)	0.50 (0.032 , 8.76)
Education + Equipment (3)	0.71 (0.44 , 1.13)	3.06 (0.61 , 14.25)		0.89 (0.099 , 9.67)
Education + Equipment + Home inspection (4)	0.89 (0.56 , 1.42)		1.27 (0.80 , 2.00)	

Abbreviations: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio.

Values above the stepped line are results from the NMA; those below the line are direct estimates from a trial or where more than one were available, a meta-analysis. Blank cells indicate that no direct evidence on specific pairwise comparisons was available.

[c] significant at the 5% level

Assessment of best intervention for the outcome of not leaving a child alone on high surface

	Not leaving a child alone on a high surface	
Intervention	Probability intervention is best	Median intervention rank (95% CrI)
Usual care (1)	0.32	2(1, 4)
Education (2)	0.10	4(1, 4)
Education + Equipment (3)	0.23	2(1, 4)
Education + Equipment + Home inspection (4)	0.35	2(1, 4)

Appendix K Sensitivity analysis for safety gate NMA
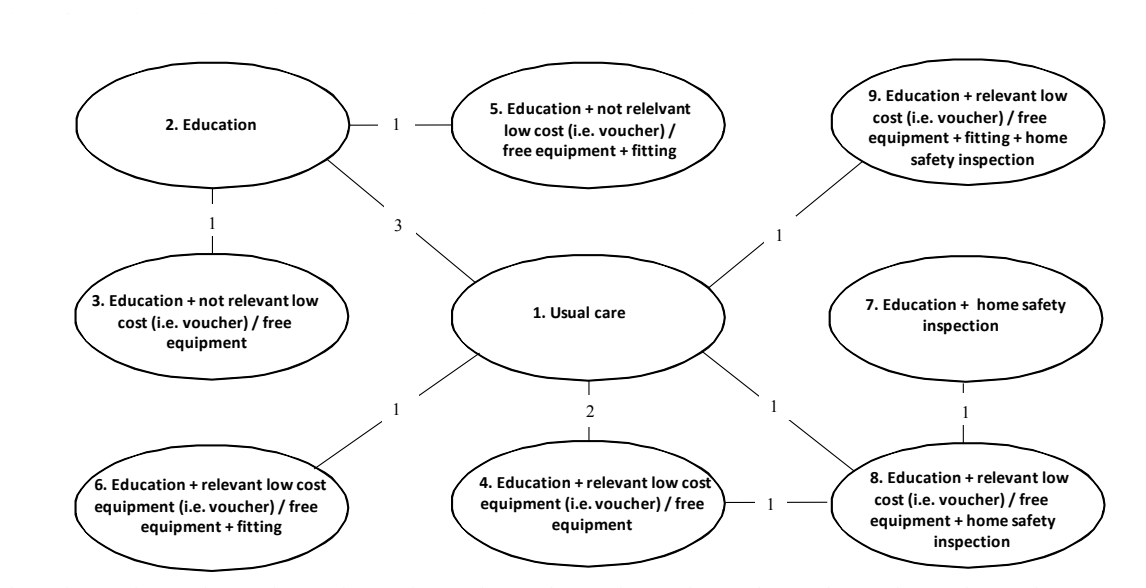
Network diagram of interventions for the possession of a fitted stair gate: sensitivity analysis, equipment component split into relevant or not relevant equipment.

Each intervention is a node in the network. The links between the nodes are pairwise intervention comparisons. The numbers along the link lines indicate the number of studies or pairs of study arms for that link in the network.

Sensitivity Analysis Results of a NMA model for interventions to increase possession of a stair gate expressed as Odds Ratios [a,b]

	Usual care (1)	Education (2)	Education + Not relevant Equipment (3)	Education + Relevant Equipment (4)	Education + Not relevant Equipment + Fitting (5)	Education + Relevant Equipment + Fitting (6)	Education + Home inspection (7)	Education + Relevant Equipment + Home inspection (8)	Education + Relevant Equipment + Fitting + Home inspection (9)
Usual care (1)		1.53 (0.79 , 3.76)	1.83 (0.42 , 9.77)	1.63 (0.70 , 4.03)	2.79 (0.38 , 23.34)	1.46 (0.44 , 4.70)	1.14 (0.28 , 6.15)	1.30 (0.53, 3.94)	7.87[c] (2.15, 28.32)
Education (2)	1.48 (0.97 , 2.25)		1.19 (0.30 , 4.55)	1.05 (0.31 , 3.15)	1.76 (0.28 , 12.63)	0.96 (0.20 , 3.56)	0.75 (0.13 , 4.1)	0.85 (0.24 , 2.91)	5.18[c] (1.10, 20.68)
Education + Not relevant Equipment (3)		1.17 (0.52 , 2.63)		0.88 (0.14 , 4.93)	1.52 (0.15 , 16.15)	0.81 (0.10 , 5.11)	0.64 (0.07 , 5.60)	0.7 (0.11 , 4.39)	4.35 (0.50 , 28.13)
Education + Relevant Equipment (4)	1.92[c] (1.11 , 3.35)				1.69 (0.19 , 17.19)	0.90 (0.20 , 3.79)	0.72 (0.16 , 3.73)	0.81 (0.30 , 2.49)	4.86 (0.99 , 22.6)
Education + Not relevant Equipment + Fitting (5)		1.63 (0.37 , 7.23)				0.53 (0.05 , 5.09)	0.43 (0.03 , 5.23)	0.48 (0.05 , 4.60)	2.87 (0.24 , 29.88)
Education + Relevant Equipment + Fitting (6)	1.45[c] (1.18 , 1.79)						0.78 (0.13 , 6.40)	0.89 (0.21 , 4.74)	5.40 (0.91 , 31.76)
Education + Home inspection (7)								1.13 (0.35 , 3.60)	6.89 (0.80 , 44.42)
Education + Relevant Equipment + Home inspection (8)	1.13 (0.82, 1.58)			1.25 (0.49, 3.17)			1.12 (0.86, 1.47)		6.09[c] (1.07, 27.46)
Education + Relevant Equipment + Fitting + Home inspection (9)	7.73[c] (4.14, 14.4)								

Model fit:

Posterior mean residual deviance   23.53 (cf 24 data points)
Between-study standard deviation   0.34 (0.022, 1.36)
Probability intervention 9 is best = 0.77 with a median intervention ranking of 1(1, 4)

Abbreviations: CrI, credible interval; CI, confidence interval; NMA, network meta-analysis; OR, odds ratio.
[a] values above the diagonal are results from the NMA, OR with 95%CrI; those below the line are direct estimates from a trial or, where more than one are available, a meta-analysis with 95%CI. Blank cells indicate that no direct evidence on specific pairwise comparisons is available. [b] Column and row headings signify intervention or comparison (intervention number).
[c] significant at the 5% level

Appendix L   NMA RE model with a covariate - WinBUGS code

#Model 1 NMA with covariate independent beta, model not given as not fitted

Model 2 NMA with covariate exchangeable betas

```
model{
for(i in 1:ns){                         # LOOP THROUGH STUDIES
          temp1[i] <- id[i]
#  w[i,1] <- 0                          # adjustment for multi-arm trials is zero for control arm
 delta[i,1] <- 0                        # intervention effect is zero for control arm
 mu[i] ~ dnorm(0,.0001)                 # vague priors for all trial baselines

 for (k in 1:na[i]) {                    # LOOP THROUGH ARMS
   r[i,k] ~ dbin(p[i,k],n[i,k])          # binomial likelihood
   logit(p[i,k]) <- mu[i] + delta[i,k]   # model for linear predictor
   rhat[i,k] <- p[i,k] * n[i,k]          # expected value of the numerators
    #Deviance contribution
          dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
             + (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
          }
 resdev[i] <- sum(dev[i,1:na[i]])        # summed res deviance contribution for trial

 for (k in 2:na[i]) {                   # LOOP THROUGH ARMS
    delta[i,k] ~ dnorm(md[i,k],prec)     # trial-specific LOR distributions
    md[i,k] <- d[t[i,k]] - d[t[i,1]] +(beta[t[i,k]]-beta[t[i,1]])*(cov[i]-meancov)
          }
}

totresdev <- sum(resdev[])               #Total Residual Deviance
d[1]<- 0               # intervention effect is zero for reference intervention
beta[1]<-0

for (k in 2:nt)  { d[k] ~ dnorm(0,.0001)         # vague priors for intervention effects
          beta[k]~dnorm(m.beta,prec.beta)}

m.beta~dnorm(0,.0001)
tau.beta~dunif(0,2)
prec.beta<-pow(tau.beta,-2)
tau ~ dunif(0,2)
prec <- pow(tau,-2)
tau.sq <-tau*tau

# intervention effect when covariate = z[j] (un-centring intervention effects)
#for (k in 1:nt){
#   for (j in 1:nz) { dz[j,k] <- d[k] - (beta[k]-beta[1])*(meancov-z[j]) }
# }

# Intervention A baseline, based on  average of the nb trials including it.
for (i in 1:ns) { mu1[i] <- mu[i] * equals(t[i,1],1) }
for (k in 1:nt)  { logit(T[k])<- sum(mu1[])/nb +d[k] }

#mn.mu1<-sum(mu1[])/nb

#Rank the intervention effects (with 1=best) & record the best intervention
for (k in 1:nt) { rk[k]<-nt+1 - rank(T[],k)
          best[k]<-equals(rk[k],1)}

#All pairwise log odds ratios and odds ratios
for (c in 1:(nt-1))
          { for (k in (c+1):nt)
                  { lor[c,k] <- d[k] - d[c]
                   log(or[c,k]) <- lor[c,k]
```

```
#estimates of effect evaluated at specified covariate values (z)
# at covariate=z[j]
#      for (j in 1:nz) {
#        orz[j,c,k] <- exp(dz[j,k] - dz[j,c])
#        lorz[j,c,k] <- (dz[j,k]-dz[j,c])
#                       }
             }
         }
 }

#Model 3  NMA with covariate common beta

model{
for(i in 1:ns){                       # LOOP THROUGH STUDIES
          temp1[i] <- id[i]
 delta[i,1] <- 0                   # intervention effect is zero for control arm
 mu[i] ~ dnorm(0,.0001)              # vague priors for all trial baselines

 for (k in 1:na[i]) {                   # LOOP THROUGH ARMS
   r[i,k] ~ dbin(p[i,k],n[i,k])              # binomial likelihood
   logit(p[i,k]) <- mu[i] + delta[i,k]         # model for linear predictor
   rhat[i,k] <- p[i,k] * n[i,k]          # expected value of the numerators
  #Deviance contribution
          dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
          + (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))
          }
 resdev[i] <- sum(dev[i,1:na[i]])      # summed res deviance contribution for trial


 for (k in 2:na[i]) {                   # LOOP THROUGH ARMS
   delta[i,k] ~ dnorm(md[i,k],prec)      # trial-specific LOR distributions
   md[i,k] <- d[t[i,k]] - d[t[i,1]] +(beta[t[i,k]]-beta[t[i,1]])*(cov[i]-meancov)
          }
}

totresdev <- sum(resdev[])                        #Total Residual Deviance
d[1]<- 0                 # intervention effect is zero for reference intervention
beta[1]<-0

for (k in 2:nt)  { d[k] ~ dnorm(0,.0001)  # vague priors for intervention effects
                 beta[k]<-B}       #common covariate effect

B~dnorm(0,.0001)                         #vague prior for common covariate effect
tau ~ dunif(0,5)
prec <- pow(tau,-2)
tau.sq<-tau*tau

# intervention effect when covariate = z[j] (un-centring intervention effects)
#for (k in 1:nt){
#   for (j in 1:nz) { dz[j,k] <- d[k] - (beta[k]-beta[1])*(meancov-z[j])
#                                                         }
#                              }

# Intervention A baseline, based on average of the nb trials including it.
for (i in 1:ns) { mu1[i] <- mu[i] * equals(t[i,1],1) }
for (k in 1:nt)  { logit(T[k])<- sum(mu1[])/nb +d[k] }

mn.mu1<-sum(mu1[])/nb

#Rank the intervention effects (with 1=best) & record the best intervention
for (k in 1:nt) { rk[k]<-nt+1 - rank(T[],k)
          best[k]<-equals(rk[k],1)}

#All pairwise log odds ratios and odds ratios
for (c in 1:(nt-1))
         { for (k in (c+1):nt)
```

```
                    { lor[c,k] <- d[k] - d[c]
                      log(or[c,k]) <- lor[c,k]

# at covariate=z[j]
#      for (j in 1:nz) {
#        orz[j,c,k] <- exp(dz[j,k] - dz[j,c])
#        lorz[j,c,k] <- (dz[j,k]-dz[j,c])
#                                                                              }
                                      }
            }
}
```

#**data** % black or minority ethnicity covariate
ns= number of studies; nt=number of interventions; nb=number of studies with baseline=1
#meancov = mean of covariate value for centring

list(nt=6, ns=10, meancov=48.9, nb=8)

t[,1]	t[,2]	r[,1]	n[,1]	r[,2]	n[,2]	na[]	id[]	cov[]
1	2	70	89	76	85	2	26	19
1	2	348.437	436.79	310.926	376.779	2	259	5
1	2	29	38	60	69	2	9019	32
1	3	50	69	52	64	2	4	10
1	3	10	41	23	54	2	344	54
1	4	214.260	323.613	223.150	323.613	2	49	12
1	5	328	718	408	742	2	345	28
1	6	78	147	131	146	2	9042	18
2	5	45	50	44	47	2	29	13
3	4	10.8711	47.4378	12.8477	47.4378	2	122	87
END								

Appendix M NMA RE model combining IPD and aggregate data – WinBUGS code

#MTC AD+IPD all no covariates

model {

Model for non-clustered ipd trial data

```
for(i in 1:n.non.cluster.subjects) {
        logit(p[i])<- mu[study[i]] + delta[index[i]] * (1-equals(treat[i],baseline[i]))
        outcome[i] ~dbern(p[i])      #likelihood for non-clustered IPD data
        temp1[i] <- id1[i]
        temp2[i]<- uniqueid[i]


        }

for(l in 1:n.non.cluster.arms) {
        md[l]<-d[treat1[l]]-d[baseline1[l]]
        delta[l]~dnorm(md[l], prec)
}

for(j in 1:n.ipd.non.cluster.trials) {
        mu[j]~dnorm(0,1.0E-6)
        }
```

Model for cluster ipd cluster trial data

```
for(i in 1:n.cluster.subjects) {
        c.outcome[i] ~ dbern(c.p[i])
        logit(c.p[i]) <- c.mu[c.study[i],c.cluster[i]] + delta[c.index[i]+n.non.cluster.arms] *
                                                (1-equals(c.treat[i],c.baseline[i]))

        temp5[i]<- id2[i]
        temp8[i] <- c.uniqueid[i]

                }

for(l in (n.non.cluster.arms+1):(n.non.cluster.arms+n.cluster.arms)) {
        md[l]<-d[treat1[l]]-d[baseline1[l]]
        delta[l]~dnorm(md[l], prec)
}

for(i in 1:n.ipd.cluster.trials) {
        for(j in 1:n.cluster.max) {
            c.mu[i,j] ~ dnorm(mu.mean[i], inv.tau.sq.mu[i])
    }
        mu.mean[i] ~ dnorm(0.0,1.0E-6)
        inv.tau.sq.mu[i]<-1/(sigma.mu[i]*sigma.mu[i])
        sigma.mu[i]~dunif(0,10)
        tau.sq.mu[i] <- sigma.mu[i]*sigma.mu[i]
        }
```

Model for non-cluster and cluster aggregate data

```
for(i in 1:n.agg.arms) {
        logit(pa[i])<- mu.ad[a.study[i]] + delta[i + n.non.cluster.arms + n.cluster.arms] *
                                        (1-equals(a.treat[i],a.base[i]))
        temp11[i] <- id4[i]
        outcome.ad[i]~dbin(pa[i],n[i])
        delta[i + n.non.cluster.arms + n.cluster.arms]~dnorm(md.ad[i], prec)
        md.ad[i]<-d[a.treat[i]]-d[a.base[i]]
        }
```

```
for(j in 1:n.agg.trials) {
        mu.ad[j]~dnorm(0,1.0E-6)
        }
```

Model for combining all estimates of intervention effect

```
d[1]<-0
for (k in 2:max.treat) {
        d[k] ~ dnorm(0,1.0E-6)
        }

tau~dunif(0,10)
tau.sq<-tau*tau
prec<-1/(tau.sq)

# pairwise ORs
for (c in 1:(max.treat - 1)) {
        for (k in (c + 1):max.treat) {
                lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
                }
        }


}
```
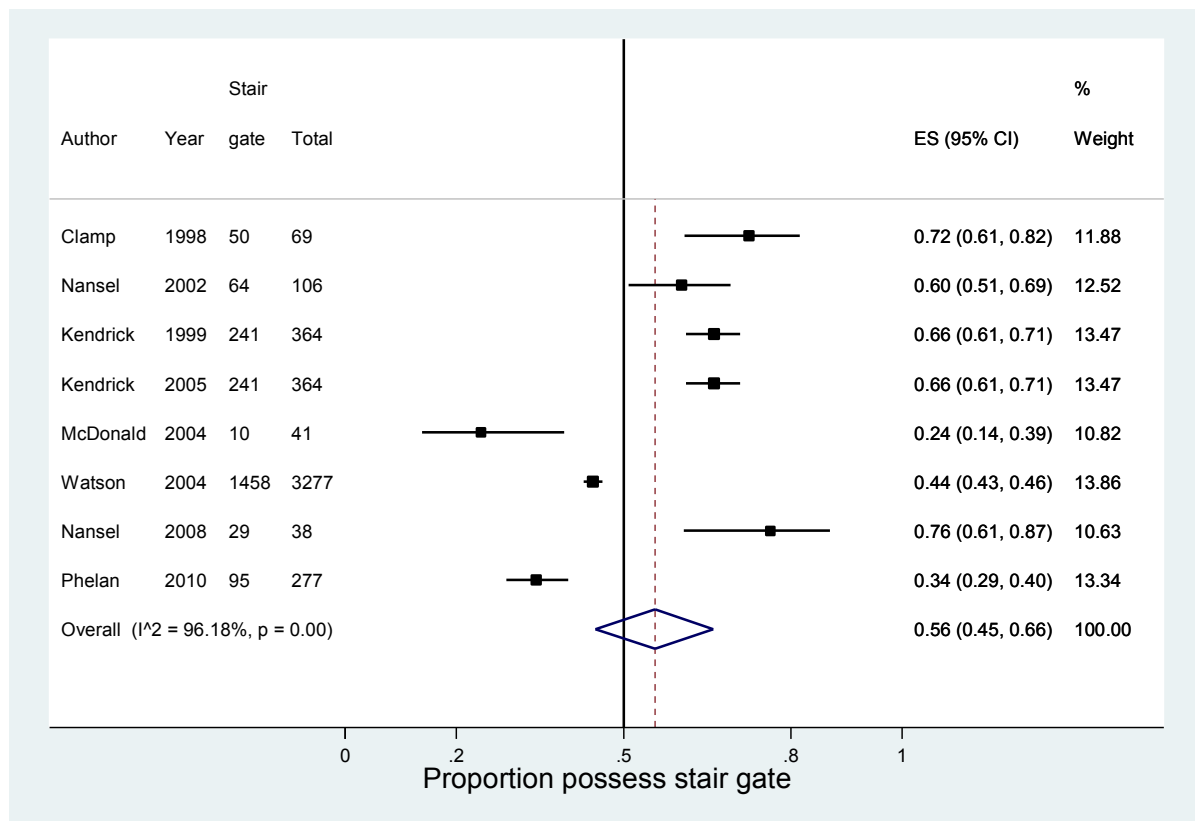
Appendix N NMA RE model combining IPD and aggregate data with a covariate – WinBUGS code

```
model {

# Model for non-clustered ipd trial data #

for(i in 1:n.non.cluster.subjects) {
        outcome[i] ~dbern(p[i])      #likelihood for non-clustered IPD data
#model
        logit(p[i])<- mu[study[i]] + delta[index[i]] * (1-equals(treat[i],baseline[i])) +
                beta_cov[study[i]] * cov[i] +
                beta.w[index[i]] * (cov[i] - meancov[i]) *(1 - equals(treat[i], baseline[i])) +
                beta.b[index[i]] * (1-equals(treat[i], baseline[i])) * meancov[i]
        temp1[i] <- id1[i]
        temp4[i]<- uniqueid[i]
        }

for(j in 1:n.ipd.non.cluster.trials) {
        beta_cov[j]~dnorm(0,1.0E-06)
        mu[j]~dnorm(0,1.0E-6)
        }


# Model for cluster ipd cluster trial data  #

for(i in 1:n.cluster.subjects) {
     temp2[i] <- id2[i]
     temp5[i] <- c.uniqueid[i]
     c.outcome[i] ~ dbern(c.p[i])
     logit(c.p[i]) <- c.mu[c.study[i],c.cluster[i]] +
       delta[c.index[i]+n.non.cluster.arms] * (1-equals(c.treat[i],c.baseline[i])) +
       c.beta_cov[c.study[i]] * c.cov[i]  +
       beta.w[c.index[i] + n.non.cluster.arms] * (c.cov[i] - c.meancov[i]) *(1-equals(c.treat[i], c.baseline[i])) +
       beta.b[c.index[i] + n.non.cluster.arms]*(1 - equals(c.treat[i], c.baseline[i])) * c.meancov[i]
                 }

for(i in 1:n.ipd.cluster.trials) {
        c.beta_cov[i] ~ dnorm(0,1.0E-6)
        for(j in 1:n.cluster.max) {
           c.mu[i,j] ~ dnorm(mu.mean[i], inv.tau.sq.mu[i])
         }
        mu.mean[i] ~ dnorm(0.0,1.0E-6)
        inv.tau.sq.mu[i] <- 1/(sigma.mu[i]*sigma.mu[i])
        sigma.mu[i] ~ dunif(0,2)
        tau.sq.mu[i] <- sigma.mu[i]*sigma.mu[i]
        }

#for all IPD
for(i in 1:(n.non.cluster.arms + n.cluster.arms)) {
        md[i] <- d[treat1[i]]-d[baseline1[i]]
        delta[i] ~ dnorm(md[i], prec)
        beta.w[i] <- bw[treat1[i]] - bw[baseline1[i]]
        beta.b[i] <- bb[treat1[i]] - bb[baseline1[i]]
}
```

Model for non-cluster and cluster aggregate data

```
for(i in 1:n.agg.arms) {
        temp3[i] <- id4[i]
        logit(pa[i])<- mu.ad[a.study[i]] +
                delta[i + n.non.cluster.arms + n.cluster.arms] * (1-equals(a.treat[i],a.base[i]))
        outcome.ad[i] ~ dbin(pa[i],n[i])
        delta[i + n.non.cluster.arms + n.cluster.arms] ~ dnorm(md.ad[i], prec)
        md.ad[i]<-d[a.treat[i]]-d[a.base[i]] + (bb[a.treat[i]] - bb[a.base[i]]) * a.cov[i]
        }

for(j in 1:n.agg.trials) {
        mu.ad[j]~dnorm(0,1.0E-6)
        }
```

Model for combining all estimates of intervention effect

```
bw[1] <- 0
bb[1] <- 0
bdiff[1]<-0
bsum[1]<-0
d[1]<-0
for (k in 2:max.treat) {
        bw[k] ~ dnorm(m.betaw, prec.betaw)
        bb[k] ~ dnorm(m.betab, prec.betab)
        d[k] ~ dnorm(0,1.0E-6)
        }

for (k in 2:max.treat) {
        bdiff[k]<- bb[k] - bw[k]
        bsum[k]<- bb[k] + bw[k]
        }

m.betaw ~ dnorm(0,1.0E-6)
tau.betaw ~ dunif(0,2)
tau.sq.betaw <- (tau.betaw*tau.betaw)
prec.betaw <- 1/(tau.sq.betaw)

m.betab ~ dnorm(0,1.0E-6)
tau.betab ~ dunif(0,2)
tau.sq.betab <- (tau.betaw * tau.betab)
prec.betab <- 1/(tau.sq.betab)

tau~dunif(0,2)
tau.sq<-tau*tau
prec<-1/(tau.sq)
```

pairwise ORs
```
for (c in 1:(max.treat - 1)) {
        for (k in (c + 1):max.treat) {
                lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
                }
        }
}
```

Appendix O Meta-analysis of baseline proportions and proportions in the control arm after intervention of households possessing fitted safety gates in the studies used in the NMA

Eight studies gave rates (ranging from 24% in McDonald 2004 to 76% in Nansel 2008, both control arm) but only three were baseline rates.

Author	Year	Stair gate	Total		ES (95% CI)	% Weight
Clamp	1998	50	69		0.72 (0.61, 0.82)	11.88
Nansel	2002	64	106		0.60 (0.51, 0.69)	12.52
Kendrick	1999	241	364		0.66 (0.61, 0.71)	13.47
Kendrick	2005	241	364		0.66 (0.61, 0.71)	13.47
McDonald	2004	10	41		0.24 (0.14, 0.39)	10.82
Watson	2004	1458	3277		0.44 (0.43, 0.46)	13.86
Nansel	2008	29	38		0.76 (0.61, 0.87)	10.63
Phelan	2010	95	277		0.34 (0.29, 0.40)	13.34
Overall (I^2 = 96.18%, p = 0.00)					0.56 (0.45, 0.66)	100.00

Proportion possess stair gate

Using WinBUGS: All eight studies including control arm

node	mean	sd	MC error	2.5%	median	97.5%
pop_sg[1]	0.5572	0.0629	7.01E-04	0.4286	0.558	0.6804
τ	4.991	2.881	0.0274	0.2333	5.015	9.759
precision	518.2	37220	370.5	0.0105	0.03977	18.46

Only three studies reported baseline possession of fitted safety gates: Phelan (34%), Watson(44%) and Nansel 2002(60%)

node	mean	sd	MC error	2.5%	median	97.5%
pop_sg[1]	0.4626	0.1888	1.39E-03	0.0529	0.4555	0.9219
τ	1.582	1.654	0.0275	0.2719	0.9757	7.053
precision	2.532	4.884	0.0581	0.0201	1.051	13.53

Author	Year	Stair gate	Total	ES (95% CI)	% Weight
Nansel	2002	64	106	0.60 (0.51, 0.69)	28.71
Watson	2004	1458	3277	0.44 (0.43, 0.46)	37.53
Phelan	2010	95	277	0.34 (0.29, 0.40)	33.76
Overall (I^2 = 91.14%, p = 0.00)				0.46 (0.35, 0.56)	100.00

Proportion possess stair gate
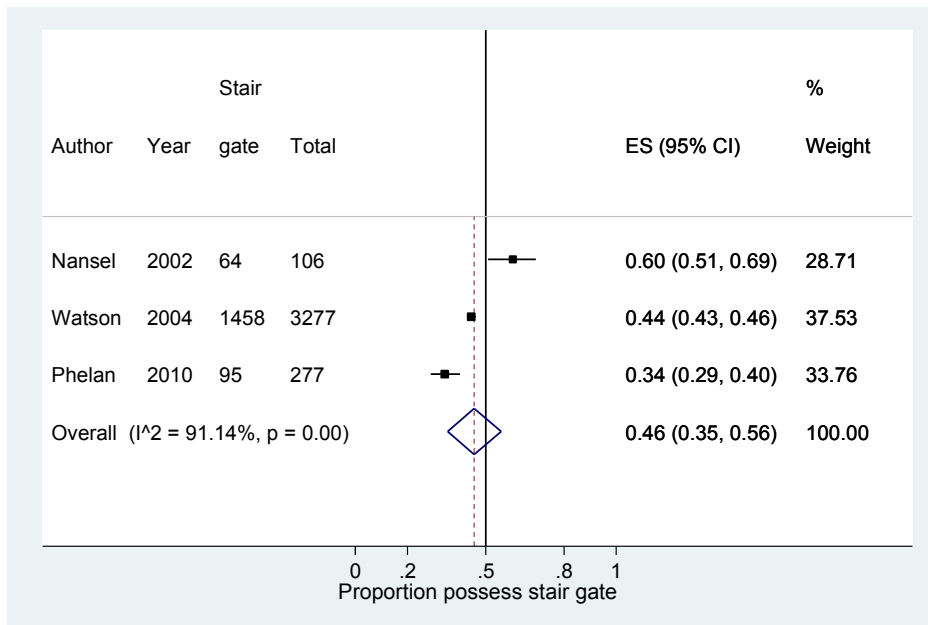
Appendix P Meta-analysis of proportions that accepted the intervention in the studies used in the NMA

Ten studies reported uptake/acceptance of the intervention
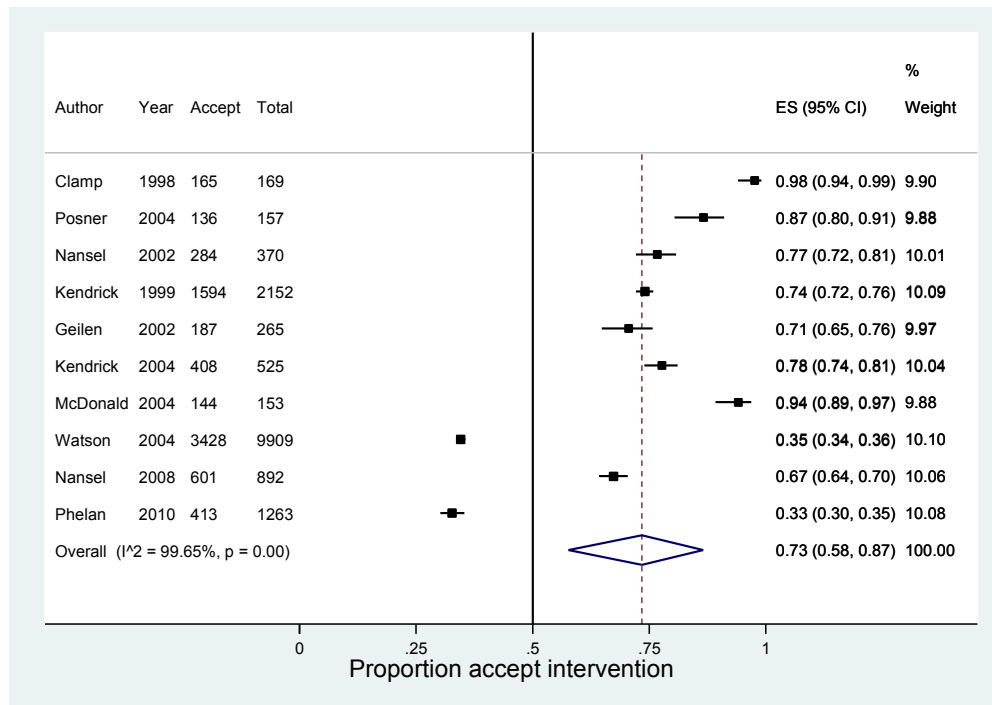
Using WinBUGS

node	mean	sd	MC error	2.5%	median	97.5%
p_accept[1]	0.7532	0.09451	8.80E-04	0.5311	0.7655	0.902
τ	1.576	0.5036	0.00861	0.9201	1.469	2.854
precision	0.5102	0.2768	0.003767	0.1229	0.4632	1.182

Appendix Q Decision Model WinBUGS code

model{

#DECISION MODEL for preventing falls down stairs

```
# NHS Perspective
# S=Health states (1=Safety gate (SG),2=NO SG,3=SG/disability,4=NO SG/disability
# 5=death from fatal fall injury and 6=Death other causes
# N=Number of households
# C=Cycle
# T=Total number of years (time horizon)
# P=INTERVENTIONS
# 1Usual care
# 2Education
# 3Education + low cost/free equipment
# 4Education + low cost/free equipment + Home safety inspection
# 5Education + low cost/free equipment + Fitting
# 6Education + Home safety inspection
# 7Education + low cost/free equipment +Home safety inspection + Fitting
```

#STAGE ONE

#Network Meta-Analysis model
#Random effects model for multi-arm trials (any number of arms)

```
for(i in 1:ns){                        # LOOP THROUGH ns STUDIES
        temp1[i] <- id[i]
 w[i,1] <- 0                # adjustment for multi-arm trials is zero for control arm
 delta[i,1] <- 0                        # intervention effect is zero for control arm
 mu[i] ~ dnorm(0,.0001)                 # vague priors for all trial baselines

 for (k in 1:na[i]) {                        # LOOP THROUGH na ARMS
  rMTC[i,k] ~ dbin(pMTC[i,k],nMTC[i,k])     # binomial likelihood
  logit(pMTC[i,k]) <- mu[i] + delta[i,k]    # model for linear predictor
  rhat[i,k] <- pMTC[i,k] * nMTC[i,k]        # expected value of the numerators
#Deviance contribution
        dev[i,k] <- 2 * (rMTC[i,k] * (log(rMTC[i,k])-log(rhat[i,k]))+(nMTC[i,k]-
                       rMTC[i,k]) * (log(nMTC[i,k]-rMTC[i,k]) - log(nMTC[i,k]-rhat[i,k])))
        }
 resdev[i] <- sum(dev[i,1:na[i]])       # summed res deviance contribution for trial

 for (k in 2:na[i]) {                        # LOOP THROUGH ARMS
   delta[i,k] ~ dnorm(md[i,k],taud[i,k])    # trial-specific LOR distributions
   md[i,k] <- d[t[i,k]] - d[t[i,1]] + sw[i,k]  # mean of LOR dist(with multi-arm correction)
   taud[i,k] <- tau *2*(k-1)/k              # precision of LOR dists(with multi-arm correction)
   w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])  # adjustment for multi-arm RCTs
   sw[i,k] <- sum(w[i,1:k-1])/(k-1)         # cumulative adjustment for multi-arm trials
        }
}

totresdev <- sum(resdev[])              #Total Residual Deviance
d[1]<- 0                                # intervention effect is zero for reference intervention
for (k in 2:nt) { d[k] ~ dnorm(0,.0001)}    # vague priors for intervention effects
sd ~ dunif(0,2)
tau <- pow(sd,-2)
```

```
# Intervention A baseline, based on  average of the 8 trials including it.
for (i in 1:ns) { mu1[i] <- mu[i] * equals(t[i,1],1) }
for (k in 1:nt)  { logit(p_MTC[k])<- sum(mu1[])/8 +d[k] }

mn.mu1<-sum(mu1[])/8

#Rank the intervention effects (with 1=best) & record the best intervention
for (k in 1:nt) { rk[k]<-nt+1 - rank(p_MTC[],k)
               best[k]<-equals(rk[k],1)}

#All pairwise log odds ratios and odds ratios
for (c in 1:(nt-1))
    { for (k in (c+1):nt)
         { lor[c,k] <- d[k] - d[c]
           log(or[c,k]) <- lor[c,k]
           RR[c,k]<-or[c,k]/(1-mn.mu1+mn.mu1*or[c,k])
         }
    }
#} ends model for NMA

#Proportion of population that have a safety gate (pop_sg)
#results from M-A of baseline and control group data

    for( i in 1 : Nstudbase ) {
        rbase[i] ~ dbin(pbase[i], nbase[i])
        logit(pbase[i]) <- mubase[i]
        mubase[i] ~ dnorm(mnbase,taubase)
        tmp1[i]<-idbase[i]
    }
    mnbase ~ dnorm(0.0,1.0E-6)
    pop_sg[1]<-(exp(mnbase))/(1+exp(mnbase))
    taubase<-1/(sigmabase*sigmabase)       #precision
    sigmabase~dunif(0,10)                          #sd

#Proportion who agree to participate, assume the same for all interventions (p_accept)
#results from M-A of studies in KCS systematic review

    for( i in 1 : Nstudaccept ) {
        raccept[i] ~ dbin(paccept[i], naccept[i])
        logit(paccept[i]) <- muaccept[i]
        muaccept[i] ~ dnorm(mnaccept,tauaccept)
        tmp2[i]<-idaccept[i]
    }
    mnaccept ~ dnorm(0.0,1.0E-6)
    p_accept[1]<-(exp(mnaccept))/(1+exp(mnaccept))
    tauaccept<-1/(sigmaaccept*sigmaaccept)          #precision
    sigmaaccept~dunif(0,10)                     #sd

#Costs of each intervention P (at time point c=1).
for (g in 1:P){
#Costs of different interventions - intervention safety gates only
c_interv[g]<- c_educ[g] + c_equipgiveaway[g]*n_units+c_hsi[g] +c_install[g]*n_units +c_travel[g]
#cost of the intervention
c_n1[g]<-n1[g]*0
#cost=0, have safety gates prior to intervention (no intervention)
c_n2[g]<-n2[g]*(c_interv[g] + c_acc)
#accept intervention (have safety gates after intervention)
```

```
c_n3[g]<-n3[g]*(c_interv[g] + c_acc)
#accept intervention (have NO safety gates after intervention)
c_n4[g]<-n4[g]*0
#decline intervention (have NO safety gates)

ct[1,g]<-c_fixed[g]+c_n1[g]+c_n2[g]+c_n3[g]+c_n4[g]   #total cost for the intervention

ut[1,g]<-u_pop[1]*N                                          #total QALYs for the intervention
        }

n1[1]<-N*pop_sg[1]                #Baseline have fitted safety gates using results from M-A
n2[1]<-0
n3[1]<-0
n4[1]<-N*(1-pop_sg[1])      #Baseline have NO fitted stair gate

for (i in 2:P){                                    # for i = 2 to 7 i.e. different interventions
    n1[i]<-N*pop_sg[1]              #have fitted safety gates prior to intervention
    n2[i]<-N*(1-pop_sg[1])*p_accept[1]*p_MTC[i]        #have  fitted safety gates after intervention
    n3[i]<-N*(1-pop_sg[1])*p_accept[1]*(1-p_MTC[i])  #have NO fitted safety gates after intervention
    n4[i]<-N*(1-pop_sg[1])*(1-p_accept[1])    #have NO fitted safety gates, refuse intervention
        }
```

#Utility for non-injured population
#Using data1 mnu_pop, u_pop.se: general background mean utility for a non-injured population
#from Kind et al, uncertainty on population norms in age groups: <25, 25-34, ...75+
#(& p_allcause all cause mortality estimates for the UK pop (ONS) only temp variable.

```
for(l in 1:T){                                          #T is 100 years follow-up
        u_pop[l]~dnorm(mnu_pop[l], prec_pop[l])
        prec_pop[l]<-1/(u_pop.se[l]*u_pop.se[l])
        tmpall[l]<-p_allcause[l]
                }
```

#STAGE TWO

#Calculate the probability of being in state i for intervention j in cycle 1 (child age 0-1)
```
for (j in 1:P){
#cycle 1, state 1 to 6, intervention 1 to P (states are for Part 2 Markov model "pre-school")
        pi[1,1,j]<-n1[j]+n2[j]                  #state 1=fitted safety gates
        pi[1,2,j]<-n3[j]+n4[j]                  #state 2=NO fitted safety gates
        pi[1,3,j]<-0                            #state 3=fitted safety gates/disability,
        pi[1,4,j]<-0                            #state 4=NO fitted safety gates/disability
        pi[1,5,j]<-0                            #state 5=death fatal injury
        pi[1,6,j]<-0                            #state 6=death other causes

        CHECK[1,j]<-pi[1,1,j]+pi[1,2,j]+pi[1,3,j]+pi[1,4,j]+pi[1,5,j]+pi[1,6,j]     #Should sum to N
        }
```

#Probability of taking each pathway through the model

#Prob minor (not admitted), moderate (admitted for observation) or severe (admitted overnight)
#multinomial distribution for injury probability distributions
#Dirichlet distribution is a multivariate generalisation of the beta distribution
#prior constrains all probabilities to be [0,1] and total 1.

```
r_injury[1:4]~dmulti(p_injuryseverity[1:4],n_injury)
p_injuryseverity[1:4]~ddirch(alpha[])
for(r in 1:4){
        alpha[r]<-1
}
```

```
#Adjusted OR from CC study for falls down stairs
logOR_sgfall~dnorm(logOR_sg, taulogOR_sg)
taulogOR_sg<-pow(selogOR_sg,-2)
OR_sgfall<-exp(logOR_sgfall)

rfallstairs~dnorm(rfallST, taufallST)
taufallST<-pow(sefallST, -2)

p_fall[1]<-(rfallstairs/n04yrs)*OR_sgfall
p_fall[2]<-(rfallstairs/n04yrs)                          #*1/OR_sgfall use
p_fall[3]<-p_fall[1]
p_fall[4]<-p_fall[2]

#Probability of ending cycle with safety gate
for(p in 1:P){
    p_func[1,p]~dunif(0.9,1)        #Prob at end of cycle with a fall of staying with functional safety gates
    p_func[2,p]~dunif(0.5,0.62)     #Prob at end of cycle  with a fall of staying with no safety gates
    p_func[3,p]~dunif(0.9,1)
    p_func[4,p]~dunif(0.5,0.62)
    p_funcnofall[1,p]<-1            #Prob at end of cycle of staying with functional safety gates
    p_funcnofall[2,p]<-1            #Prob at end of cycle of staying with no safety gates
    p_funcnofall[3,p]<-1
    p_funcnofall[4,p]<-1

#Pathways
for(c in 2:C){                                  #cycles
    for(k in 1:4){        #1=functional SG, 2=no functional SG, 3=functional SG / disability,
                          #4=no functional SG / disability
        o1[c,k,p]<-(1-p_fall[k])*p_allcause[c]
        o2[c,k,p]<-(1-p_fall[k])*(1-p_allcause[c])*p_funcnofall[k,p]                #stay in func state
        o3[c,k,p]<-(1-p_fall[k])*(1-p_allcause[c])*(1-p_funcnofall[k,p])    #move from func state

        o4[c,k,p]<-p_fall[k]*p_fatal[k]
        o5[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[1]*(1-p_allcause[c])*p_func[k,p]
        o6[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[1]*(1-p_allcause[c])*(1-p_func[k,p])

        o7[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[1]*p_allcause[c]
        o8[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[2]*(1-p_allcause[c])*p_func[k,p]
        o9[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[2]*(1-p_allcause[c])*(1-p_func[k,p])

        o10[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[2]*p_allcause[c]
        o11[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[3]*(1-p_allcause[c])*p_func[k,p]
        o12[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[3]*(1-p_allcause[c])*(1-p_func[k,p])

        o13[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[3]*p_allcause[c]
        o14[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[4]*(1-p_allcause[c])*p_func[k,p]
        o15[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[4]*(1-p_allcause[c])*(1-p_func[k,p])
        o16[c,k,p]<-p_fall[k]*(1-p_fatal[k])*p_injuryseverity[4]*p_allcause[c]

TOT[c,k,p]<-o1[c,k,p] +o2[c,k,p] +o3[c,k,p] +o4[c,k,p] +o5[c,k,p] +o6[c,k,p] +o7[c,k,p] +o8[c,k,p] +o9[c,k,p]
+o10[c,k,p] +o11[c,k,p] +o12[c,k,p]+o13[c,k,p]+o14[c,k,p]+o15[c,k,p]+o16[c,k,p]
#Check sums to 1
                }       #k
          }             #c
}                       #p
```

```
#Cycle length 1 year, run model for ages 1 to 3 year olds
#Transition probabilities for cycles 1 to 3

for (p in 1:P){
    for(c in 2:C){
    #From 'functioning' state
        lambda[c,1,1,p]<-o2[c,1,p]+o5[c,1,p]+o8[c,1,p]+o11[c,1,p]  #to func
        lambda[c,1,2,p]<-o3[c,1,p]+o6[c,1,p]+o9[c,1,p]+o12[c,1,p]  #to non-func
        lambda[c,1,3,p]<-o14[c,1,p]#to func disabled
        lambda[c,1,4,p]<-o15[c,1,p]#to non-func disabled
        lambda[c,1,5,p]<-o4[c,1,p]     #to fatal
        lambda[c,1,6,p]<-o1[c,1,p]+o7[c,1,p]+o10[c,1,p]+o13[c,1,p]+o16[c,1,p]  #to all-cause
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,1,p]<-lambda[c,1,1,p] +lambda[c,1,2,p] +lambda[c,1,3,p] +lambda[c,1,4,p]
                +lambda[c,1,5,p] +lambda[c,1,6,p]

    #From 'non-functioning' state
        lambda[c,2,1,p]<-o2[c,2,p]+o5[c,2,p]+o8[c,2,p]+o11[c,2,p]  #to func state
        lambda[c,2,2,p]<-o3[c,2,p]+o6[c,2,p]+o9[c,2,p]+o12[c,2,p]  #to non-func
        lambda[c,2,3,p]<-o14[c,2,p]  #to func disabled
        lambda[c,2,4,p]<-o15[c,2,p]  #to non-func disabled
        lambda[c,2,5,p]<-o4[c,2,p]     #to fatal
        lambda[c,2,6,p]<-o1[c,2,p]+o7[c,2,p]+o10[c,2,p]+o13[c,2,p]+o16[c,2,p]   #to all-cause
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,2,p]<-lambda[c,2,1,p] +lambda[c,2,2,p] +lambda[c,2,3,p] +lambda[c,2,4,p]
                +lambda[c,2,5,p] +lambda[c,2,6,p]

    #From 'functioning / disability' state
        lambda[c,3,1,p]<-0
        lambda[c,3,2,p]<-0
        lambda[c,3,3,p]<-o2[c,3,p]+o5[c,3,p]+o8[c,3,p]+o11[c,3,p]+o14[c,3,p]
        lambda[c,3,4,p]<-o3[c,3,p]+o6[c,3,p]+o9[c,3,p]+o12[c,3,p]+o15[c,3,p]
        lambda[c,3,5,p]<-o4[c,3,p]
        lambda[c,3,6,p]<-o1[c,3,p]+o7[c,3,p]+o10[c,3,p]+o13[c,3,p]+o16[c,3,p]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,3,p]<-lambda[c,3,1,p] +lambda[c,3,2,p] +lambda[c,3,3,p] +lambda[c,3,4,p]
                +lambda[c,3,5,p] +lambda[c,3,6,p]


    #From 'non-functioning/ disability' state
        lambda[c,4,1,p]<-0
        lambda[c,4,2,p]<-0
        lambda[c,4,3,p]<- o2[c,4,p]+o5[c,4,p]+o8[c,4,p]+ o11[c,4,p]+o14[c,4,p]
        lambda[c,4,4,p]<-o3[c,4,p]+o6[c,4,p]+o9[c,4,p]+ o12[c,4,p]+o15[c,4,p]
        lambda[c,4,5,p]<-o4[c,4,p]
        lambda[c,4,6,p]<-o1[c,4,p]+o7[c,4,p]+o10[c,4,p]+o13[c,4,p]+o16[c,4,p]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,4,p]<-lambda[c,4,1,p]+lambda[c,4,2,p]+lambda[c,4,3,p]+ lambda[c,4,4,p]
                +lambda[c,4,5,p]+lambda[c,4,6,p]

    #From 'fatal' state
        lambda[c,5,1,p]<-0
        lambda[c,5,2,p]<-0
        lambda[c,5,3,p]<-0
        lambda[c,5,4,p]<-0
        lambda[c,5,5,p]<-1
        lambda[c,5,6,p]<-0
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,5,p]<-lambda[c,5,1,p] +lambda[c,5,2,p] +lambda[c,5,3,p] +lambda[c,5,4,p]
                +lambda[c,5,5,p] +lambda[c,5,6,p]

    #From 'all cause' state
        lambda[c,6,1,p]<-0
        lambda[c,6,2,p]<-0
        lambda[c,6,3,p]<-0
        lambda[c,6,4,p]<-0
```

225

```
                    lambda[c,6,5,p]<-0
                    lambda[c,6,6,p]<-1
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,6,p]<-lambda[c,6,1,p] +lambda[c,6,2,p] +lambda[c,6,3,p] +lambda[c,6,4,p]
                    +lambda[c,6,5,p] +lambda[c,6,6,p]
                    }
            }

#Number of individuals in each state at time t>1
for (p in 1:P){
    for(c in 2:C){
        for (s in 1:S){
            pi[c,s,p]<-inprod(pi[(c-1),,p],lambda[c, ,s,p])
            }

CHECK[c,p]<-pi[c,1,p]+pi[c,2,p]+pi[c,3,p]+pi[c,4,p]+pi[c,5,p]+pi[c,6,p]  #Check sums to N
        }
}
```

#Costs in each state
```
#Usual care
#1=functional, 2=non-functional, 3=functional / disability, 4=non-functional / disability

for(k in 1:4){
        c_equip[1,k]<-0
}
c_stairgate<-0                 #Use when not including private costs otherwise add stairgate cost

# All other intervention groups
for (p in 2:P){                           #intervention P
#Equipment costs of having fitted stair gate at end of each cycle
        c_equip[p,1]<- 0
        c_equip[p,2]<- c_stairgate
        c_equip[p,3]<- c_equip[p,1]
        c_equip[p,4]<- c_equip[p,2]
}

#Cost of disability per year
c_dispyr~dgamma(c_dispyra, c_dispyrb)

#Minor
var_minor<-se_minor*se_minor
mn_minorsq<-mn_minor*mn_minor
mu.minor<-log(mn_minor)-0.5*log(1+var_minor/mn_minorsq)
sigmasq.minor<-log(1+var_minor/mn_minorsq)
prec.minor<-1/sigmasq.minor
c_minor~dlnorm(mu.minor, prec.minor)

#Moderate
var_mod<-se_moderate*se_moderate
mn_modsq<-mn_moderate*mn_moderate
mu.mod<-log(mn_moderate)-0.5*log(1+var_mod/mn_modsq)
sigmasq.mod<-log(1+var_mod/mn_modsq)
prec.mod<-1/sigmasq.mod
c_moderate~dlnorm(mu.mod, prec.mod)
#Severe
var_sev<-se_severe*se_severe
mn_sevsq<-mn_severe*mn_severe
mu.sev<-log(mn_severe)-0.5*log(1+var_sev/mn_sevsq)
sigmasq.sev<-log(1+var_sev/mn_sevsq)
prec.sev<-1/sigmasq.sev
c_severe~dlnorm(mu.sev, prec.sev)

#Short stay
var_shortstay<-se_shortstay*se_shortstay
mn_shortstaysq<-mn_shortstay*mn_shortstay
```

```
mu.shortstay<-log(mn_shortstay)-0.5*log(1+var_shortstay/mn_shortstaysq)
sigmasq.shortstay<-log(1+var_shortstay/mn_shortstaysq)
prec.shortstay<-1/sigmasq.shortstay
c_shortstay~dlnorm(mu.shortstay, prec.shortstay)

#Long stay
var_longstay<-se_longstay*se_longstay
mn_longstaysq<-mn_longstay*mn_longstay
mu.longstay<-log(mn_longstay)-0.5*log(1+var_longstay/mn_longstaysq)
sigmasq.longstay<-log(1+var_longstay/mn_longstaysq)
prec.longstay<-1/sigmasq.longstay
c_longstay~dlnorm(mu.longstay, prec.longstay)

#Ambulance
var_ambulance<-se_ambulance*se_ambulance
mn_ambulancesq<-mn_ambulance*mn_ambulance
mu.ambulance<-log(mn_ambulance)-0.5*log(1+var_ambulance/mn_ambulancesq)
sigmasq.ambulance<-log(1+var_ambulance/mn_ambulancesq)
prec.ambulance<-1/sigmasq.ambulance
c_ambulance~dlnorm(mu.ambulance, prec.ambulance)

for (p in 1:P){                                    #intervention
    for (c in 2:C){                    #cycle
        for(k in 1:2){                             #1=functional, 2=non-functional
            c_o1[c,p,k]<-o1[c,k,p]*c_nofall
            c_o2[c,p,k]<-o2[c,k,p]*c_nofall
            c_o3[c,p,k]<-o3[c,k,p]*c_nofall
            c_o4[c,p,k]<-o4[c,k,p]*c_fatal
            c_o5[c,p,k]<-o5[c,k,p]*(c_minor+(c_ambulance*p_amb))
            c_o6[c,p,k]<-o6[c,k,p]*(c_minor+(c_ambulance*p_amb))
            c_o7[c,p,k]<-o7[c,k,p]*(c_minor+(c_ambulance*p_amb))

            c_o8[c,p,k]<-o8[c,k,p]*(c_moderate+c_shortstay+(c_ambulance*p_amb))
            c_o9[c,p,k]<-o9[c,k,p]*(c_moderate+c_shortstay+(c_ambulance*p_amb))
            c_o10[c,p,k]<-o10[c,k,p]*(c_moderate+c_shortstay+(c_ambulance*p_amb))

            c_o11[c,p,k]<-o11[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))
            c_o12[c,p,k]<-o12[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))
            c_o13[c,p,k]<-o13[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))

            c_o14[c,p,k]<-o14[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))
            c_o15[c,p,k]<-o15[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))
            c_o16[c,p,k]<-o16[c,k,p]*(c_severe+c_longstay+(c_ambulance*p_amb))
        }          #k

        for(k in 3:4){          #3=functional / disability, 4=non-functional / disability
            c_o1[c,p,k]<-o1[c,k,p]*c_nofall
            c_o2[c,p,k]<-o2[c,k,p]*(c_nofall+c_dispyr)
            c_o3[c,p,k]<-o3[c,k,p]*(c_nofall+c_dispyr)
            c_o4[c,p,k]<-o4[c,k,p]*c_fatal

            c_o5[c,p,k]<-o5[c,k,p]*(c_minor +(c_ambulance*p_amb)+ c_dispyr)
            c_o6[c,p,k]<-o6[c,k,p]*(c_minor +(c_ambulance*p_amb)+ c_dispyr)
            c_o7[c,p,k]<-o7[c,k,p]*c_minor+(c_ambulance*p_amb)

            c_o8[c,p,k]<-o8[c,k,p]*(c_moderate +c_shortstay+(c_ambulance*p_amb) + c_dispyr)
            c_o9[c,p,k]<-o9[c,k,p]*(c_moderate +c_shortstay+(c_ambulance*p_amb) + c_dispyr)
            c_o10[c,p,k]<-o10[c,k,p]*c_moderate+c_shortstay+(c_ambulance*p_amb)

            c_o11[c,p,k]<-o11[c,k,p]*(c_severe +c_longstay+(c_ambulance*p_amb)+ c_dispyr)
            c_o12[c,p,k]<-o12[c,k,p]*(c_severe +c_longstay+(c_ambulance*p_amb)+ c_dispyr)
            c_o13[c,p,k]<-o13[c,k,p]*c_severe+c_longstay+(c_ambulance*p_amb)

            c_o14[c,p,k]<-o14[c,k,p]*(c_severe +c_longstay+(c_ambulance*p_amb)+ c_dispyr)
            c_o15[c,p,k]<-o15[c,k,p]*(c_severe +c_longstay+(c_ambulance*p_amb)+ c_dispyr)
            c_o16[c,p,k]<-o16[c,k,p]*c_severe+c_longstay+(c_ambulance*p_amb)
```

```
                }                      #k

cost[c,1,p]<-
    c_o1[c,p,1]+c_o2[c,p,1]+c_o3[c,p,1]+c_o4[c,p,1]+c_o5[c,p,1]+c_o6[c,p,1]+c_o7[c,p,1]+c_o8[c,p,1]+c_o9[c,
    p,1]+c_o10[c,p,1]+c_o11[c,p,1]+c_o12[c,p,1]+c_o13[c,p,1]+c_o14[c,p,1]+c_o15[c,p,1]+c_o16[c,p,1]

cost[c,2,p]<-
    c_o1[c,p,2]+c_o2[c,p,2]+c_o3[c,p,2]+c_o4[c,p,2]+c_o5[c,p,2]+c_o6[c,p,2]+c_o7[c,p,2]+c_o8[c,p,2]+c_o9[c,
    p,2]+c_o10[c,p,2]+c_o11[c,p,2]+c_o12[c,p,2]+c_o13[c,p,2]+c_o14[c,p,2]+c_o15[c,p,2]+c_o16[c,p,2]

cost[c,3,p]<-
    c_o1[c,p,3]+c_o2[c,p,3]+c_o3[c,p,3]+c_o4[c,p,3]+c_o5[c,p,3]+c_o6[c,p,3]+c_o7[c,p,3]+c_o8[c,p,3]+c_o9[c,
    p,3]+c_o10[c,p,3]+c_o11[c,p,3]+c_o12[c,p,3]+c_o13[c,p,3]+c_o14[c,p,3]+c_o15[c,p,3]+c_o16[c,p,3]

cost[c,4,p]<-
    c_o1[c,p,4]+c_o2[c,p,4]+c_o3[c,p,4]+c_o4[c,p,4]+c_o5[c,p,4]+c_o6[c,p,4]+c_o7[c,p,4]+c_o8[c,p,4]+c_o9[c,
    p,4]+c_o10[c,p,4]+c_o11[c,p,4]+c_o12[c,p,4]+c_o13[c,p,4]+c_o14[c,p,4]+c_o15[c,p,4]+c_o16[c,p,4]

cost[c,5,p]<-0

cost[c,6,p]<-0
            }            #c

for (c in C+1:T){
        cost[c,1,p]<- 0
        cost[c,2,p]<- 0
        cost[c,3,p]<- c_dispyr
        cost[c,4,p]<- c_dispyr
        cost[c,5,p]<- 0
        cost[c,6,p]<- 0
         }            #c
}                    #p


#Utilities in each state - using Uniform distribution
u_deficit~dbeta(u_deficita, u_deficitb)

u_Min~dunif(0,0.1)
u_Sev~dunif(0.1,0.3)

for (p in 1:P){
    u[1,1,p]<-u_pop[1]
    u[1,2,p]<-u_pop[1]
    u[1,3,p]<-u_pop[1]
    u[1,4,p]<-u_pop[1]
    u[1,5,p]<-0
    u[1,6,p]<-0

    for (c in 2:C){
        for(k in 1:2){          #1=functional SG, 2=non-functional SG
            u_o1[c,p,k]<-o1[c,k,p]*0
            u_o2[c,p,k]<-o2[c,k,p]*u_pop[c]
            u_o3[c,p,k]<-o3[c,k,p]*u_pop[c]
            u_o4[c,p,k]<-o4[c,k,p]*0

            u_o5[c,p,k]<-o5[c,k,p]*(u_pop[c]-u_Min)
            u_o6[c,p,k]<-o6[c,k,p]*(u_pop[c]-u_Min)
            u_o7[c,p,k]<-o7[c,k,p]*0

            u_o8[c,p,k]<-o8[c,k,p]*(u_pop[c]-u_Mod)
            u_o9[c,p,k]<-o9[c,k,p]*(u_pop[c]-u_Mod)
            u_o10[c,p,k] <- o10[c,k,p]*0

            u_o11[c,p,k] <- o11[c,k,p]*(u_pop[c]-u_Sev)
            u_o12[c,p,k] <- o12[c,k,p]*(u_pop[c]-u_Sev)
```

```
                    u_o13[c,p,k] <- o13[c,k,p]*0

                    u_o14[c,p,k] <-o14[c,k,p]*(u_pop[c]-u_Sev)
                    u_o15[c,p,k] <-o15[c,k,p]*(u_pop[c]-u_Sev)
                    u_o16[c,p,k]<-o16[c,k,p]*0
                            }           #k

            for(k in 3:4){          #3=functional SG / disability, 4=non-functional SG / disability
                u_o1[c,p,k]<-o1[c,k,p]*0
                u_o2[c,p,k]<-o2[c,k,p]*(u_pop[c]-u_deficit)
                u_o3[c,p,k]<-o3[c,k,p]*(u_pop[c]-u_deficit)
                u_o4[c,p,k]<-o4[c,k,p]*0

                u_o5[c,p,k]<-o5[c,k,p]*((u_pop[c]-u_deficit)-u_Min)
                u_o6[c,p,k]<-o6[c,k,p]*((u_pop[c]-u_deficit)-u_Min)
                u_o7[c,p,k]<-o7[c,k,p]*0

                u_o8[c,p,k]<-o8[c,k,p]*((u_pop[c]-u_deficit)-u_Mod)
                u_o9[c,p,k]<-o9[c,k,p]*((u_pop[c]-u_deficit)-u_Mod)
                u_o10[c,p,k] <- o10[c,k,p]*0

                u_o11[c,p,k] <- o11[c,k,p]*((u_pop[c]-u_deficit)-u_Sev)
                u_o12[c,p,k] <- o12[c,k,p]*((u_pop[c]-u_deficit)-u_Sev)
                u_o13[c,p,k] <- o13[c,k,p]*0

                u_o14[c,p,k] <-o14[c,k,p]*((u_pop[c]-u_deficit)-u_Sev)
                u_o15[c,p,k] <-o15[c,k,p]*((u_pop[c]-u_deficit)-u_Sev)
                u_o16[c,p,k]<-o16[c,k,p]*0
                            }           #k

u[c,1,p]<-
    u_o1[c,p,1]+u_o2[c,p,1]+u_o3[c,p,1]+u_o4[c,p,1]+u_o5[c,p,1]+u_o6[c,p,1]+u_o7[c,p,1]+u_o8[c,p,1]+u_o9
    [c,p,1]+u_o10[c,p,1]+u_o11[c,p,1]+u_o12[c,p,1]+u_o13[c,p,1]+u_o14[c,p,1]+u_o15[c,p,1]+u_o16[c,p,1]

u[c,2,p]<-
    u_o1[c,p,2]+u_o2[c,p,2]+u_o3[c,p,2]+u_o4[c,p,2]+u_o5[c,p,2]+u_o6[c,p,2]+u_o7[c,p,2]+u_o8[c,p,2]+u_o9
    [c,p,2]+u_o10[c,p,2]+u_o11[c,p,2]+u_o12[c,p,2]+u_o13[c,p,2]+u_o14[c,p,2]+u_o15[c,p,2]+u_o16[c,p,2]

u[c,3,p]<-
    u_o1[c,p,3]+u_o2[c,p,3]+u_o3[c,p,3]+u_o4[c,p,3]+u_o5[c,p,3]+u_o6[c,p,3]+u_o7[c,p,3]+u_o8[c,p,3]+u_o9
    [c,p,3]+u_o10[c,p,3]+u_o11[c,p,3]+u_o12[c,p,3]+u_o13[c,p,3]+u_o14[c,p,3]+u_o15[c,p,3]+u_o16[c,p,3]

u[c,4,p]<-
    u_o1[c,p,4]+u_o2[c,p,4]+u_o3[c,p,4]+u_o4[c,p,4]+u_o5[c,p,4]+u_o6[c,p,4]+u_o7[c,p,4]+u_o8[c,p,4]+u_o9
    [c,p,4]+u_o10[c,p,4]+u_o11[c,p,4]+u_o12[c,p,4]+u_o13[c,p,4]+u_o14[c,p,4]+u_o15[c,p,4]+u_o16[c,p,4]

u[c,5,p]<-0

u[c,6,p]<-0
                    }                   #c

        for (c in C+1:T){
            u[c,1,p]<- u_pop[c]
            u[c,2,p]<- u_pop[c]
            u[c,3,p]<- u_pop[c] - u_deficit                 #u_deficit - utility lost due to disability
            u[c,4,p]<- u_pop[c] - u_deficit
            u[c,5,p]<- 0
            u[c,6,p]<- 0
                    }               #c
}                               #p
```

#STAGE THREE

```
#Transition matrix for cycles 4 to 100
for (p in 1:P){
    for(c in C+1:T){
            lambda[c,1,1,p]<-1-p_allcause[c]
            lambda[c,1,2,p]<-0
            lambda[c,1,3,p]<-0
            lambda[c,1,4,p]<-0
            lambda[c,1,5,p]<-0
            lambda[c,1,6,p]<-p_allcause[c]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,1,p]<-lambda[c,1,1,p] +lambda[c,1,2,p] +lambda[c,1,3,p] +lambda[c,1,4,p] +lambda[c,1,5,p]
    +lambda[c,1,6,p]


            lambda[c,2,1,p]<-0
            lambda[c,2,2,p]<-1-p_allcause[c]
            lambda[c,2,3,p]<-0
            lambda[c,2,4,p]<-0
            lambda[c,2,5,p]<-0
            lambda[c,2,6,p]<-p_allcause[c]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,2,p]<-lambda[c,2,1,p] +lambda[c,2,2,p] +lambda[c,2,3,p] +lambda[c,2,4,p] +lambda[c,2,5,p]
    +lambda[c,2,6,p]


            lambda[c,3,1,p]<-0
            lambda[c,3,2,p]<-0
            lambda[c,3,3,p]<-1-p_allcause[c]
            lambda[c,3,4,p]<-0
            lambda[c,3,5,p]<-0
            lambda[c,3,6,p]<-p_allcause[c]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,3,p]<-lambda[c,3,1,p] +lambda[c,3,2,p] +lambda[c,3,3,p] +lambda[c,3,4,p] +lambda[c,3,5,p]
    +lambda[c,3,6,p]


            lambda[c,4,1,p]<-0
            lambda[c,4,2,p]<-0
            lambda[c,4,3,p]<-0
            lambda[c,4,4,p]<-1-p_allcause[c]
            lambda[c,4,5,p]<-0
            lambda[c,4,6,p]<-p_allcause[c]
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,4,p]<-lambda[c,4,1,p] +lambda[c,4,2,p] +lambda[c,4,3,p] +lambda[c,4,4,p] +lambda[c,4,5,p]
    +lambda[c,4,6,p]


            lambda[c,5,1,p]<-0
            lambda[c,5,2,p]<-0
            lambda[c,5,3,p]<-0
            lambda[c,5,4,p]<-0
            lambda[c,5,5,p]<-1
            lambda[c,5,6,p]<-0
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,5,p]<-lambda[c,5,1,p] +lambda[c,5,2,p] +lambda[c,5,3,p] +lambda[c,5,4,p] +lambda[c,5,5,p]
    +lambda[c,5,6,p]


            lambda[c,6,1,p]<-0
            lambda[c,6,2,p]<-0
            lambda[c,6,3,p]<-0
            lambda[c,6,4,p]<-0
            lambda[c,6,5,p]<-0
            lambda[c,6,6,p]<-1
#Check each row of the transition matrix (lambda) sums to 1
TOTAL[c,6,p]<-lambda[c,6,1,p] +lambda[c,6,2,p] +lambda[c,6,3,p] +lambda[c,6,4,p] +lambda[c,6,5,p]
    +lambda[c,6,6,p]
            }                       #c
}                                   #p
```

```
#Number of individuals in each state at time >C
for (p in 1:P){
    for(c in C+1:T){
        for (s in 1:S){
            pi[c,s,p]<-inprod(pi[(c-1),,p],lambda[c, ,s,p])
                    }           #s

CHECK[c,p]<-pi[c,1,p]+pi[c,2,p]+pi[c,3,p]+pi[c,4,p]+pi[c,5,p]+pi[c,6,p]  #Check sums to N
        }           #c
}                   #p


#Costs in each cycle of model

for (p in 1:P){
    for(c in 2:T) {
        ct[c,p]<-inprod(pi[c, ,p],cost[c, ,p])/pow((1+disc.c),(c-1))
    }               #c
}                   #p

#Utlities in each cycle of model
for (p in 1:P){
    for(c in 2:T) {
        ut[c,p]<-inprod(pi[c, ,p],u[c, ,p])/pow((1+disc.u),(c-1))
    }   #c

    TotC[p]<-sum(ct[,p])
    mean.C[p]<-TotC[p]/N

    TotU[p]<-sum(ut[,p])
    mean.U[p]<-TotU[p]/N
}           #p

Cost.diff[2]<-mean.C[2]-mean.C[1]               #Intervention2 compared to usual care
Cost.diff[3]<-mean.C[3]-mean.C[1]               #Intervention3 compared to usual care
Cost.diff[4]<-mean.C[4]-mean.C[1]               #Intervention4 compared to usual care
Cost.diff[5]<-mean.C[5]-mean.C[1]               #Intervention5 compared to usual care
Cost.diff[6]<-mean.C[6]-mean.C[1]               #Intervention6 compared to usual care
Cost.diff[7]<-mean.C[7]-mean.C[1]               #Intervention7 compared to usual care

Util.diff[2]<-mean.U[2]-mean.U[1]               #Intervention2 compared to usual care
Util.diff[3]<-mean.U[3]-mean.U[1]               #Intervention3 compared to usual care
Util.diff[4]<-mean.U[4]-mean.U[1]               #Intervention4 compared to usual care
Util.diff[5]<-mean.U[5]-mean.U[1]               #Intervention5 compared to usual care
Util.diff[6]<-mean.U[6]-mean.U[1]               #Intervention6 compared to usual care
Util.diff[7]<-mean.U[7]-mean.U[1]               #Intervention7 compared to usual care

#Cost-effectiveness
for (b in 2:7){
        ICER[b]<-Cost.diff[b]/Util.diff[b]                  #Iincremental cost-effectiveness ratio (ICER)
}

for(k in 1:K){
    Rc[k]<-(k-1)*2000
    for (p in 1:P){
        NB[p,k]<-Rc[k]*mean.U[p]-mean.C[p]  #Net monetary benefit
        pCE[p,k]<-equals(rank(NB[,k],p),P)      #Probability CE for cost-effectiveness
                                                #Acceptability Curves

        }           #p
}                   #k

}           #END OF MODEL
```

Util.diff[2] chains 1:2 sample: 20000

Util.diff[3] chains 1:2 sample: 20000

Util.diff[4] chains 1:2 sample: 20000

Util.diff[5] chains 1:2 sample: 20000

Util.diff[6] chains 1:2 sample: 20000

Util.diff[7] chains 1:2 sample: 20000

Cost.diff[2] chains 1:2 sample: 20000

Cost.diff[3] chains 1:2 sample: 20000

Cost.diff[4] chains 1:2 sample: 20000

Cost.diff[5] chains 1:2 sample: 20000

Cost.diff[6] chains 1:2 sample: 20000

Cost.diff[7] chains 1:2 sample: 20000

232

Appendix R Numbers of households in each state at the start of cycles 1-4

		Household with safety gate	Household with no safety gate	Child with permanent fall disability with safety gate	Child with permanent fall disability with no safety gate	Child death from fall	Child death from other causes
Cycle	Intervention	S1	S2	S3	S4	S5	S6
1	1	55950	44050				
1	2	80350	19650				
1	3	81080	18920				
1	4	79510	20490				
1	5	80730	19270				
1	6	80250	19750				
1	7	87490	12510				
2	1	99600	363	0.8	0.35	0.002	37
2	2	99780	183	0.7	0.17	0.002	37
2	3	99780	178	0.7	0.17	0.002	37
2	4	99770	189	0.7	0.18	0.002	37
2	5	99780	181	0.7	0.17	0.002	37
2	6	99780	184	0.7	0.17	0.002	37
2	7	99830	131	0.7	0.12	0.001	37
3	1	99900	38	1.8	0.03	0.003	59
3	2	99900	36	1.5	0.03	0.003	59
3	3	99900	37	1.5	0.03	0.003	59
3	4	99900	37	1.6	0.03	0.003	59
3	5	99900	36	1.5	0.03	0.003	59
3	6	99900	36	1.5	0.03	0.003	59
3	7	99900	36	1.5	0.03	0.003	59
4	1	99880	38	1.8	0.03	0.003	75
4	2	99890	36	1.5	0.03	0.003	75
4	3	99890	37	1.5	0.03	0.003	75
4	4	99890	37	1.6	0.03	0.003	75
4	5	99890	36	1.5	0.03	0.003	75
4	6	99890	36	1.5	0.03	0.003	75
4	7	99890	36	1.5	0.03	0.003	75

Extending meta-analysis to answer public health policy questions using network meta-analysis, including covariates and incorporating individual participant data

Stephanie Hubbard (sjh62@le.ac.uk)
Senior Teaching Fellow in Medical Statistics and part-time PhD student
Department of Health Sciences
University of Leicester

www.le.ac.uk

Acknowledgements

- Other members of the Keeping Children Safe team
  Including Alex Sutton, Nicola Cooper, Felix Achana, Denise Kendrick(Nottingham), Pedro Saramago(York)

"This study presents independent research commissioned by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10231). The views expressed are those of the author(s) and are not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health, United Kingdom."

Background to the Keeping Children Safe project

- Unintentional injury is a major public health challenge in young children
  - falls, poisoning and thermal injuries

- *"Better Safe than Sorry"*:

  - "little evidence of a systematic approach to child injury prevention within the NHS"

- Overall aim is to increase evidence-based NHS prevention of falls, poisoning and thermal injuries in young children at home

Example

- Prevention of falls accidents in pre-school children

- Effectiveness of home safety education and the provision of stair gates

- Studies identified as part of a general Cochrane review of safety equipment (updated 2011)

Challenges

- Analysis fully uses all the information available

- Accounts for the heterogeneous evidence

Systematic review – Possession of a fitted stair gates

	Author	Year	Design	Clustered	IPD
1	Clamp	1998	rct	no	yes
2	Geilen	2002	rct	yes	no
3	Kendrick	1999	non-rct	yes	yes
4	Kendrick	2005	rct	yes	yes
5	King J	2001	rct	no	no
6	McDonald	2004	rct	no	yes
7	Nansel	2002	rct	no	yes
8	Nansel	2008	non-rct	no	yes
9	Phelan	2010	rct	no	yes
10	Posner	2004	rct	no	yes
11	Snazjder	2003	rct	no	yes
12	Watson	2004	rct	no	yes

Analyses

Binary outcome – possession of a fitted stair gate

1. Meta-analysis (random effects)
   - summary data
   - comparing a control arm to an intervention arm

2. Meta-analysis (random effects)
   - combining IPD and summary data
   - comparing a control arm to an intervention arm
   - effect of intervention by social groups

3. Network meta-analysis (mixed treatment comparison)
   - account for different interventions

1. Meta-analysis using summary data only

- Results adjusted for clustering
- *Overall OR=1.60 (95%CI: 1.18 to 2.18) indicating families in the home safety intervention arms more likely to have a fitted stair gate than control group families*
- *Significant heterogeneity between studies (p<0.001)*
- *No evidence of publication bias*
- *Sensitivity analysis showed findings were robust to study exclusion*

2. Meta-analysis using summary and IPD data

- Used to investigate the effect of social variables all previously shown to be associated with risk of injury:

  child age, male gender, black or minority ethnicity, single parent family, at least one parent unemployed, residing in rented accommodation

- IPD and study summary level data analysed simultaneously to :
  - minimise bias
  - maximise power   (Lambert 2002)

2. Meta-analysis using summary and IPD data
We used the methods developed by Sutton et al (2008)

| Non-cluster IPD |
| Cluster Allocated IPD |
| Non-cluster Summary Data |
| Cluster Allocated Summary Data |

→ Estimates of shared intervention effect are combined in a random effects meta-analysis

2. Meta-analysis using summary and IPD data

- IPD analysis used individual covariate values
- Summary analysis used proportion of study population with covariate of interest
- Ecological bias: study-level and participant-level analyses give different results
  - split the variability between studies and within studies (Riley et al 2010)

2. Meta-analysis using summary and IPD data

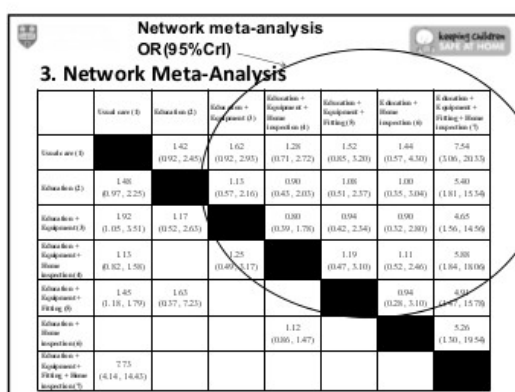*OR for accommodation type with 95% Credible Interval*

- *Difference between and within -0.31(-1.40 to 0.78)*
- *Combining between and within study variances*

Rented	1.98 (1.48 to 2.66)
Owner occupied	1.22 (0.96 to 1.61)
Interaction term	1.62 (1.18 to 2.24)

There was evidence that the intervention was more effective amongst families living in rented accommodation.

3. Network Meta-Analysis

- Previous analyses every study taken to consist of an intervention group and a control group but this was not the case
- The interventions compared in the studies were:
  1. *Usual care*
  2. *Education*
  3. *Education + Equipment*
  4. *Education + Equipment + Home Inspection*
  5. *Education + Equipment + Fitting*
  6. *Education + Home Inspection*
  7. *Education + Equipment + Fitting + Home Inspection*

3. Network Meta-Analysis

Network Diagram of Interventions for Stairgates

Each intervention is a node in the network. The links between the nodes are studies or pairs of study arms. The numbers along the link lines indicate the number of studies or pairs of study arms for that link in the network.

---

3. Network Meta-Analysis

	Usual care (1)	Education (2)	Education + Equipment (3)	Education + Equipment + Home inspection (4)	Education + Equipment + Fitting (5)	Education + Home inspection (6)	Education + Equipment + Fitting + Home inspection (7)
Usual care (1)							
Education (2)	1.48 (0.97, 2.25)						
Education + Equipment (3)	1.92 (1.05, 3.51)	1.17 (0.52, 2.63)					
Education + Equipment + Home inspection (4)	1.13 (0.82, 1.58)		1.25 (0.49, 3.17)				
Education + Equipment + Fitting (5)	1.45 (1.18, 1.79)	1.63 (0.57, 7.23)					
Education + Home inspection (6)				1.12 (0.86, 1.47)			
Education + Equipment + Fitting + Home inspection (7)	7.73 (4.14, 14.45)						

Pair-wise meta-analysis or single study results OR(95% CI)

Network meta-analysis OR (95%CrI)

3. Network Meta-Analysis

	Usual care (1)	Education (2)	Education + Equipment (3)	Education + Equipment + Home inspection (4)	Education + Equipment + Fitting (5)	Education + Home inspection (6)	Education + Equipment + Fitting + Home inspection (7)
Usual care (1)		1.42 (0.92, 2.45)	1.62 (0.92, 2.93)	1.28 (0.71, 2.72)	1.52 (0.85, 3.20)	1.44 (0.57, 4.30)	7.54 (3.06, 20.33)
Education (2)	1.48 (0.97, 2.25)		1.15 (0.57, 2.16)	0.90 (0.43, 2.03)	1.08 (0.51, 2.57)	1.00 (0.35, 3.04)	5.40 (1.81, 15.54)
Education + Equipment (3)	1.92 (1.05, 3.51)	1.17 (0.52, 2.63)		0.80 (0.39, 1.78)	0.94 (0.42, 2.54)	0.90 (0.32, 2.80)	4.65 (1.56, 14.56)
Education + Equipment + Home inspection (4)	1.13 (0.82, 1.58)		1.25 (0.49, 3.17)		1.19 (0.47, 3.10)	1.11 (0.52, 2.46)	5.08 (1.84, 18.06)
Education + Equipment + Fitting (5)	1.45 (1.18, 1.79)	1.63 (0.57, 7.23)				0.94 (0.28, 3.10)	4.91 (1.47, 15.78)
Education + Home inspection (6)				1.12 (0.86, 1.47)			5.26 (1.30, 19.54)
Education + Equipment + Fitting + Home inspection (7)	7.73 (4.14, 14.45)						

---

3. Network Meta-Analysis

	Possession of a stair gate	
Intervention	Probability intervention is best	Median intervention rank (95% CrI)
Usual care (1)	0.00	7(5, 7)
Education (2)	0.002	4(2, 7)
Education + Equipment (3)	0.004	3(2, 7)
Education + Equipment + Home inspection (4)	0.001	5(2, 7)
Education + Equipment + Fitting (5)	0.008	4(2, 7)
Education + Home inspection (6)	0.013	4(2, 7)
Education + Equipment + Fitting + Home inspection (7)	0.97	1(1, 2)

---

3. Network Meta-Analysis

- Most effective intervention is most intensive (*Education + Equipment + Fitting + Home Inspection*)

- No evidence of inconsistency

- Still some lumping together of interventions; education differed, different levels of usual care

---

Extensions

- Network meta-analysis using IPD and summary data incorporating covariates.
  - Saramago et al (accepted for publication Stats in Med)

- Use the results in a cost-effectiveness analysis to inform the NHS on whether to implement a home safety programme

Summary

- Lots of challenges in synthesising the evidence from public health interventions

- Use of IPD increases the power of a meta-analysis and minimises bias when investigating covariate effects

- Network meta-analyses can account for some of the heterogeneity of the evidence

- Analyses should use as much of the information available as possible

References

- Kendrick D, Young B, Mason-Jones AJ, Ilyas N, Achana FA, Cooper NJ, Hubbard SJ, Sutton AJ, Smith S, Wynn P, Mulvaney C, Watson MC, Coupland C. Home safety education and provision of safety equipment for injury prevention. Updated Cochrane review not yet published.
- Sutton AJ, Kendrick D, Coupland CA. Meta-analysis of individual- and aggregate-level data. Stat Med 2008 Feb 28;27(5):651-669. Cochrane
- Kendrick D, Watson MC, Mulvaney CA, Smith SJ, Sutton AJ, Coupland CA, et al. Preventing childhood falls at home: meta-analysis and meta-regression. Am J Prev Med 2008 Oct;35(4):370-379.
- Lambert PC, Sutton A, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002; 55(1): 86-94.
- Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. Res Syn Meth 2010; 1: 2-19.
- Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated August 2011. Available from http://www.nicedsu.org.uk
- Cooper NJ, Kendrick D, Achana F, Dhiman P, He Z, Wynn P, Le Cozannet E, Saramago P, Sutton A. Network Meta-Analysis to evaluate the effectiveness of different interventions to increase the prevalence of functioning smoke alarms in households with children. Epidemiol. Rev. 2012 34: 32-45.
- Saramago P, Sutton A, Cooper N, Manca A. Mixed Treatment Comparisons using Aggregate- and Individual-Participant Level Data. Accepted for publication April 2012 in Stats in Med.

Appendix U Papers

ACHANA, F., HUBBARD, S., SUTTON, A., KENDRICK, D. and COOPER, N., 2014. An exploration of synthesis methods in public health evaluations of interventions concludes that the use of modern statistical methods would be beneficial. *Journal of clinical epidemiology,* **67**(4), pp. 376-390.

Contribution: I analysed the data and wrote section 3, Exposition of new synthesis

methodology applied in a PH evaluation context, and contributed to the discussion.

Hubbard, S., Cooper, N., Kendrick, D., Young, B., Wynn, P. M., He, Z., . . . Sutton, A. (2015). Network meta-analysis to evaluate the effectiveness of interventions to prevent falls in children under age 5 years. *Injury Prevention*, *21*(2), 98-108. doi:10.1136/injuryprev-2013-041135

Contribution: I analysed the data and wrote the paper except for the background and

systematic review information.

An exploration of synthesis methods in public health evaluations of interventions concludes that the use of modern statistical methods would be beneficial

Felix Achana[a], Stephanie Hubbard[a], Alex Sutton[a], Denise Kendrick[b], Nicola Cooper[a],[*]

[a]*Department of Health Sciences, University of Leicester, University Road, Leicester LE1 7RH, UK*
[b]*Division of Primary Care, Community Health Sciences, Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

Abstract

Objectives: To review the methods currently used to synthesize evidence in public health evaluations and demonstrate the availability of more sophisticated approaches.

Study Design and Setting: A systematic review of National Institute for Health and Care Excellence (NICE) public health appraisals published between 2006 and 2012 was performed to assess the methods used for the synthesis of effectiveness evidence. The ability of new developments in evidence synthesis methodology to address the challenges and opportunities present in a public health context is demonstrated.

Results: Nine (23%) of the 39 NICE appraisals included in the review performed pairwise meta-analyses as part of the effectiveness review with one of these also including a network meta-analysis. Of the remainder, 29 (74.4%) presented narrative summaries of the evidence only, and 1 (2.6%) appraisal did not present any review of effectiveness and/or cost-effectiveness evidence. Heterogeneity of outcomes, methods, and interventions were the main reasons given for not pooling the data. Exploration of quantitative synthesis methods shows that pairwise meta-analyses can be extended to incorporate individual participant data (when it is available), extend the number of interventions being compared using a network meta-analysis, and adjust for both subject- and summary-level covariates. All these can contribute to ensuring the analysis answers directly the policy-relevant questions.

Conclusion: More sophisticated methods in evidence synthesis should be considered to make evaluations in public health more useful for decision makers. © 2014 Elsevier Inc. All rights reserved.

Keywords: Public health evaluation; Network meta-analysis; Decision making; Meta-analysis; Systematic review

1. Introduction

Systematic reviews and economic evaluations conducted within a decision modeling framework are two important tools in health-care evaluation [1,2]. Systematic reviews with or without meta-analyses have been accepted as providing a transparent and consistent way of obtaining research evidence on effectiveness of interventions in a way that minimizes bias [3]. Decision analytical models offer an additional framework through which effectiveness evidence, ideally from a systematic review, may be integrated with other relevant evidence and information on resource utilization to derive comparative estimates of cost-effectiveness. By providing a framework for assessing effectiveness and cost-effectiveness, these methods enable policy-relevant questions such as which interventions represent the best use of scarce health-care resources to be answered [4].

A key component of a systematic review is how the evidence, on outcomes such as effectiveness and adverse events, is synthesized. Meta-analysis, when used in a systematic review to combine quantitative information from multiple well-conducted randomized controlled trials (RCTs), is considered at the top of the hierarchy of evidence for intervention effectiveness [5]. An alternative approach to evidence synthesis, when meta-analysis is considered

What is new?

Key findings
- Quantitative synthesis is not carried out in the systematic reviews for most public health (PH) evaluations.

- When quantitative synthesis *is* done, it tends to use the simplest methods, for example, a fixed- or random-effects meta-analyses comparing two groups, which potentially limits the scope of the analysis.

What this adds to what was known?
- Demonstrates how more sophisticated synthesis methods can be used in PH appraisals to more realistically model the data and answer the relevant policy questions.

What is the implication and what should change now?
- Researchers working on PH evaluations should consider expanding their toolbox and using more sophisticated methods many of which have recently been developed, motivated, and applied in pharmaceutical evaluations.

inappropriate, is narrative synthesis (also referred to as qualitative synthesis [6]). In this approach, individual studies identified in the review are summarized using a variety of formats without combining results quantitatively [7].

Meta-analysis is widely applied in reviews of the effectiveness of clinical interventions, treatments, and medical device technologies where the interventions and health outcomes are usually well defined and evaluated in well-conducted RCTs [8]. In other fields of health-care evaluation, however, things may not always be as clear cut. A good example is public health (PH), where interventions are often more complex and less well defined than clinical interventions [9]. There may also be a lack of good-quality evidence, particularly from RCTs in PH, for a number of well-documented reasons [10,11] including limited generalizability of the findings of RCTs to the wider population due to highly selected study populations, a narrow definition of intervention strategies and outcomes, and a focus on the individual instead of the community that is of interest in PH. Even when feasible, many have argued that RCTs may not always be possible to conduct in PH for other reasons, for example, ethical concerns may be raised regarding not offering the control population a possibly beneficial intervention [10]. Also, many of the RCTs conducted in PH tend to be cluster randomized trials and hence

have more complex designs that need adjusting for in the analysis. In addition, the best available PH evidence may often come from observational nonrandomized studies [9], despite the increased risk of bias associated with the lack of randomization. For these reasons, the use of quantitative evidence synthesis methods such as meta-analysis in PH raises a number of methodological challenges. These include (1) increased methodological heterogeneity and risk of bias as a result of including studies with different study designs (RCTs, cluster RCTs, controlled before-and-after studies, and other observational nonrandomized studies), (2) the interventions or "program" being evaluated is often described in little detail, (3) a wide range of outcomes measures are often used, which may be variously defined across studies, and (4) the use of intermediate and/or surrogate outcome measures.

There are growing calls for PH decision making to be based on the best available evidence whenever possible. For example, a 2004 Department of Health report [12] on improving health and reducing health inequalities in England called for economic evaluations of PH interventions to ensure judicious use of scarce resources. Following this report, the remit of the UK National Institute for Health and Care Excellence (NICE), which already evaluated pharmaceutical interventions, was expanded to include the development of guidance for PH based on sound appraisals of intervention effectiveness and cost-effectiveness [13]. Consequently, a number of PH appraisals have been produced by NICE since 2006 on a wide range of issues including smoking cessation, alcohol use, and, particularly of relevance to the example used in this article, unintentional injuries in children.

To help address specific methodological challenges and provide advice on the technical aspects of the appraisal development process, NICE published a manual of methods for PH evaluation in 2006 [14], which was subsequently updated in 2009 [15] (a further update was published in September 2012 [16] after this review was completed, but the guidance was not changed). The guidance recommended "Meta-analysis data may be used to produce a graph if the data (usually from RCTs) are sufficiently homogenous and if there are enough relevant and valid data from comparable (or the same) outcome measures. Where such data are not available, the synthesis may have to be restricted to a narrative overview of individual studies looking at the same question," "Before pooling or combining the results of different studies, the degree of heterogeneity in the data should be assessed to determine how the results have been affected by the circumstances in which studies were carried out," and "Publication bias [17,18] should be critically assessed and reported in the interpretation of the meta-analysis results." These recommendations match well to the challenges in systematic review/meta-analysis in PH highlighted by the Cochrane Collaboration [9] and the 2011 Institute of Medicine report on standards for systematic reviews [6].

In view of the aforementioned challenges facing PH evaluations and recommendations for synthesis of PH evidence contained in the NICE manuals of methods, a review of all NICE PH appraisals published since 2006 was conducted. The aim of this article is twofold: (1) to identify the current situation (ie, what is already done and/or not done) with regards to addressing problems in synthesis of PH evidence and (2) to illustrate the application of new synthesis methods (ie, beyond those recommended by NICE [14–16] and Cochrane [9]) including methods from other fields such as health technology assessment to PH evidence that we believe have the potential to address many of the challenges in PH evaluation as aforementioned and thus improve the quality of evidence syntheses in PH interventions.

2. Systematic review of NICE PH appraisals

2.1. Methods

Completed PH appraisals published between March 1, 2006 and September 25, 2012 were identified for inclusion in the review through the NICE Web site (http://www.nice.org.uk/Guidance/PHG/Published). Each PH appraisal consisted of a number of articles such as qualitative reviews, epidemiologic reviews, expert opinions, field reports, and other similar nonquantitative review reports, quantitative systematic reviews of intervention effectiveness and cost-effectiveness, and decision analytical modeling reports. These were retrieved from the "background information" sections and assessed for eligibility. The "how this guidance was produced" sections were also searched for relevant articles if none were identified under "background information." Articles meeting the inclusion criteria were systematic reviews of the quantitative effectiveness and cost-effectiveness evidence and/or decision analytical modeling reports. Qualitative evidence reviews, epidemiologic reviews, field reports, expert opinions, and other similar nonquantitative evidence review reports were excluded. In addition, the final appraisal/guidance documents developed for each PH appraisal area were also excluded as these did not contain relevant information on the conduct of the evidence synthesis and decision modeling, which is of interest in this review. All except two (PH1 and PH2) of the appraisals were published after the 2006 NICE manual of methods [14] so should have followed the guidance for quantitative effectiveness evidence synthesis techniques.

Information extracted from the retrieved articles was used to assess the methods used to synthesize the effectiveness evidence and subsequent incorporation of the evidence into the decision models (when developed) that informed the PH appraisal. The assessment criteria for the synthesis methods were

1. Type of systematic review—narrative summary vs. meta-analysis;

2. Included studies—RCT vs. observational (non-randomized) studies;
3. Methods used to synthesize the evidence (if undertaken), including specification of the statistical model (including fixed- and/or random-effects models), heterogeneity, publication bias, and the outcome measures used, as well as presentation of results; and
4. How evidence from the systematic review was used to inform any cost-effectiveness analysis.

2.2. Results of systematic review

Thirty-nine completed PH appraisals published since 2006 were identified from the NICE PH Web site. Within these 39 appraisals, 371 potentially relevant articles were retrieved, and after screening the titles and reading the introduction and/or abstract sections, 164 were excluded as they failed to meet the inclusion criteria. Fifty-two articles, identified as duplicates and supplementary appendices, were combined with the corresponding main report and counted as one article leaving a total of 155 articles for inclusion in this review. The median number of included articles per appraisal was 4 (range 0 to 10). [No relevant supporting document meeting our inclusion criteria existed for one appraisal (PH36—prevention and control of hospital infection).]

2.2.1. Type of review

Table 1 lists all 39 PH appraisals by summary of the evidence synthesis and cost-effectiveness analyses undertaken to inform each appraisal development. One appraisal (PH36) reported neither effectiveness and cost-effectiveness evidence reviews nor a decision model, two appraisals (PH33 and PH34) reported reviews of evidence but conducted no cost-effectiveness analysis, and the fourth appraisal (PH7) reported evidence reviews and decision models; however, no estimates of cost-effectiveness of interventions were presented.

Twenty-nine (74.4%) of the 39 appraisals contained systematic reviews in which only a narrative summary of the evidence was conducted, another seven (18%) conducted both narrative summary and meta-analysis, two appraisals (5%) conducted only meta-analysis, and one (2.6%) appraisal had no systematic review and hence no evidence synthesis. In the narrative summary approach, the review findings were summarized study by study in the text and through tables. Sometimes, forest plots were used to display results of primary studies, but no overall mean or pooled result was presented (see PH4 for an example). Eight of the 29 appraisals using only a narrative summary approach did not report the reasons for not pooling the data, 2 included only review-level evidence from the overview of reviews, and 19 cited heterogeneity as the reason why meta-analysis was not considered appropriate. The reported causes of heterogeneity are presented in Appendix at www.jclinepi.com.

Table 1. NICE public health appraisals and summary of evidence synthesis methods and decision modeling used to inform their development

NICE public health appraisal title	Review of the effectiveness and cost-effectiveness evidence and decision analysis used to inform each appraisal					
	Systematic review of effectiveness (narrative summary)	Systematic review of effectiveness (at least one M-A)[a]	Cost-effectiveness reviews	Study quality	Decision model	Source of effectiveness estimate used in decision model[b]
Brief interventions and referral for smoking cessation (PH1)	✓	x[nr]	✓	✓	✓	Published review
Four commonly used methods to increase physical activity (PH2)	✓	x[nr]	✓	✓	✓	Individual study[1]
Prevention of sexually transmitted infections and under 18 conceptions (PH3)	✓	✓	✓	✓	✓	Published review
Interventions to reduce substance misuse among vulnerable young people (PH4)	✓	x[o]	✓	✓	✓	Individual study[1]
Workplace interventions to promote smoking cessation (PH5)	✓	x[m]	✓	✓	✓	Individual study[3]
Behaviour change (PH6)	✓	x[nr]	✓	✓	✓	Individual study[4]
School-based interventions on alcohol (PH7)	✓	✓	✓	✓	✓	Individual study[1]
Physical activity and the environment (PH8)	✓	x[l,m,o]	✓	✓	✓	Individual study[3]
Community engagement (PH9)	✓	x[l,m,o,p]	✓	✓	x	Not applicable
Smoking cessation services (PH10)	✓	x[m]	✓	✓	✓	New meta-analysis
Maternal and child nutrition (PH11)	✓	x[m]	✓	✓	✓	Individual study[5]
Social and emotional wellbeing in primary education (PH12)	✓	✓	✓	✓	✓	Individual study[5]
Promoting physical activity in the workplace (PH13)	✓	x[nr]	✓	✓	✓	Individual study[5]
Preventing the uptake of smoking by children and young people (PH14)	✓	x[m,o]	✓	✓	✓	Individual study[1]
Identifying and supporting people most at risk of dying prematurely (PH15)	✓	x[l,m,p]	✓	✓	✓	Individual study[1]
Mental wellbeing and older people (PH16)	✓✓	x[l,m,o]	✓✓	✓✓	✓✓	Individual study[1]
Promoting physical activity for children and young people (PH17)	✓✓	x[l,m,o]	✓✓	✓✓	✓✓	Analyst estimate[4]
Needle and syringe programmes (PH18)	✓✓	x[nr]	✓✓	✓✓	✓✓	Individual study[3]
Management of long-term sickness and incapacity for work (PH19)	✓✓	✓	✓✓	✓✓	✓✓	New meta-analysis
Social and emotional wellbeing in secondary education (PH20)	✓	x[l,m,o]	✓	✓	✓	Individual study[2]
Reducing differences in the uptake of immunisations (PH21)	✓	x[nr]	✓	✓	✓	Individual study[4]
Promoting mental wellbeing at work (PH22)	✓✓	x[l]	✓✓	✓✓	✓✓	Individual study[3]
School-based interventions to prevent smoking (PH23)	✓✓	✓	✓✓	✓✓	✓✓	New meta-analysis
Alcohol-use disorders - preventing harmful drinking (PH24)	✓	x[nr]	✓	✓	✓	Published review
Prevention of cardiovascular disease (PH25)	✓✓	x[nr]	✓✓	✓✓	✓✓	Individual study[5]
Quitting smoking in pregnancy and following childbirth (PH26)	✓✓	x[l,m,o]	✓✓	✓✓	✓✓	Published review

(Continued)

243

Table 1. Continued

NICE public health appraisal title	Review of the effectiveness and cost-effectiveness evidence and decision analysis used to inform each appraisal					
	Systematic review of effectiveness (narrative summary)	Systematic review of effectiveness (at least one M-A)[a]	Cost-effectiveness reviews	Study quality	Decision model	Source of effectiveness estimate used in decision model[b]
Weight management before, during and after pregnancy (PH27)	✓	✓	✓	✓	✓	Not clear[5] reported
Looked-after children and young people (PH28)	✓✓	x[m,o,p]	✓✓	✓✓	✓✓	Individual study[3]
Strategies to prevent unintentional injuries among under-15s (PH29)	✓✓	x[j]	✓✓	✓✓	✓✓	Individual study[3]
Preventing unintentional injuries among under-15s in the home (PH30)	✓	x[j,o]	✓	✓	✓	Individual study[3]
Preventing unintentional road injuries among under-15s: road design (PH31)	✓	x[j]	✓	✓	✓	Individual study[2]
Skin cancer prevention: information, resources and environmental changes (PH32)	✓	x[j,m]	✓	✓	✓	Individual study[3]
Increasing the uptake of HIV testing among black Africans in England (PH33)	✓	x[m]	✓	x	x	Not applicable
Increasing the uptake of HIV testing among men who have sex with men (PH34)	✓	x[nr]	✓	x	x	Not applicable
Preventing type 2 diabetes—population and community interventions (PH35)	✓	✓	✓	✓	✓	New meta-analysis
Prevention and control of healthcare-associated infections (PH36)		x	x	x	x	Not applicable
Tuberculosis - hard-to-reach groups (PH37)	✓✓	x[nr]	✓✓	✓✓	✓✓	Individual study[5]
Preventing type 2 diabetes - risk identification and interventions for individuals at high risk (PH38)		✓	✓✓	✓✓	✓✓	New meta-analysis
Smokeless tobacco cessation - South Asian communities (PH39)	✓	x[s]	✓	✓	✓	Published review

Abbreviations: NICE, National Institute for Health and Care Excellence; M-A, meta-analysis.

Ticks indicate a systematic review of evidence, meta-analysis, or decision model have been conducted, whereas x indicates analysis have not been conducted.

[a] Reported reason why meta-analysis was not done [i = heterogeneity of interventions, m = heterogeneity of methods, design, and settings, o = heterogeneity of outcome measures, p = heterogeneity of study populations, s = heterogeneity of studies (specific cause not reported), and nr = not reported or reported that studies do not support a meta-analysis].

[b] Selection of individual study estimate of intervention effect used in the decision model (1 = used a prespecified criteria reported in the decision model report, 2 = discussion with NICE or estimates selected based on quality grading of evidence using the guide manuals of methods, 3 = selected studies based on the relevance of the intervention to the decision problem, 4 = assumption/analyst estimated based on an assumption, and 5 = not clearly reported).

2.2.2. Included studies—RCTs vs. nonrandomized studies

Two (PH23 and PH38) of the 38 appraisals (containing a systematic review) included evidence from RCTs only in the effectiveness review. The remaining 36 appraisals were informed by reviews of both randomized and observational (nonrandomized) evidence identified from individual study reports and/or published systematic review reports. All 38 appraisals (containing a systematic review) graded the quality of primary studies and assessed the applicability of the evidence adhering to the guidelines for PH appraisal methods [14,15].

2.2.3. Quantitative evidence synthesis

Only 9 of the 39 appraisals (23%) contained one or more systematic review with a meta-analysis (Table 2). In total, there were 10 systematic reviews and/or decision analytical modeling reports with at least one meta-analysis within the nine appraisal areas. (Note: PH10 has two systematic review reports in which a meta-analysis was conducted.) Four of the 10 meta-analyses included RCTs only, and six included both RCT and observational (non-RCT) studies. Six of the 10 meta-analyses were conducted on "final outcomes"; that is, the main outcome measures on which the corresponding cost-effectiveness analyses were based (eg, PH10 Smoking abstinence). The remaining four meta-analyses were conducted on "intermediate outcomes" (eg, PH3 Uptake of Chlamydia screening in schools rather than prevention of chlamydia).

There was evidence that interventions may have been "lumped" [19,20] into two broad intervention groups to facilitate inclusion of more studies in 7 of the 10 reports with a meta-analysis. For example, in PH23, which investigated the effect of school-based interventions on alcohol consumption, seemingly different interventions (such as lessons delivered by teachers or other professionals as part of the curriculum; peer-led education by other pupils; external contributions from, for example, the police, life education center staff; and implementation of school policy—type interventions) were lumped together to form one "intervention group," which was then compared with the no intervention control in a pairwise meta-analysis.

Seven of the 10 review reports conducted random-effects pairwise meta-analysis, one conducted fixed- and random-effects analyses, one conducted random-effects mixed treatment comparisons [20] (also referred to as network meta-analysis [21,22]—see later) alongside the pairwise analysis, and another one did not clearly present the statistical model used. Six of the 10 systematic reviews presented forest plots with heterogeneity statistics displayed on them, two (PH3 and PH1) presented forest plots without heterogeneity statistics, and one review (PH35) did not present a forest plot. Only one review (PH23) assessed publication bias using funnel plot and Egger's test for asymmetry.

2.2.4. How the evidence from the systematic reviews was incorporated into the model

Thirty-five (89.7%) of the 39 appraisals were informed by cost-effectiveness evaluations contained in one or more decision analytical modeling reports (Table 1). Twenty-three (66%) of these used estimates of intervention effectiveness derived from individual studies identified in the systematic review to inform the decision analysis (reasons for using the studies selected given in Table 1), 5 (14%) used previously published systematic review results, another 5 (14%) used estimates from a meta-analysis of studies identified in the systematic review, 1 used expert opinion/analyst estimate, and another one did not clearly report the source(s) of the intervention effect.

3. Exposition of new synthesis methodology applied in a PH evaluation context

In this section, we outline new developments in evidence synthesis methodology, many of them motivated by the evaluation of medical interventions and others motivated specifically by challenges in PH. We also show how such methods can be applied in a PH context to help address challenges and opportunities that exist in this context and thus, in some situations, raise the quality bar (established in the first part of this article) for PH interventions.

We use, for illustration, a topic area in which the authors have actively been working for several years—accident prevention among preschool children at home. This area of accident prevention among children at home was recently appraised by NICE PH30 (Table 1) using only narrative summaries for the systematic review of intervention effectiveness and thus using estimates from individual trials to inform the cost-effectiveness analyses. We have found accident prevention to have many of the issues typical of PH appraisals including studies of different designs, heterogeneity in both study design [eg, specific nature of interventions, level of randomization (individual or cluster), etc.] and study results, and interest in differential treatment effects across degrees of population inequality such as accommodation type, proportion of black and minority ethnicity, and proportion of single-parent families.

The account discussed later follows an approximately chronological path and details the development and adaptation of methods to synthesize the evidence by making the best use of available data. In this study, we restrict our attention to strategies to reduce falls among children at home, in particular, to increase the possession of a fitted stair gate(s) in homes.

We start by discussing the analyses performed in a recently updated Cochrane review [23] of interventions to prevent unintentional injuries to children at home—pairwise meta-analysis, subgroup analyses to explore heterogeneity, and meta-regression incorporating individual participant data (IPD). We then present a network meta-

Table 2. Review of quantitative methods used to synthesize public health evidence for NICE public health appraisal

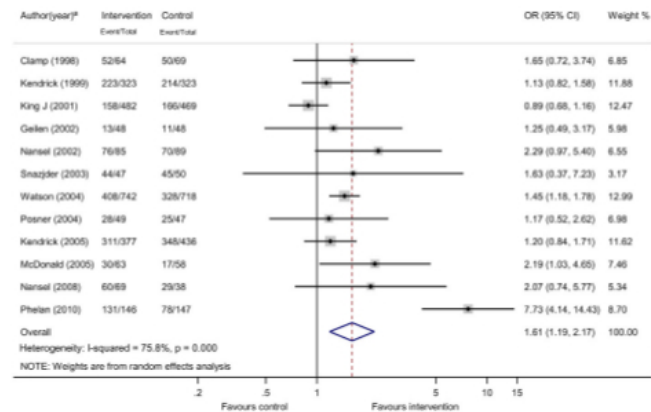| Appraisal title	Systematic review report title	Included RCTs only	Main outcome	Description of main outcome	Outcome measure: statistic	Type of synthesis	Model type	Lumping[a] of interventions	Presentation of results	Assessed publication bias	Software	Used result of M-A in decision model
Prevention of sexually transmitted infections and under 18 conceptions (PH3)	Review 2 - Review of evidence for the effectiveness of screening for genital chlamydial infection in sexually active young women and men	No	Intermediate	Uptake of proactive chlamydia screening using home-collected specimens	Screening response rate (%)	M-A	Random effects	No	FP/Txt	No	RevMan, Stata	No
School-based interventions on alcohol (PH7)	Alcohol and schools: effectiveness and cost-effectiveness review	No	Final	Alcohol use	Weighted mean difference	M-A	Random effects	Yes	FP/Txt	No	Not stated	No
Smoking cessation services (PH10)	Cut down to quit' with nicotine replacement therapies	Yes	Final	6 or more months' sustained abstinence	Relative risk and Cohen's d	M-A	Random effects	Yes	FP/T/Txt	No	RevMan	Yes
Smoking cessation services (PH10)	Final report	No	Final	6 or more months' sustained abstinence	Cohen's d	M-A	Fixed and random effects	Yes	FP/T/Txt	No	RevMan	No
Social and emotional wellbeing in primary education (PH12)	Teesside review	Yes	Intermediate	Social problem solving	Standardized mean difference	M-A	Random effects	Yes	FP/T	No	RevMan	No
Management of long-term sickness and incapacity for work (PH19)	PH19 Management of long-term sickness and incapacity for work: Economic analysis report	No	Yes	Number returning to work following sickness	Relative risk	M-A	Random	Yes	FP/T/Txt	No	RevMan	Yes
School-based interventions to prevent smoking (PH23)	School-based interventions to prevent smoking: quantitative effectiveness review	Yes	Final	Smoking uptake	Odds ratio	M-A	Random effects	Yes	FP/Txt	Yes	Stata	Yes
Weight management before, during and after pregnancy (PH27)	Weight management before, during and after pregnancy: evidence review	No	Intermediate	Number exceeding IoM[b] guidelines for healthy weight gain	Relative risk	M-A	Random effects	Yes	FP/T/Txt	No	RevMan	No

Fig. 1. Forest plot illustrating the findings of the random-effects meta-analysis of interventions aimed at increasing the uptake of safety equipment for the outcome ''possession of a fitted stair gate.'' [a] References for all the studies can be found in the updated Cochrane review [23]. OR, odds ratio; CI, confidence interval.

analysis that allows the interventions to be ranked and provides more informative evidence for a cost-effectiveness analysis.

3.1. Pairwise meta-analysis

A random-effects meta-analysis was used to synthesize the evidence for the possession of fitted stair gate(s) outcome, which comprised 12 studies [10 RCTs (2 clusters allocated) and 2 non-RCTs (1 cluster allocated)]. Because the original reporting of the cluster randomized studies had ignored the effect of clustering in their analysis, the meta-analysis was adjusted using external data to estimate the likely effects of such clustering on the certainty of the results [24]. Fig. 1 displays a forest plot of the results. Intervention arms were more likely to possess fitted stair gate(s) than the control arms [odds ratio (OR), 1.61; 95% confidence interval (CI): 1.19, 2.17]. Considerable heterogeneity was observed between study results ($I^2 = 76\%$) [25].

3.2. Subgroup analyses

Potential sources of heterogeneity were explored using subgroup analyses based on a priori explanations, which were (1) whether the intervention included the provision of safety equipment, (2) follow-up period (up to and including 3 months and 4 or more months), (3) whether the intervention was delivered in a clinical setting or at home or community, (4) use of a randomized or nonrandomized design, and (5) study quality (allocation concealment, blinding of outcome assessment, and at least 80% follow-up in each treatment arm). Some of the heterogeneities were partly explained by different settings and the provision

of stair gates, but significant heterogeneity remained in the different subgroups.

3.3. Meta-regression using IPD and summary data

In an attempt to explain further variability between study results—to address whether differential intervention effects could be discerned to be related to indicators of deprivation—and thus try and answer questions relating to inequalities in health care, a number of subject-level covariates were explored. To achieve this, the IPD were requested from the researchers responsible for all the relevant primary studies. By obtaining IPD, the power of meta-regression to explore subject-level covariates (eg, if the subject lived in owned or rented accommodation, etc.) is much increased over the use of summary data (eg, the percentage of subjects living in an owned house in a particular study) [26]; in fact, obtaining IPD is considered the gold standard way to carry out meta-analysis generally [27].

IPD were successfully obtained for approximately half of the studies across all types of injury prevention included in the review, with varying degree of success for the different injury prevention domains. But this partial success presented an analysis challenge. We wanted to not only use the IPD but also include the other studies in the analysis for which only summary data were available. This involved using a model developed for the original Cochrane review in this area [28], which essentially "married" summary and IPD meta-analysis models including covariates within a single analysis based on all available data [29]. This approach also accounted for the correct analysis of the cluster-allocated studies through appropriate reanalysis of the IPD (when available) and through utilization of adjustment

methods for the summary data as aforementioned. Importantly, using IPD allowed the use of data on outcomes that had not been reported in the articles; for example, some studies had reported composite measures of home safety and not individual safety practices, but the IPD included data on these individual safety practices.

For the possession of fitted stair gate(s) outcome, IPD were obtained for 10 of the 12 studies. Treatment interactions were investigated for child age, ethnic group, gender, family type (single or two parents), housing tenure (rented or owned), and parental unemployment. Most of the findings indicated little difference between the subgroups, except for the analysis of housing tenure, which combined the analyses of IPD for two cluster and five noncluster studies, and one study for which only summary data were available. The OR for intervention effect in non−owner-occupied households was 1.98 [95% credible interval (CrI, which is similar to a CI generated using Bayesian statistics): 1.48, 2.66], and in owner-occupied households, the OR was 1.22 (95% CrI: 0.96, 1.61), providing evidence to suggest that the intervention effect was larger in non−owner-occupied households (ratio of ORs, 1.62; 95% CrI: 1.18, 2.24).

It is interesting to note that such covariates could have been investigated without obtaining IPD through the use of meta-regression on summary-level covariates (ie, percentage of study participants in non−owner-occupied households), but such an analysis has much diminished power and is more prone to ecological/aggregation biases [30]. Running such an analysis on the same eight studies, but not using any IPD, produces an exponentiated regression coefficient of 1.01 (95% CI: 0.998, 1.022), indicating that there is no evidence of an increase in the odds of possession of fitted stair gate(s) for a one−percentage point increase in percentage of families living in non−owner-occupied household. This result is very different from the findings from the IPD analysis, which suggest that the odds of possessing fitted stair gate(s) are 62% higher among those in non−owner-occupied household than those in owner-occupied household.

3.4. Network meta-analysis

Our next refinement to the analysis, not included in the Cochrane review, came from concerns with the interpretability of the effect sizes from pairwise analyses of the type presented previously. We were aware that the interventions to increase the uptake of safety practices varied between studies (eg, interventions ranged from educational initiatives, through vouchers to reduce the price of equipment, through to the free provision and fitting of equipment), and therefore, by fitting the data into a meta-analysis framework of "intervention" vs. "usual care," the interpretation of the resulting pooled effect was unclear—exactly what does the pooled effect relate to? This was especially important as the effectiveness results were to be used to inform

the cost-effectiveness of injury prevention interventions evaluated via a decision model, which would require explicit interventions to be defined and costed. Thus, an analysis in which the different interventions were kept as unique was required. Once this was established, it became possible to include further relevant literature, known about but not used in the initial meta-analysis, in the analysis, namely, studies which compared different interventions to increase safety equipment uptake directly (but which had no "usual care" control group—hence their omission thus far). Further literature searches were conducted to identify all such studies. Network meta-analysis, which was being increasingly used in the evaluation of pharmaceuticals for funding bodies such as NICE [31], presented an analysis approach that would both keep interventions distinct and include trials with direct comparisons.

The meta-analysis of possession of fitted stair gate(s) outcome presented in the Cochrane review included all studies that compared a control group with an enhanced intervention group, but these controls and interventions varied considerably as outlined in Fig. 2A. In fact, seven distinct controls and interventions (including usual care) were identified across the included studies. To better understand the structure of the evidence base, when interventions are defined in this more refined way, a network diagram of the form presented in Fig. 2B [32] can be constructed. Network meta-analysis methods allow a simultaneous analysis of all the comparisons presented on the network. Table 3 [33] presents the ORs for the pairwise comparisons between the interventions produced both from the network meta-analysis and the direct comparisons from a trial or, when there was more than one trial, a pairwise meta-analysis of that particular comparison. In the network meta-analysis, the most intensive intervention (education + low-cost/free equipment + home safety inspection + fitting) was most effective for the possession of a fitted stair gate outcome compared with all other interventions. The probability that each intervention is best and the median rank (with uncertainty) of each intervention [31] calculated from the network meta-analysis are presented in Table 4. These data show that the most intensive intervention clearly had the highest probability (0.97) of being the most effective and a median rank of 1 (95% CI: 1, 2).

Although we believe such an analysis is more refined, interpretable, complete, and thus more helpful than the standard pairwise meta-analysis presented initially, it only considered summary study data, some of which were obtained from IPD, and did not include any potential treatment-modifying covariates. We had developed models to include covariates in network meta-analysis of summary data [34] (including a special model to deal with the inclusion of the control group event rate as a covariate in network meta-analysis [35], which is not illustrated here but potentially very useful in a PH context in which inequalities are of interest, particularly when IPD are not
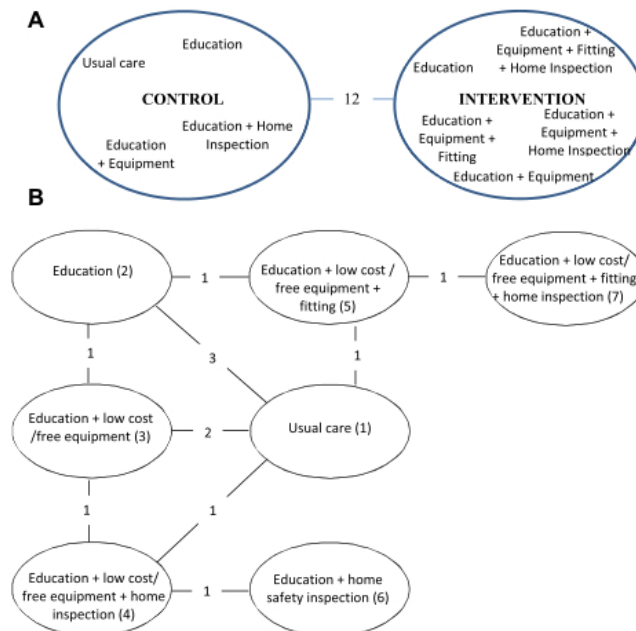
Fig. 2. Network diagrams indicating how intervention groups were defined and the number of studies in the (A) Cochrane review and (B) network meta-analysis.

available). We have also extended the network meta-analysis covariate model to allow the inclusion of IPD and thus subject-level covariates when possible [36].

4. Discussion

This review of completed NICE PH appraisals illustrates the current situation regarding the use of evidence synthesis methods to inform PH decision making in the United Kingdom. It identified that effectiveness evidence was mostly synthesized using narrative summaries and that quantitative synthesis was not carried out for most evaluations in PH systematic reviews. Of the 39 appraisals published since 2006, only 9 (23%) appraisals were informed by at least one systematic review with a meta-analysis. The other 30 appraisals may have refrained from meta-analysis because of a lack of randomized trials or heterogeneity in study design (ie, a mix of RCTs and non-RCTs). Moreover, systematic reviews opting for a quantitative summary tended to use the simplest methods such as fixed- or random-effects pairwise meta-analyses, which only enables comparison between two interventions at any one time and thus potentially limiting the scope of the analysis and the utility of the findings. These findings would seem to indicate that despite great advances in quantitative synthesis techniques, application in PH evaluation is still very much in its infancy and appears to lag behind other areas of health care such as the evaluation of clinical interventions. There are several reasons for this, not least due to the often heterogeneous nature of PH evidence including variations in many aspects of study design, including (1) the exact nature of the interventions, (2) outcome measures, (3) the wider scope of many PH research questions, and (4) the quantitative skills of the researchers involved.

Underlying our desire for PH reviews to become more quantitative, in the face of the challenges encountered, is a firm belief that a structured and transparent description and analysis of the decision question is desirable. Our review found that nearly 80% of NICE PH appraisals did not attempt a quantitative synthesis at all because of, what investigators believe but we want to challenge, insurmountable problems due to the heterogeneous nature of the evidence base. We believe that the more complex synthesis models, described in Section 3, can often more appropriately model the types of data commonly available in PH appraisals than carrying out less focused and detailed reviews of the literature.

NICE guidance states that ''Meta-analysis data may be used to produce a graph if the data (*usually from RCTs*)

Table 3. Results of a network meta-analysis (above stepped line) and pairwise meta-analysis (below stepped line) for possession of a fitted stair gate expressed as odds ratios (95% CrI) [a,b]

	Usual care (1)	Education (2)	Education + equipment (3)	Education + equipment + home inspection (4)	Education + equipment + fitting (5)	Education + home inspection (6)	Education + equipment + fitting + home inspection (7)
Usual care (1)		1.43 (0.90, 2.49)	1.63 (0.93, 3.03)	1.28 (0.69, 2.79)	1.52 (0.84, 3.38)	1.43 (0.56, 4.42)	7.80[c] (3.08, 21.3)
Education (2)	1.48 (0.97, 2.25)		1.14 (0.56, 2.23)	0.90 (0.41, 2.07)	1.07 (0.51, 2.41)	1.01 (0.33, 3.25)	5.46[c] (1.75, 16.12)
Education + equipment (3)	1.92[c] (1.05, 3.51)	1.17 (0.52, 2.63)		0.78 (0.38, 1.77)	0.94 (0.42, 2.41)	0.88 (0.32, 2.80)	4.77[c] (1.56, 15.18)
Education + equipment + home inspection (4)	1.13 (0.82, 1.58)		1.25 (0.49, 3.17)		1.20 (0.45, 3.25)	1.12 (0.52, 2.49)	6.13[c] (1.75, 18.71)
Education + equipment + fitting (5)	1.45[c] (1.18, 1.79)	1.63 (0.37, 7.23)				0.94 (0.27, 3.28)	5.07[c] (1.47, 15.93)
Education + home inspection (6)	7.73[c] (4.14, 14.43)			1.12 (0.86, 1.47)			5.48[c] (1.23, 20.73)
Education + equipment + fitting + home inspection (7)							

Abbreviation: CrI, credible interval.

Blank cells indicate that no direct evidence on specific pairwise comparisons was available.

[a] Values above the stepped line are results from the NMA; those below the line are direct estimates from a trial or, when more than one was available, a meta-analysis.

[b] Column and row headings signify intervention or comparison (intervention number).

[c] Significant at the 5% level.

Table 4. Assessment of which intervention is best for possession of a fitted stair gate

Intervention	Possession of a stair gate	
	Probability treatment is best	Median treatment rank (95% CrI)
Usual care (1)	0.00	7 (5, 7)
Education (2)	0.002	4 (2, 7)
Education + equipment (3)	0.004	3 (2, 7)
Education + equipment + home inspection (4)	0.001	5 (2, 7)
Education + equipment + fitting (5)	0.008	4 (2, 7)
Education + home inspection (6)	0.013	4 (2, 7)
Education + equipment + fitting + home inspection (7)	0.97	1 (1, 2)

is sufficiently homogenous'' (Section 5.4.4.2 n NICE guidance 2012 [16]). For PH reviews, the evidence from RCTs is often limited, and the best available evidence may be from non-RCTs, which reviewers may be reluctant to pool because of the risk of bias (Cochrane chapter 13 [9], Valentine and Thompson[37], and Moher et al. [38]). However, provided reviewers quality assess non-RCTs (as they would RCTs) to identify well-conducted studies, to limit confounding by selection bias, then meta-analysis can be considered.

Although concluding the evidence base to be "too heterogeneous for meta-analysis" may be better than carrying out a naive simple meta-analysis, not being able to present a quantitative analysis severely restricts the utility of the review, particularly for decision making. Exploring heterogeneity and attempting to account for it should be part of the analysis, and greater awareness of modern methods, and greater expertise in using them, will yield fruit for future PH reports. There are several other reasons why conducting a meta-analysis may not be advisable, however, for example, a small number of studies may mean that statistical heterogeneity is underestimated; some studies are too biased to draw a conclusion from them; there is evidence of publication bias; and insufficient reporting of outcomes.

We acknowledge that although softwares to undertake pairwise meta-analysis are widely available (eg, RevMan, Comprehensive Meta-Analysis), analyses such as the most complex ones previously described require advanced statistical expertise in evidence synthesis to implement (and some groundwork regarding the Bayesian theory underpinning such an approach may be required by nonstatistical PH specialists). Our software package of choice is WinBUGS. This is a freely available Bayesian simulation package [39] and is extremely powerful for fitting models not immediately available in other packages. (It even allows economic decision models to be included in the same program as the synthesis model, allowing a truly comprehensive assessment [40].) With the recent publication of the NICE technical support documents on evidence synthesis methodology [32,41,42] including all WinBUGS code to implement the models, together with more widely available specialist training courses and the new introductory

WinBUGS book [43], the time is ripe for getting to grips with the more complex evidence synthesis methodologies currently being embraced by health technology appraisals [1,44]. A detailed discussion of specific technical challenges in Bayesian random-effects synthesis models is available elsewhere [45].

This article is limited to only considering NICE PH appraisals in the review and does not claim to have all the answers to all evidence synthesis challenges that exist in PH evaluation. For example, none of the above analyses considers directly the influence of the study quality/validity of the individual studies going into an analysis, although others are doing work in other contexts that could be adapted, for example, including different, both observational and randomized, evidence [46].

Regarding the specific injury prevention context, for the analyses presented previously, even when categorizing the interventions into seven distinct groups, there is still residual heterogeneity in intervention definition, for example, education may be a leaflet designed for the prevention of an injury at home, it may also include a face-to-face interview, a computer-based questionnaire producing tailored advice based on the user answers, and so forth. We are developing further modeling extensions including how to extrapolate across a series of evidence networks to allow information sharing on the effectiveness of interventions in promoting other safety practices for the prevention of falls. We hope such analyses will be more efficient and robust than individual analyses of each outcome. Note that all the data considered only relate to an increase in the *uptake of safety practice* and not to reduction in accidents per se. Therefore, a further initiative is to develop models, which extend those presented to include the direct evidence between safety practices and injury data. This problem is similar to the use of surrogate end points in clinical evaluation, and we plan to adapt methods developed there.

PH evaluations are notoriously messy and complex, with many factors to consider. But if a decision has to be made, explicit, transparent, and appropriate analysis of the data should be preferred to current alternatives. Just as evaluations of clinical interventions are becoming more sophisticated, we think there is a pressing need to do the same for

PH contexts and we hope this article can contribute to the initiation of such an initiative.

Appendix

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2013.09.018.

References

[1] Cooper NJ, Peters J, Lai MCW, Juni P, Wandel S, Palmer S, et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? Value Health 2011;14(2): 371—80.

[2] Novielli N, Cooper NJ, Abrams KR, Sutton AJ. How is evidence on test performance synthesized for economic decision models of diagnostic tests? A systematic appraisal of health technology assessments in the UK since 1997. Value Health 2010;13(8):952—7.

[3] Higgins JPT, Green SE. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration 2011.

[4] Drummond MF, Sculpher M, Torrance GW, O'Brien B, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford, UK: Oxford University Press; 2005.

[5] Sutton AJ, Donegan S, Takwoingi Y, Garner P, Gamble C, Donald A. An encouraging assessment of methods to inform priorities for updating systematic reviews. J Clin Epidemiol 2009;62:241—51.

[6] Eden J, Levit L, Berg A, Morton S. Finding what works in health care: standards for systematic reviews. 2012. Available at http://www.nap.edu/catalog.php?record_id=13059. Accessed January 1, 2013.

[7] Rodgers M, Sowden A, Petticrew M, Arai L, Roberts H, Britten N, et al. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews. Evaluation 2009;15(1):49—73.

[8] Weatherly H, Drummond M, Claxton K, Cookson R, Ferguson B, Godfrey C, et al. Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. Health Policy 2009;93(2—3):85—92.

[9] Armstrong R, Waters E, Doyle Je. Chapter 21: reviews in health promotion and public health. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Chichester, UK: John Wiley & Sons; 2008.

[10] Rosen L, Manor O, Engelhard D, Zucker D. In defense of the randomized controlled trial for health promotion research. Am J Public Health 2006;96:1181—6.

[11] Petticrew M, Egan M. Relevance, rigour and systematic reviews. In: Popay J, ed. Moving beyond effectiveness in evidence synthesis. Methodological issues in the synthesis of diverse sources of evidence. 2006. Available at http://www.nice.org.uk/niceMedia/docs/Moving_beyond_effectiveness_in_evidence_synthesis2.pdf. Accessed June 6, 2012.

[12] Wanless D. Securing good health for the whole population: final report—February 2004 (Chapter 6). 2004. Available at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4074426. Accessed June 6, 2012.

[13] Chalkidou K, Culyer A, Naidoo B, Littlejohns P. Cost-effective public health guidance: asking questions from the decision-maker's viewpoint. Health Econ 2008;17(3):441—8.

[14] NICE. Methods for the development of NICE public health guidance. 2006. Available at http://www.nice.org.uk/media/FB9/59/PHMethodsManual2006.pdf. Accessed January 1, 2013.

[15] NICE. Methods for the development of NICE public health guidance. 2nd ed. 2009. Available at http://www.nice.org.uk/media/CE1/F7/CPHE_Methods_manual_LR.pdf. Accessed January 1, 2013.

[16] NICE. Methods for the development of NICE public health guidance. 3rd ed. 2012. Available at http://publications.nice.org.uk/methods-for-the-development-of-nice-public-health-guidance-third-edition-pmg4. Accessed January 1, 2013.

[17] Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H. Publication bias and clinical-trials. Control Clin Trials 1987;8:343—53.

[18] Sutton A, Abrams K, Jones DR, Sheldon TA, Song F. Methods for meta-analysis in medical research. Chichester, UK: Wiley; 2000.

[19] Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331:897—900.

[20] Caldwell DM, Welton NJ, Ades AE. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. J Clin Epidemiol 2010;63:875—82.

[21] Ades AE, Mavranezouli I, Dias S, Welton NJ, Whittington C, Kendall T. Network meta-analysis with competing risk outcomes. Value Health 2010;13(8):976—83.

[22] Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med 2002;21:2313—24.

[23] Kendrick D, Young B, Mason-Jones AJ, Ilyas N, Achana FA, Cooper NJ, et al. Home safety education and provision of safety equipment for injury prevention. Cochrane Database Syst Rev 2012;(9):CD005014.

[24] Kendrick D, Sutton A, Smith S, Mulvaney C, Woods A, Watson M, et al. Does the effect of home safety education, with or without the provision of safety equipment, differ by social group? Systematic review and meta-analysis. Report to the Department of Health 2006.

[25] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557—60.

[26] Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. J Clin Epidemiol 2002;55:86—94.

[27] Stewart LA. Practical methodology of meta-analyses (overviews) using updated individual patient data. Stat Med 1995;14:2057—79.

[28] Kendrick D, Coupland C, Mulvaney C, Simpson J, Smith SJ, Sutton A, et al. Home safety education and provision of safety equipment for injury prevention. Cochrane Database Syst Rev 2007;(1): CD005014.

[29] Sutton AJ, Kendrick D, Coupland CAC. Meta-analysis of individual-and aggregate-level data. Stat Med 2008;27:651—69.

[30] Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 2010;340: c221.

[31] Dias S, Sutton AJ, Ades AE, Welton NJ. A generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making 2013;33:607—17.

[32] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2011. Available at http://www.nicedsu.org.uk. Accessed January 1, 2013.

[33] Hubbard S, Sutton A, Kendrick D, Young B, Achana F, Cooper N. Network meta-analysis to evaluate the effectiveness of interventions to prevent falls injuries in children under 5. Leicester, UK: University of Leicester; 2013.

[34] Cooper NJ, Kendrick D, Achana F, Dhiman P, He Z, Wynn P, et al. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms. Epidemiol Rev 2012;34: 32—45.

[35] Achana FA, Cooper NJ, Dias S, Lu GB, Rice SJC, Kendrick D, et al. Extending methods for investigating the relationship between

treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. Stat Med 2013;32:752—71.

[36] Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. Stat Med 2012;31:3516—36.

[37] Valentine JC, Thompson SG. Issues relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. Res Synth Methods 2013; 4(1):26—35.

[38] Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Lancet 1999;354: 1896—900.

[39] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10(4):325—37.

[40] Cooper NJ, Abrams KR, Sutton AJ, Turner D, Lambert PC. A Bayesian approach to Markov modelling in cost-effectiveness analyses: application to taxane use in advanced breast cancer. J R Stat Soc Ser A Stat Soc 2003;166(3):389—405.

[41] Dias S, Sutton AJ, Welton NJ, Ades AE. NICE DSU technical support document 3: heterogeneity: subgroups, meta-regression, bias and bias-adjustment. 2011. Available at http://www.nicedsu.org.uk. Accessed November 8, 2011.

[42] Dias S, Welton NJ, Sutton AJ, Ades AE. NICE technical support document 4: inconsistency in networks of evidence based on randomised trials. 2011. Available at http://www.nicedsu.org.uk. Accessed January 1, 2013.

[43] Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton, FL; CRC Press; 2012.

[44] Cooper NJ, Spiegelhalter D, Bujkiewicz S, Dequen P, Sutton A. A systematic review of the deliberate and accidental use of Bayesian methods in health technology assessment. Int J Technol Assess Health Care 2013; 29:336—42.

[45] Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc 2009; 172:137—59.

[46] Turner RM, Spiegelhalter DJ, Smith GCS, Thompson SG. Bias modelling in evidence synthesis. J R Stat Soc Ser A Stat Soc 2009; 172:21—47.

Network meta-analysis to evaluate the effectiveness of interventions to prevent falls in children under age 5 years

Stephanie Hubbard,[1] Nicola Cooper,[1] Denise Kendrick,[2] Ben Young,[2] Persephone M Wynn,[2] Zhimin He,[2] Philip Miller,[2] Felix Achana,[1] Alex Sutton[1]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/injuryprev-2013-041135).

[1]Department of Health Sciences, University of Leicester, Leicester, UK
[2]Division of Primary Care, University of Nottingham, Nottingham, UK

Correspondence to
Stephanie Hubbard, Department of Health Sciences, Adrian Building, University of Leicester, University Road, Leicester LE1 7RH, UK; sjh62@le.ac.uk

ABSTRACT

Background This study aimed to simultaneously evaluate the effectiveness of a range of interventions to increase the possession of safety equipment or behaviours to prevent falls in children under 5 years of age in the home.

Methods A recently published systematic review identified studies to be included in a network meta-analysis; an extension of pairwise meta-analysis that enables comparison of all evaluated interventions *simultaneously*, including comparisons not directly compared in individual studies.

Results 29 primary studies were identified, of which 16 were included in at least 1 of 4 network meta-analyses. For increasing possession of a fitted stair gate, the most intensive intervention (including education, low cost/free home safety equipment, home safety inspection and fitting) was the most likely to be the most effective, with an OR versus usual care of 7.80 (95% CrI 3.08 to 21.3). For reducing possession or use of a baby walker: education only was most likely to be most effective, with an OR versus usual care of 0.48 (95% CrI 0.31 to 0.84). Little difference was found between interventions for possession of window locks (most intensive intervention versus usual care OR=1.56 (95% CrI 0.02 to 89.8)) and for not leaving a child alone on a high surface (education vs usual care OR=0.89 (95% CrI 0.10 to 9.67)). There was insufficient evidence for network meta-analysis for possession and use of bath mats.

Conclusions These results will inform healthcare providers of the most effective components of interventions and can be used in cost-effectiveness analyses.

INTRODUCTION

Across the world, falls are a leading cause of morbidity and mortality in children.[1–4] Mortality rates from falls in childhood are highest in children under 1 year old[1] and, among 0–4-year-olds, they are the 12th leading cause of disability-adjusted life years lost.[5] Most falls occur at home in the 0–4-year-olds.[6] Falls place a considerable burden on healthcare systems globally as they are the most common injury among emergency department attenders, comprising between 25% and 52% of all treated child injuries.[7][8] Falls are the leading cause of injury-related healthcare costs in the USA, accounting for just over one quarter of all injury-related healthcare costs in childhood, totalling US$81 billion in 1996.[9]

Despite the high burden of injury attributable to falls, there is little evidence that home safety interventions reduce falls rates or promote falls prevention practices. A recent overview of systematic reviews on preventing childhood falls within the home identified one meta-analysis and 13 systematic reviews, and included a total of 29 relevant primary studies[10] Evidence of the effect of interventions on falls or fall injuries was sparse, with one of three primary studies reporting this outcome finding a significant reduction in falls.[11] Interventions were effective in promoting the use of safety gates and furniture corner covers. There was some evidence of a reduction in baby walker use. The effect on the use of window safety devices, non-slip bath mats/decals and the reduction of tripping hazards were mixed. One meta-analysis was included in the overview[12] which found that families receiving home safety interventions were: significantly more likely to have a fitted stair gate; less likely to use a baby walker; not significantly different in their possession of window locks, non-slip bath mats or decals; or reporting leaving a child alone on a high surface. There was significant heterogeneity between effect sizes for the fitted stair gate and baby walker outcomes. This meta-analysis evaluated any intervention against a 'usual care or no intervention' comparison group. The interventions, and in some studies the control arm, comprised various combinations of education, home safety inspection, provision of free or low-cost safety equipment, and fitting of safety equipment. Some interventions were aimed at only preventing fall-related injuries, while others aimed to prevent a range of injuries. In reality, healthcare commissioners, and housing providers, among others, have to make policy decisions on the 'best' intervention(s) for preventing fall-related injuries, so lumping together interventions is not particularly useful.

Standard (pairwise) meta-analyses are usually restricted to finding a pooled estimate of effectiveness comparing two groups, often an intervention group with a control group and, hence, only identifies one as being superior to the other. This can make only a limited contribution to policy decisions.[13] An extension to this that enables comparison of all evaluated interventions *simultaneously* within a single coherent analysis is network meta-analysis (NMA),[14] also known as mixed treatment comparison.[13][15] NMA allows all interventions to be compared with one another, including comparisons not directly evaluated within any of

254

the primary studies. Interventions can also be ranked in order of effectiveness. Such an approach is being increasingly used in health technology assessment when deciding on the optimal intervention strategy for a given medical condition.[16 17]

Suppose we have studies providing effect estimates for control versus an intervention A, and for intervention A versus an intervention B. NMA allows us to estimate the pooled effects where pairwise evidence exists (direct comparison between control and A and between A and B), and also allows us to estimate effects where interventions are not directly compared but are linked through a connected network of studies (indirect comparison between control and B). If evidence is available on all comparisons between control, A and B, the indirect evidence is pooled with direct data from the studies, hence, inference is based on more evidence and uncertainty should be reduced. All evidence is combined in a single model and details can be found in the Nice DSU Technical Support Document 2.[18] NMA is particularly relevant to the field of injury prevention, where interventions are often complex and multifaceted, and the number of studies evaluating the same comparisons is small.

The objective of our research was to evaluate the effectiveness of different interventions to increase the possession of safety equipment by households or increase falls prevention behaviours. This application is part of a series of NMAs evaluating a range of interventions to prevent injuries in preschool children in the home. The first paper to be published in this series reported NMA to evaluate the effectiveness of interventions to increase the uptake of smoke alarms.[19]

METHODS
Study identification
A recently published overview of systematic reviews and a systematic review of primary studies published since the most recent comprehensive review[10] sourced studies for the NMA. The primary studies included in the published review reported non-legislative interventions aimed at primary or secondary prevention of falls at home among children aged 0–19 years, and reported medically or non-medically attended falls, possession or use of home safety equipment to prevent falls, or other falls prevention practices. The sources searched for the review are summarised in online supplementary appendix table 1A, and more details (including the search strategies used, and the inclusion and exclusion criteria applied) are available in Young et al.[10]

Study quality of the primary studies was assessed using: allocation concealment, blinding of outcome assessment, and completeness of follow-up for randomised studies; blinding of outcome assessment, completeness of follow-up, and balance of confounders between treatment arms for non-randomised studies; and the Newcastle-Ottawa scale[20] for controlled observational (case-control and cohort) studies.

Statistical methods
The main outcomes analysed for the NMAs were: possession of a fitted stair gate; possession or use of a baby walker; possession of a bath mat or decals; possession of window locks; and never leaving a child alone on a high surface. Possession of a fitted stair gate could refer to a gate at the top and/or bottom of the stairs or a safety gate preventing access to an unsafe area; most of the original studies did not specify and only looked at possession. For the baby walker outcome, we considered it appropriate to combine baby walker possession or use as one of the included studies[21] found that 94% of those owning a walker used it and 96% of those who used a walker owned one.

NMA[13–15] was implemented to compare a range of different interventions including 'usual care' and to include studies where the control arm was another intervention. It allows us to estimate the pooled effects where pairwise evidence exists (direct evidence) and also to estimate effects where interventions are not directly compared but linked through a connected network of studies (indirect evidence). For example, if we wanted to compare the following four interventions—usual care, education, safety equipment giveaway, home safety inspection— this could be achieved by using studies containing the following direct pairwise comparisons: usual care versus education, education versus safety equipment giveaway, and safety equipment giveaway versus home safety inspection (figure 1A), and by tracing a comparison pathway through the direct pairwise comparisons to estimate, for example, the indirect comparison of usual care versus safety equipment giveaway. However, the network would be disconnected, and the analysis impossible, if only studies of usual care versus education, and safety equipment giveaway versus home safety inspection existed (figure 1B).

For our analyses, a standard NMA random-effects model with a binary outcome[13–15] was fitted to the data that allows us to include trials with three or more arms by accounting for the correlation structure.[22 23] If studies did not adjust for clustered allocation of intervention, the effective sample size was
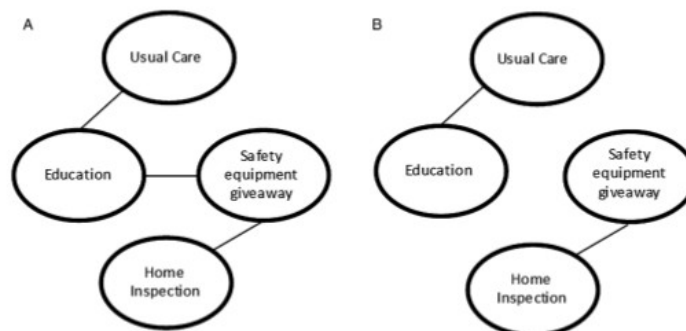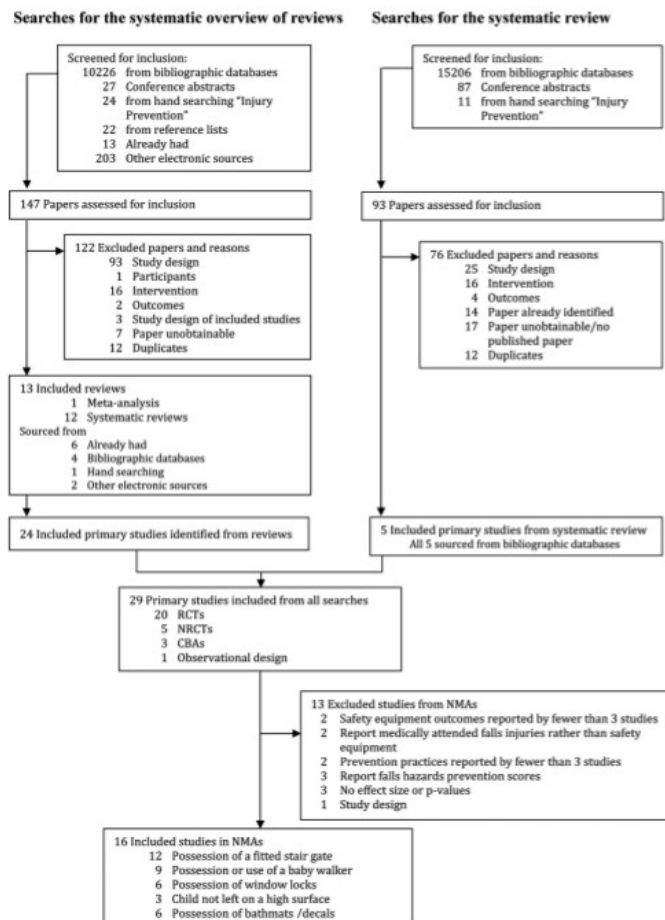

Figure 1 (A) Connected Network; (B) Disconnected Network. Line represent available direct evidence.

Figure 2 PRISMA flow chart for the systematic overview of reviews and systematic review of primary studies.

estimated based on the design effect using published intraclass correlation coefficients (ICC),[24] or ICCs estimated from individual participant data where the author(s) provided it. We obtained pooled estimates of intervention effects, expressed as ORs, and 95% credible intervals (CrI) for all combinations of pairwise comparisons from the NMAs using a combination of direct and indirect evidence, and indirect evidence only. For completeness, pooled estimates from the direct evidence only are presented for each pairwise comparison where study data was available (using a fixed effect meta-analysis model when only two studies were available for a particular pairwise comparison, and a random effects model where three or more studies were available, and where only one study had evaluated a particular pairwise comparison, the results from this study alone). From the NMA, intervention effectiveness was ranked based on absolute intervention effects (derived by using an

underlying rate based on the usual-care arms) and the probability that each intervention was best for a particular outcome was calculated.[14]

To assess the goodness of fit of the model to the data, the posterior mean residual deviance[25][26] was calculated. For an adequately fitting model it will be approximately equal to the number of treatment arms across all studies.[27][28] Heterogeneity of the network (variability in treatment effects within pairwise comparisons above that expected by chance) was quantified by using the between-study SD parameter where a SD of below 0.5 indicated fairly low heterogeneity, and above 1, substantial heterogeneity.[23][29][30] This model assumes that the degree of between-study, within-comparison heterogeneity is constant across all intervention comparisons in the network.

Inconsistency, where direct and indirect evidence are available and do not agree (beyond chance), was assessed using the

256

Table 1 Summary of Studies and their data included in the NMA of the interventions to prevent falls injuries in children under 5 (Numbers adjusting for clustering in parentheses)

Comparison (intervention number)	First author, year (reference No.)	Study design	Follow-up (months)	Study quality	Intervention	Stair gate		Baby walker		Window locks		High surfaces		Bath mats	
						Number of households with stair gates	Total number of households	Number of households with or using baby walkers	Total number of households	Number of households with window locks	Total number of households	Number of households not leaving child on high surface	Total number of households	Number of households with bath mats	Total number of households
Usual care (1) versus education (2)	Nansel 2002[33]*	RCT	0.75	A=Y, B=Y, F=Y	1 / 2	70 / 76	89 / 85	30 / 19	89 / 85						
	Kendrick 2005[21]	RCT	9	A=Y, B=N, F=Y	1 / 2	418 (348.44) / 373 (310.93)	524 (436.80)† / 452 (376.78)	230 (105.27) / 131 (59.96)	543 (248.52)† / 463 (211.90)						
	Nansel 2008[34]†	Non-RCT	1	B=N, F=N, C=N	1 / 2	29 / 60	38 / 69	12 / 13	38 / 69			21 / 55	24 / 62		
	Tan 2004[35]§	Non-RCT	5	B=U, F=Y, C=Y	1 / 2			393 / 143	480 / 228						
Usual care (1) versus Education +Low/free equipment (3)	Babul 2007[36]¶	RCT	10	A=Y, B=N, F=N	1 / 3			31 / 22	148 / 162			69 / 89	148 / 161		
	Clamp 1998[37]	RCT	1.5	A=Y, B=N, F=Y	1 / 3	50 / 52	69 / 64			72 / 80	82 / 83				
	McDonald 2005[38]	RCT	1	A=Y, B=U, F=N	1 / 3	10 / 23	41 / 54								
Usual care (1) versus Education +Low/free equipment +Home safety inspection (4)	Babul 2007[36]¶	RCT	10	A=Y, B=N, F=N	1 / 4			31 / 26	148 / 173			69 / 84	148 / 170		
	Kendrick 1999[39]	Non-RCT	25	B=N, F=N, C=Y	1 / 4	241 (214.26) / 251 (223.15)	364 (323.61)** / 364 (323.61)			339 / 323	366 / 362				
	Hendrickson 2002[40]	RCT	1.5	A=N, B=N, F=Y	1 / 4					21 / 24	39 / 34				
Usual care (1) versus Education +Low/free equipment +Fitting (5)	Watson 2005[41]	RCT	12	A=Y, B=N, F=N	1 / 5	328 / 408	718 / 742			493 / 550	741 / 767				
Usual care (1) versus Education +Home safety inspection (6)	Petridou 1997[42]	CBA	20	B=N, F=Y, C=Y	1 / 6									64 (48.91) / 66 (50.44)	128 (97.83)†† / 131 (100.12)
Usual care (1) versus	Phelan 2010[43]	RCT	12	A=Y, B=N, F=N	1 / 7	78 / 131	147 / 146	29 / 24	138 / 140	145 / 146	150 / 149			56 / 59	150 / 149

Continued

Table 1 Continued

Comparison	Study	Design	Follow-up	A	B	F								
Education +Low/free equipment +Fitting+Home safety inspection (7)				A=Y, B=N, F=Y										
Education (2) versus Education +Low/free equipment (3)	Posner 2004[44]	RCT	2.5	A=Y	B=N	F=Y	2 / 3	25 / 28	47 / 49	4 / 4	8 / 7	44 / 12	50 / 17	34 / 44	47 / 49
Education (2) versus Education +Low/free equipment +Fitting (5)	Sznajder 2003[45]	RCT	1.5 to 2	A=Y	B=N	F=N	2 / 5	45 / 44	50 / 47	14 / 19	50 / 47			37 / 31	49 / 48
Education +low/free equipment (3) versus Education +low/free equipment +Home safety inspection (4)	Gielen 2002[46]	RCT	12	A=U	B=U	F=N	3 / 4	11 (12.85) / 13 (10.87)	48 (47.44)‡‡ / 48 (47.44)	22 / 26	162 / 173				
Education +Low/free equipment +Home safety inspection (4)	Babul 2007[38]¶	RCT	10	A=Y	B=N	F=N	3 / 4					89 / 84	161 / 170		
Education +Home safety inspection (4) versus Education +Home safety inspection (6)	King 2001[47]	RCT	12	A=Y	B=Y	F=Y	4 / 6	158 / 166	482 / 469	29 / 33	482 / 469	285 / 299	469 / 482		

*Two intervention arms were combined (tailored advice and tailored advice+care provider feedback).

†ICC calculated from individual participant data (IPD) from Kendrick 2005.[21]

‡Generic safety advice was counted as usual care.

§Two control arms were combined (usual care and usual care+baseline questionnaire).

¶Babul has been included in the NMA as a three-arm trial but is listed above as three separate comparisons.

**ICC calculated from IPD from Kendrick 1999.[39]

††Adjusted for clustering using an ICC estimated as the mid-point of a range of ICCs published for health outcomes at the level of health authority, local authority or town.[24]

‡‡ICC calculated from IPD from Kendrick 1999[39] and Kendrick 2005.[21] A, adequate allocation concealment; B, blinded outcome assessment; C, prevalence of confounders does not differ by more than 10% between treatment arms; CBA, controlled before-and-after study; F, at least 80% participants followed up in each arm; ICC, intraclass correlation coefficient; N, no; NMA, network meta-analysis; RCT, randomised clinical trial; U, unclear; Y, yes.
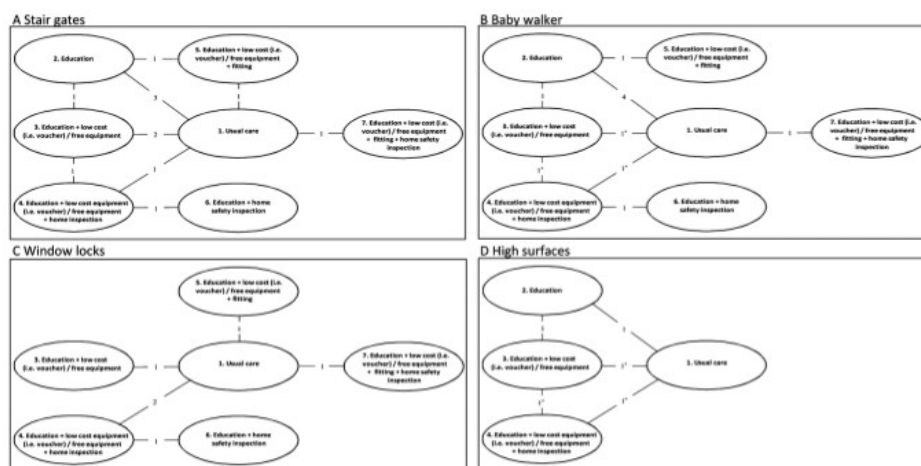
Figure 3 Network diagrams of interventions to increase safety practices to prevent falls in pre-school children in the home. Each intervention is a node in the network. The links between the nodes are pairwise intervention comparisons. The numbers along the link lines indicate the number of studies or pairs of study arms for that link in the network. The network for bath mats is presented in the online supplementary appendix. * Babul three-arm trial (A) Stair gates (B) Baby walker (C) Window locks (D) High surfaces.

method based on node splitting[31] which compares the results from the direct evidence to the results from the NMA with the direct evidence excluded.

To explore the effect of the variable study quality of the included studies, analyses were repeated restricted to randomised clinical trials (RCT) and, to explore the variability of the interventions, analyses were repeated by splitting the interventions further.

The NMAs were conducted using a Markov chain Monte Carlo method[14] with minimally informative prior distributions fitted in the WinBUGS software[28] and tests for inconsistency were carried out in R.[31] Further technical details of the analysis, together with the WinBUGS and R code, are available from the corresponding author. The analysis and subsequent reporting adhere to the PRISMA statement guidelines.[32]

RESULTS

The published overview of systematic reviews and systematic review of primary studies[10] identified 29 primary studies eligible for inclusion in the NMA (24 from the overview of systematic reviews and 5 from the systematic review of primary studies). Of these, 13 were excluded from the NMA (figure 2). Online supplementary appendix table 1B shows the bibliographic databases in which the articles appeared. The characteristics of the 16 studies included in the NMAs are reported in table 1 together with their study quality which was observed to be variable across studies. In total, seven interventions were evaluated (although not all interventions had been trialled for each outcome):
1. usual care
2. education
3. education+low cost (ie, voucher)/free equipment
4. education+low cost (ie, voucher)/free equipment+home safety inspection
5. education+low cost (ie, voucher)/free equipment+fitting
6. education+home safety inspection

7. education+low cost (ie, voucher)/free equipment+fitting +home safety inspection.

Possession of a fitted stair gate

Ten (83%) of the 12 studies reporting the possession of a fitted stair gate outcome were RCTs, and 2 (17%) were non-RCTs (table 1). Figure 3A displays a connected network diagram for this NMA.

The NMA estimated the 21 possible pairwise comparisons between the seven interventions. The pooled estimates, along with the available direct within-trial estimates are reported in table 2A.

The most intensive intervention (education+low cost/free equipment+home safety inspection+fitting) was most likely to be effective (probability best=0.97, table 3) in increasing possession, with, for example, families in the intensive intervention group more likely to possess a fitted stair gate compared with those in the usual care group (OR=7.80 (95% CrI 3.08 to 21.3)).

The effect of study design on the NMA results was assessed by repeating the above analysis using only data from the 10 RCTs. The result was similar with the most intensive intervention identified as being the most likely to be effective (probability best=0.87), with an estimated OR for possession versus usual care of 7.93 (95% CrI 2.76 to 23.6).

Possession or use of a baby walker

Seven (78%) of the nine studies for the possession or use of a baby walker outcome were RCTs and two (22%) were non-RCTs (table 1, figure 3B).

The NMA estimated the 21 possible pairwise comparisons between the 7 interventions (table 2B). The education-only intervention was the most effective (probability best=0.65, table 3), with families in the education-only intervention group less likely to possess or use a baby walker compared to usual care (OR 0.48, 95% CrI 0.31 to 0.84).

Table 2 Results of NMA models for interventions to increase safety practices to prevent falls in preschool children in the home expressed as ORs. Direct comparison estimates are also displayed.*,†

(A) Stair gates

(A) Stair gates	Usual care (1)	Education (2)	Education+Equipment (3)	Education+Equipment+Home inspection (4)	Education+Equipment+Fitting (5)	Education+Home inspection (6)	Education+Equipment+Fitting+Home inspection (7)
Usual care (1)		1.43 (0.90 to 2.49)	1.63 (0.93 to 3.03)	1.28 (0.69 to 2.79)	1.52 (0.84 to 3.38)	1.43 (0.56 to 4.42)	7.80‡ (3.08 to 21.3)
Education (2)	1.48 (0.97 to 2.25)		1.14 (0.56 to 2.23)	0.90 (0.41 to 2.07)	1.07 (0.51 to 2.41)	1.01 (0.33 to 3.25)	5.46‡ (1.75 to 16.1)
Education+Equipment (3)	1.92‡ (1.05 to 3.51)	1.17 (0.52 to 2.63)		0.78 (0.38 to 1.77)	0.94 (0.42 to 2.41)	0.88 (0.32 to 2.80)	4.77‡ (1.56 to 15.2)
Education+Equipment+Home inspection (4)	1.13 (0.82 to 1.58)		1.25 (0.49 to 3.17)		1.20 (0.45 to 3.25)	1.12 (0.52 to 2.49)	6.13‡ (1.75 to 18.7)
Education+Equipment+Fitting (5)	1.45‡ (1.18 to 1.79)	1.63 (0.37 to 7.23)		1.12 (0.86 to 1.47)		0.94 (0.27 to 3.28)	5.07‡ (1.47 to 15.9)
Education+Home inspection (6)							5.48‡ (1.23 to 20.7)
Education+Equipment+Fitting+Home inspection (7)	7.73‡ (4.14 to 14.4)						

(B) Baby walkers

(B) Baby walkers	Usual care (1)	Education (2)	Education+Equipment (3)	Education+Equipment+Home inspection (4)	Education+Equipment+Fitting (5)	Education+Home inspection (6)	Education+Equipment+Fitting+Home inspection (7)
Usual care (1)		0.48‡ (0.31 to 0.84)	1.51 (0.56 to 3.65)	1.07 (0.37 to 2.89)	0.85 (0.29 to 3.35)	1.28 (0.29 to 5.06)	0.78 (0.27 to 2.22)
Education (2)	0.46‡ (0.36 to 0.58)		3.15‡ (1.02 to 8.38)	2.25 (0.66 to 6.24)	1.80 (0.53 to 6.24)	2.68 (0.55 to 10.93)	1.63 (0.47 to 4.96)
Education+Equipment (3)	0.59 (0.33 to 1.08)	1.33 (0.17 to 10.3)		0.70 (0.26 to 2.02)	0.56 (0.13 to 2.97)	0.84 (0.21 to 3.52)	0.52 (0.13 to 2.22)
Education+Equipment+Home inspection (4)	0.67 (0.38 to 1.19)		0.89 (0.48 to 1.64)		0.79 (0.16 to 4.71)	1.20 (0.43 to 3.26)	0.73 (0.17 to 3.24)
Education+Equipment+Fitting (5)	0.78 (0.43 to 1.42)	1.75 (0.75 to 4.08)				1.51 (0.20 to 9.30)	0.91 (0.16 to 4.47)
Education+Home inspection (6)							0.61 (0.10 to 3.69)
Education+Equipment+Fitting+Home inspection (7)					0.85 (0.51 to 1.42)		

(C) Window locks

(C) Window locks	Usual care (1)	Education+Equipment (3)	Education+Equipment+Home inspection (4)	Education+Equipment+Fitting (5)	Education+Home inspection (6)	Education+Equipment+Fitting+Home inspection (7)
Usual care (1)		4.09 (0.27 to 67.9)	1.05 (0.19 to 6.89)	1.28 (0.11 to 14.2)	1.10 (0.057 to 25.2)	1.74 (0.11 to 30.5)
Education+Equipment (3)	0.71 (0.46 to 1.11)		0.26 (0.010 to 7.29)	0.31 (0.007 to 11.4)	0.27 (0.004 to 17.1)	0.42 (0.008 to 21.0)
Education+Equipment+Home inspection (4)	0.93 (0.31 to 2.80)			1.24 (0.054 to 22.1)	1.06 (0.092 to 12.6)	1.65 (0.055 to 44.1)
Education+Equipment+Fitting (5)	0.78 (0.63 to 0.98)	0.95 (0.72 to 1.24)			0.85 (0.020 to 44.8)	1.38 (0.034 to 57.2)
Education+Home inspection (6)	0.60 (0.09 to 3.13)					1.56 (0.024 to 89.8)
Education+Equipment+Fitting+Home inspection (7)						

260

(D) High surfaces	Usual care (1)	Education (2)	Education+Equipment (3)	Education+Equipment+Home inspection (4)
Usual care (1)		1.94 (0.20 to 14.4)	1.13 (0.12 to 6.75)	1.01 (0.089 to 9.18)
Education (2)	0.89 (0.18 to 5.84)		0.56 (0.064 to 4.65)	0.50 (0.032 to 8.76)
Education+Equipment (3)	0.71 (0.44 to 1.13)	3.06 (0.61 to 14.3)		0.89 (0.099 to 9.67)
Education+Equipment+Home inspection (4)	0.89 (0.56 to 1.42)		1.27 (0.80 to 2.00)	

*Values above the diagonal are results from the NMA, OR with 95% CrI; those below the line are direct estimates from a trial or, where more than one were available, a meta-analysis with 95% CI. Blank cells indicate that no direct evidence on specific pairwise comparisons was available.
†Column and row headings signify intervention or comparison (intervention number).
‡Significant at the 5% level.
CrI, credible interval; NMA, network meta-analysis.

The effect of study design on the results of the NMA results was assessed by repeating the above analysis using only data from the seven RCTs. The result was similar with the education-only intervention identified as being the most likely to be effective (probability best=0.45), with an estimated OR versus usual care of 0.58 (95% CrI 0.21 to 1.87).

Possession of window locks

Five (83%) of the six studies for the possession of window locks outcome were RCTs and one (17%) was a non-RCT (table 1, figure 3C).

The NMA estimated the 15 possible pairwise comparisons between six interventions (excluding the education-only intervention) (table 2C). Education+Home safety inspection was most likely to be effective (probability best=0.26, table 3), but there was very little difference between any of the interventions (OR for Education+Low cost/free equipment+Fitting vs Usual care of 1.56 (95% CrI 0.02 to 89.8)). Repeating the analysis using only data from the five RCTs gave similar results.

Child not left on a high surface

Only three studies reported the numbers who left a child on a high surface; two (67%) were RCTs and one was a non-RCT (table 1, figure 3D).

The NMA estimated the six possible pairwise comparisons between four interventions. The pooled estimates, along with the available direct within-trial estimates are reported in table 2D. There was very little difference between any of the interventions, but education-only was the least likely to be effective in preventing children being left on high surfaces (probability best=0.10, table 3 and OR vs Education+Low cost/free equipment of 0.56 (95% CrI 0.06 to 4.65) and versus Education +Low cost/free equipment+Home inspection of 0.50 (95% CrI 0.03 to 8.76).

Possession of bath mats

Four studies reported the possession of bath mats (table 1). Three (75%) of the studies were RCTs and one (25%) was a controlled before-and-after study (CBA). Online supplementary appendix figure 1 displays a network diagram for this NMA showing two unconnected networks of three interventions, so we were unable to use NMA for this outcome.

Evaluation of models

Overall, the NMA models fitted the data well with the posterior mean residual deviance being close to the number of data points in each network (table 3).

The between-study SDs for each of the NMA models are reported in table 3. The uncertainty in their estimation reflects the relatively low number of studies providing direct evidence for each pairwise comparison, especially for the not leaving a child alone on a high surface outcome. Consistency was checked between the direct and indirect evidence by using node-splitting methodology. This can only be done when a pair of interventions is part of a closed loop in the network. Any closed loops in the networks were checked for consistency between the direct and indirect evidence. There was no evidence of inconsistency for either the stair gate (see online supplementary appendix table 4) or baby walker outcomes (see online supplementary appendix table 5). There were no closed loops for the possession of window locks, high surfaces and bath mat outcomes.

261

Table 3 Assessment of best intervention and model fit

Intervention	Possession of a stair gate		No possession or use of a baby walker		Possession of window locks		Not leaving a child alone on a high surface	
	Probability treatment is best	Median treatment rank (95% CrI)	Probability treatment is best	Median treatment rank (95% CrI)	Probability treatment is best	Median treatment rank (95% CrI)	Probability treatment is best	Median treatment rank (95% CrI)
Usual care (1)	0.00	7 (5 to 7)	0.001	4 (2 to 7)	0.15	3 (1 to 5)	0.32	2 (1 to 4)
Education (2)	0.002	4 (2 to 7)	0.65	1 (1 to 4)			0.10	4 (1 to 4)
Education+Equipment (3)	0.004	3 (2 to 7)	0.007	6 (2 to 7)	0.052	6 (1 to 6)	0.23	2 (1 to 4)
Education+Equipment +Home safety inspection (4)	0.001	5 (2 to 7)	0.027	5 (1 to 7)	0.16	3 (1 to 6)	0.35	2 (1 to 4)
Education+Equipment +Fitting (5)	0.008	4 (2 to 7)	0.13	3 (1 to 7)	0.19	4 (1 to 6)		
Education+Home safety inspection (6)	0.013	4 (2 to 7)	0.049	6 (1 to 7)	0.26	3 (1 to 6)		
Education+Equipment +Fitting+Home safety inspection (7)	0.97	1 (1 to 2)	0.13	3 (1 to 7)	0.18	4 (1 to 6)		
Model fit								
Posterior mean residual deviance	22.51 (cf 24 data points)		17.95 (cf 19 data points)		12.49 (cf 12 data points)		7.09 (cf 7 data points)	
Between-study SD	0.23 (0.015 to 0.87)		0.24 (0.0094 to 1.14)		1.02 (0.065 to 1.95)		1.31 (0.26 to 374.9)	

Sensitivity analysis

The network for the stair gate outcome was extended to comprise nine different intervention groups by splitting the low cost/free safety equipment giveaway included in interventions into relevant and not relevant/not stated (see online supplementary appendix figure 2). The findings from this analysis (see online supplementary appendix table 4) were similar in that the most intensive intervention clearly was the most effective in increasing the possession of a stair gate, and it also showed that there was very little difference between the interventions with low cost/free relevant or not relevant equipment. There is no relevant equipment for reducing the possession of baby walkers, and the networks were too sparse for the other outcomes to be extended further.

DISCUSSION

NMA was used to compare and evaluate the different interventions to increase a range of safety practices to prevent falls in preschool children in the home. Using this method enabled all strategies, including those not addressed in any of the individual primary studies, to be compared, and interventions could be ranked to identify the most effective intervention(s) in promoting safety practices to prevent falls. The findings showed that the most intensive intervention was most effective in increasing stair gate possession, and education-only was most effective in reducing baby walker possession and use. The findings were inconclusive for the possession of window locks, child not left on a high surface, and possession of bath mats.

An updated Cochrane review[24] based on pairwise meta-analyses of the same studies found that families in the home safety intervention group were more likely to possess a fitted stair gate, less likely to have or use a baby walker and more likely to possess window locks than families in a control arm. Because of the complexity and number of different interventions in the intervention and control arms and the number of studies available for the different combinations, the results from the NMA are more likely to be useful to policy makers, service commissioners and providers when making choices between multiple alternatives, than those from the pairwise meta-analyses.

A key limitation of our analyses is that, although we were able to categorise the interventions we studied to a greater degree than in previous meta-analyses,[12] there is potentially still some 'lumping' of interventions within these categories. For example, the intensity of education interventions may differ markedly between studies (ie, from providing only educational information, such as leaflets, to providing intensive face-to-face education on home safety), and the low cost/free safety equipment may not have been relevant to the outcome concerned (ie, equipment may have included socket covers and smoke alarms that would not prevent fall injuries, and there was no directly relevant safety equipment for the possession and use of a baby walker outcome). Similarly, different levels of usual care exist across the populations recruited in the primary studies. Additionally, the definitions of low-cost equipment used by our included studies varied between studies. Costs may also not be comparable between studies conducted in different countries, or between populations within one country with very different income levels and economic conditions. Insufficient detail was presented in many of primary reports to enable us to subcategorise the interventions further. It would be helpful if future studies provided sufficient details of interventions to enable more detailed NMAs to be conducted. Sensitivity analyses showed that findings were robust to splitting the interventions further and to study quality. Due to small numbers of studies, we were not able to explore the impact of allocation concealment, blinding, and percentage follow-up separately; however, rerunning the NMA analysis using data from RCTs only had minimal impact on the results.

Potential extensions to the NMA modelling could be explored including examining differential effects by child and family

factors, exploring in more depth the effect of study quality, and categorising educational interventions more finely. However, such analyses would be severely limited by the quantity of data currently available. The NMAs presented in this paper relied on only a small number of studies. To enhance the evidence base, further studies are required to increase precision of effect estimates, along with more details on the interventions trialled to reduce heterogeneity. Also, methods have been developed to incorporate individual-level data into NMA analyses,[48] which would greatly increase the power of analyses to explore the subject-level covariates identified above.

Knowing which interventions are the most effective is important, but cost-effectiveness is an essential part of any decision-making process. The effect sizes from this NMA will be used in subsequent decision analyses to determine the most cost-effective interventions for preventing falls in preschool children in the home. For example, our analysis found the most intensive interventions to be the most effective for increasing stair gate possession; however, as these interventions will also be the most costly, it is crucial to establish which interventions provide the best value for money.

What is already known on this subject

▶ Falls at home are a leading cause of injury in children under age 5 years, and a major burden on healthcare costs.
▶ Interventions to prevent falls have shown to be effective in promoting the use of safety equipment.
▶ Because interventions are often complex, direct comparisons are sparse and not easily combined in a standard (pairwise) meta-analysis.

What this study adds

▶ Using a network meta-analysis, a range of complex interventions to prevent falls in children under age 5 years have been compared, including those comparisons not directly evaluated in primary studies.
▶ The interventions have been ranked according to their effectiveness, thus providing directly useful information for decision makers to inform falls-prevention policy.

REFERENCES

1 Peden M, Oyegbite K, Ozanne-Smith J, *et al*. World report on child injury prevention. Geneva: WHO, 2008.
2 Innocenti Research Centre F. A league table of child deaths by injury in rich nations. 2001; http://www.unicef-icdc.org/publications/pdf/repcard2e.pdf (accessed Jul 2013), Innocenti Report Card No.2.
3 Agran PF, Anderson C, Winn D, *et al*. Rates of pediatric injuries by 3-month intervals for children 0 to 3 years of age. *Pediatrics* 2003;111:e683–92.
4 Pickett W, Streight S, Simpson K, *et al*. Injuries experienced by infant children: a population-based epidemiological analysis. *Pediatrics* 2003;111:e365–70.
5 Peden M, McGee K, Krug E. *Injury: a leading cause of the Global Burden of Disease*. Geneva, Switzerland: WHO, 2002.
6 Phelan KJ, Khoury J, Kalkwarf H, *et al*. Residential Injuries in U.S. children and adolescents. *Public Health Rep* 2005;120:63–70.
7 Khambalia A, Joshi P, Brussoni M, *et al*. Risk factors for unintentional injuries due to falls in children aged 0–6 years: a systematic review. *Inj Prev* 2006;12:378–81.
8 Bartlett SN. The problem of children's injuries in low-income countries: a review. *Health Policy Plan* 2002;17:1–13.
9 Miller TR, Romano EO, Spicer RS. The cost of childhood unintentional injuries and the value of prevention. *Future Child* 2000;10:137–63.
10 Young B, Wynn PM, He Z, *et al*. Preventing childhood falls within the home: overview of systematic reviews and a systematic review of primary studies. *Accid Anal Prev* 2013;60:158–71.
11 Jordan EA, Duggan AK, Hardy JB. Injuries in children of adolescent mothers: home safety education associated with decreased injury risk. *Pediatrics* 1993;(914): 481–7.
12 Kendrick D, Wynn P, Mulvaney C, *et al*. Home safety education and provision of safety equipment for injury prevention. *Cochrane Database Syst Rev* 2007;(1): CD005014.
13 Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897–900.
14 Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313–24.
15 Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.
16 Cooper NJ, Peters J, Lai MC, *et al*. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? *Value Health* 2011;14: 371–80.
17 Cooper NJ, Spiegelhalter D, Bujkiewicz S, *et al*. Use of implicit and explicit Bayesian methods in health technology assessment. *Int J Technol Assess Health Care* 2013;29:336–42.
18 Dias S, Welton NJ, Sutton AJ, *et al*. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; last updated August 2011. http://www.nicedsu.org.uk
19 Cooper NJ, Kendrick D, Achana F, *et al*. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms. *Epidemiol Rev* 2012;34:32–45.
20 Wells G, Shea B, O'Connell D, *et al*. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. (http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm) (accessed Jul 2013).
21 Kendrick D, Illingworth R, Woods A, *et al*. Promoting child safety in primary care: a cluster randomised controlled trial to reduce baby walker use. *Br J Gen Prac* 2005;55:582–8.
22 Lu G, Ades A. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.
23 Cooper N, Sutton A, Morris D, *et al*. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med* 2009;28:1861–81.
24 Kendrick D, Young B, Mason-Jones A, *et al*. Home safety education and provision of safety equipment for injury prevention. *Cochrane Database Syst Rev* 2012;(9): CD005014.
25 Spiegelhalter DJ, Best NG, Carlin BP, *et al*. Bayesian measures of model complexity and fit. *J Royal Stat Soc Ser B* 2002;64:583–639.
26 Spiegelhalter D, Thomas A, Best N, *et al*. WinBUGS user manual v1.4. MRC Biostatistics Unit. 2003; http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf (accessed Jul 2013).
27 Congdon P. *Applied Bayesian modelling*. Chichester, UK: Wiley, 2003.
28 McCullagh P, Nelder JA. *Generalized linear models*. London, UK: Chapman & Hall, 1989.
29 Cooper N, Sutton A, Lu G, *et al*. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006;166:1269–75.
30 Borenstein M, Hedges L, Higgins J, *et al*. *Introduction to meta-analysis*. Chichester, UK: Wiley, 2009.
31 Dias S, Welton NJ, Caldwell DM, *et al*. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932–44.

32 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

33 Nansel TR, Weaver N, Donlin M, et al. Baby, Be safe: the effect of tailored communications for pediatric injury prevention provided in a primary care setting. Patient Educ Couns 2002;46:175–90.

34 Nansel TR, Weaver NL, Jacobsen HA, et al. Preventing unintentional pediatric injuries: a tailored intervention for parents and providers. Health Educ Res 2008;23:656–69.

35 Tan NC, Lim NMLH, Gu K. Effectiveness of nurse counselling in discouraging the use of the infant walkers. Asia Pac J Public Health 2004;16:104–7.

36 Babul S, Olsen L, Janssen P, et al. A randomized trial to assess the effectiveness of an infant home safety programme. Int J Inj Control Saf Prom 2007;14:109–17.

37 Clamp M, Kendrick D. A randomised controlled trial of general practitioner safety advice for families with children under 5 years. BMJ 1998;316:1576–9.

38 McDonald EM, Solomon B, Shields W, et al. Evaluation of kiosk-based tailoring to promote household safety behaviors in an urban pediatric primary care practice. Patient Educ Couns 2005;58:168–81.

39 Kendrick D, Marsh P, Fielding K, et al. Preventing injuries in children: cluster randomised controlled trial in primary care. BMJ 1999;318:980–3.

40 Hendrickson SG. Reaching an underserved population with a randomly assigned home safety intervention. Inj prev 2005;11:313–17.

41 Watson M, Kendrick D, Coupland C, et al. Providing child safety equipment to prevent injuries: randomised controlled trial. BMJ 2005;330:178–81.

42 Petridou E, Tolma E, Dessypris N, et al. A controlled evaluation of a community injury prevention project in two Greek islands. Int J Epidemiol 1997;26:173–9.

43 Phelan KJ, Khoury J, Xu Y, et al. A randomised, controlled trial of home injury hazard reduction: the HOME injury study. Arch Pediatr Med 2011;165:339–45.

44 Posner JC, Hawkins LA, Garcia-Espana F, et al. A randomized, clinical trial of a home safety intervention based in an emergency department setting. Pediatrics 2004;113:1603–8.

45 Sznajder M, Janvrin MP, Albonico V, et al. Evaluation of the effectiveness of an injury prevention kit delivery for toddlers in four French cities. Inj Prev 2003;9:261–5.

46 Gielen AC, McDonald EM, Wilson ME, et al. Effects of improved access to safety counseling, products, and home visits on parents' safety practices: results of a randomized trial. Arch Pediatr Adolesc Med 2002;156:33–40.

47 King WJ, Klassen TP, LeBlanc J, et al. The effectiveness of a home visit to prevent childhood injury. Pediatrics 2001;108:382–8.

48 Saramago P, Sutton AJ, Cooper NJ, et al. Mixed treatment comparisons using aggregate and individual participant level data. Stat Med 2012;31:3516–36.

Fall prevention programmes save money

A paper in The Journal of Safety Research found three fall prevention programmes to be effective and to save medical costs. The programmes were: Tai Chi: Moving for Better Balance; Stepping On; and The Otago Exercise Program. Importantly, effectiveness was only based on randomised controlled trial results. The return on investment ranged from 36% for Otago to 509% for Tai Chi.

End to drug overdose epidemic predicted

Farr's law on epidemic patterns may apply to infectious diseases and drug overdoses. A study in Injury Epidemiology predicts that the tenfold increase in overdose deaths since 1980 will peak in 2017 and then decline to about 6000 deaths in 2035.

Prolonged rest for concussions

A new study in Paediatrics indicates that following a concussion prolonged rest until acute symptoms are gone may and actually worsened them. Patients were randomised to strict-rest and short-rest groups. Strict rest failed to improve symptoms and also worsened them.

264

Bibliography

ABO-ZAID, G., GUO, B., DEEKS, J.J., DEBRAY, T.P.A., STEYERBERG, E.W., MOONS, K.G.M. and RILEY, R.D., 2013. Individual participant data meta-analyses should not ignore clustering. *Journal of clinical epidemiology,* **66**(8), pp. 865-873.e4.

ACHANA, F.A., COOPER, N.J., DIAS, S., LU, G., RICE, S.J., KENDRICK, D. and SUTTON, A.J., 2012. Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. *PubMed, .*

ACHANA, F.A., COOPER, N.J., BUJKIEWICZ, S., HUBBARD, S.J., KENDRICK, D., JONES, D.R. and SUTTON, A.J., 2014. Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. *BMC medical research methodology,* **14**(1), pp. 92.

ACHANA, F., HUBBARD, S., SUTTON, A., KENDRICK, D. and COOPER, N., 2014. An exploration of synthesis methods in public health evaluations of interventions concludes that the use of modern statistical methods would be beneficial. *Journal of clinical epidemiology,* **67**(4), pp. 376-390.

ADES, A.E., CALDWELL, D.M., REKEN, S., WELTON, N.J., SUTTON, A.J. and DIAS, S., 2013. Evidence Synthesis for Decision Making 7: A Reviewer's Checklist. *Medical Decision Making,* **33**(5), pp. 679-691.

ARENDS, L.R., HOES, A.W., LUBSEN, J., GROBBEE, D.E. and STIJNEN, T., 2000. Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Statistics in medicine,* **19**(24), pp. 3497.

ARMSTRONG, R. and WATERS, E., 2007-last update, Guidelines: Systematic Reviews of Health Promotion and Public Health Interventions. Version 2. Available: http://ph.cochrane.org/sites/ph.cochrane.org/files/public/uploads/Guidelines%20HP_PH%20reviews.pdf [03/08, 2018].

ARMSTRONG, R., WATERS, E. and DOYLE, J., 2011. Chapter 21: Reviews in health promotion and public health. In: J. HIGGINS and S. GREEN, eds, *Cochrane Handbook for Systematic Reviews of Interventions.* Version 5.1.0 edn. www.handbook.cochrane.org: The Cochrane Collaboration, .

AUDIT COMMISSION, 2007. *Better safe than sorry. Preventing unintentional injury to children.* Audit Commission.

BABUL, S., OLSEN, L., JANSSEN, P., MCINTEE, P. and RAINA, P., 2007. A randomized trial to assess the effectiveness of an infant home safety programme. *International journal of injury control and safety promotion,* **14**(2), pp. 109-117.

BAFETA, A., TRINQUART, L., SEROR, R. and RAVAUD, P., 2014. Reporting of results from network meta-analyses: methodological systematic review. *BMJ (Clinical research ed.),* **348**(mar11 5), pp. g1741-g1741.

BERLIN, J.A., SANTANNA, J., SCHMID, C.H., SZCZECH, L.A., FELDMAN, H.I. and ANTI-LYMPHOCYTE ANTIBODY INDUCTION THERAPY STUDY,GROUP, 2002. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in medicine,* **21**(3), pp. 371-387.

BORENSTEIN, M., HEDGES, L.V., HIGGINS, J.P.T. and ROTHSTEIN, H.R., 2009. *Introduction to Meta-Analysis.* Chichester, UK: Wiley.

BRAZIER, J., 2007. *Measuring and Valuing Health Benefits for Economic Evaluation.* Oxford: OUP Oxford.

BRIGGS, A.H., CLAXTON, K. and SCULPHER, M.J., 2006. *Decision modelling for health economic evaluation.* Oxford: Oxford University Press.

BRIGGS, A.H., WEINSTEIN, M.C., FENWICK, E.A.L., KARNON, J., SCULPHER, M.J. and PALTIEL, A.D., 2012. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health,* **15**(6), pp. 835-842.

BROOKS, S.P. and GELMAN, A., 1998. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics,* **7**(4), pp. 434-455.

BRUSSONI, M., KRUSE, S. and WALKER, K., 2013. Validity and reliability of the EQ-5D-3L™ among a paediatric injury population. *Health and quality of life outcomes,* **11**(1), pp. 157.

BUJKIEWICZ, S., THOMPSON, J.R., SPATA, E., ABRAMS, K.R. and BALAKRISHNAN, N., 2017. Uncertainty in the Bayesian meta- analysis of normally distributed surrogate endpoints. *Statistical methods in medical research,* **26**(5), pp. 2287-2318.

CALDWELL, D.M., ADES, A.E. and HIGGINS, J.P.T., 2005. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ (Clinical research ed.),* **331**(7521), pp. 897-900.

CALDWELL, D.M. and WELTON, N.J., 2016. Approaches for synthesising complex mental health interventions in meta-analysis. *Evidence Based Mental Health,* **19**(1), pp. 16.

CHAIMANI, A., HIGGINS, J.P.T., MAVRIDIS, D., SPYRIDONOS, P. and SALANTI, G., 2013. Graphical Tools for Network Meta- Analysis in STATA.(Research Article). *PLoS ONE,* **8**(10), pp. e76654.

CHEN, G., FLYNN, T., STEVENS, K., BRAZIER, J., HUYNH, E., SAWYER, M., ROBERTS, R. and RATCLIFFE, J., 2015. Assessing the Health- Related Quality of Life of Australian Adolescents: An Empirical Comparison of the Child Health Utility 9D and EQ-5D-Y Instruments. *Value in Health,* **18**(4), pp. 432-438.

CLAMP, M. and KENDRICK, D., 1998. A Randomised Controlled Trial of General Practitioner Safety Advice for Families with Children under 5 Years. *BMJ: British Medical Journal,* **316**(7144), pp. 1576-1579.

CLAXTON, K., NEUMANN, P.J., ARAKI, S. and WEINSTEIN, M.C., 2001. BAYESIAN VALUE-OF-INFORMATION ANALYSIS An Application to a Policy Model of Alzheimers Disease. *International Journal of Technology Assessment in Health Care; J.of Inter.Tech.of Health Care,* **17**(1), pp. 38-55.

CONGDON, P., 2003. *Applied Bayesian Modelling.* Chichester,United Kingdom.: Wiley.

COOPER NJ, SUTTON AJ, LU G,KHUNTI K., 2006. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Archives of Internal Medicine,* **166**(12), pp. 1269-1275.

COOPER, N.J., KENDRICK, D., TIMBLIN, C., HAYES, M., MAJSAK-NEWMAN, G., METEYARD, K., HAWKINS, A. and KAY, B., 2016. The short- term cost of falls, poisonings and scalds occurring at home in children under 5 years old in England: multicentre longitudinal study. *Injury Prevention,* **22**(5), pp. 334.

COOPER, N.J., PETERS, J., LAI, M.C., JUNI, P., WANDEL, S., PALMER, S., PAULDEN, M., CONTI, S., WELTON, N.J., ABRAMS, K.R., BUJKIEWICZ, S., SPIEGELHALTER, D. and SUTTON, A.J., 2011. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation?. *Value in Health,* **14**(2), pp. 371-380.

COOPER, N.J., ABRAMS, K.R., SUTTON, A.J., TURNER, D. and LAMBERT, P.C., 2003. A Bayesian Approach to Markov Modelling in Cost-Effectiveness Analyses: Application to Taxane Use in Advanced Breast Cancer. *Journal of the Royal Statistical Society.Series A (Statistics in Society),* **166**(3), pp. 389-405.

COOPER, N.J., KENDRICK, D., ACHANA, F., DHIMAN, P., HE, Z., WYNN, P., LE COZANNET, E., SARAMAGO, P. and SUTTON, A.J., 2012. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms. *Epidemiologic reviews,* **34**(1), pp. 32-45.

COOPER, N.J., SPIEGELHALTER, D., BUJKIEWICZ, S., DEQUEN, P. and SUTTON, A.J., 2013. Use of Implicit and Explicit Bayesian Methods in Health Technology Assessment. *International Journal of Technology Assessment in Health Care,* **29**(03), pp. 336-342.

COOPER, N.J., SUTTON, A.J., ABRAMS, K.R., TURNER, D. and WAILOO, A., 2004. Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach. *Health Economics,* **13**(3), pp. 203-226.

COOPER, N.J., SUTTON, A.J., MORRIS, D., ADES, A.E. and WELTON, N.J., 2009. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Statistics in medicine,* **28**(14), pp. 1861-1881.

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZZARETH, I. and PETTICREW, M., 2006. *Developing and evaluating complex interventions: the new Medical Research Council guidance.*

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. and PETTICREW, M., 2013. Developing and evaluating complex interventions: the new Medical Research Council guidance. *International journal of nursing studies,* **50**(5), pp. 587.

CURTIS, L., 2017-last update, Unit costs of health & social care. Available: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/ [06/20, 2018].

CURTIS, L., 2012-last update, Unit costs of health & social care. Available: http://www.pssru.ac.uk/project-pages/unit-costs/2012 [07/04, 2016].

DANIELS, M.J. and HUGHES, M.D., 1997. Meta- analysis for the evaluation of potential surrogate markers. *Statistics in medicine,* **16**(17), pp. 1965-1982.

DAVIS, S., STEVENSON, M., TAPPENDEN, P. and WAILOO, A.J., 2014. *NICE DSU Technical Support Document 15: Cost-effectiveness modelling using patient-level simulation.*

DEEKS, J., ALTMAN, D., and BRADBURN, M., 2008. Statistical Methods for Examining Heterogeneity and Combining Results from Several Studies in Meta-Analysis. In: M. EGGER, G. DAVEY-SMITH and D. ALTMAN, eds, *Systematic Reviews in Health Care: Meta-analysis in Context* . 2nd edn. London, UK: London, UK: BMJ Publishing Group, pp. 285-312.

DEPARTMENT FOR TRADE AND INDUSTRY., 2003. *24th (Final) Report of the Home and Leisure Accident Surveillance System. 2000, 2001 and 2002 Data. London: Department for Trade and Industry.*

DERSIMONIAN, R. and KACKER, R., 2007. Random- effects model for meta- analysis of clinical trials: An update. *Contemporary Clinical Trials,* **28**(2), pp. 105-114.

DERSIMONIAN, R. and LAIRD, N., 2015. Meta- analysis in clinical trials revisited. *Contemporary Clinical Trials,* **45**, pp. 139-145.

DERSIMONIAN, R. and LAIRD, N., 1986. *Meta-analysis in clinical trials.*

DIAS, S., SUTTON, A.J., WELTON, N.J. and ADES, A.E., 2012.
NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment.

DIAS, S., SUTTON, A., ADES, A. and WELTON, N., 2013a. Evidence Synthesis for Decision Making 2: A Generalized Linear Modeling Framework for Pairwise and Network Meta-analysis of Randomized Controlled Trials. *Medical Decision Making,* **33**(5), pp. 607-617.

DIAS, S., SUTTON, A., WELTON, N. and ADES, A., 2013b. Evidence Synthesis for Decision Making 3: Heterogeneity—Subgroups, Meta-Regression, Bias, and Bias-Adjustment. *Medical Decision Making,* **33**(5), pp. 618-640.

DIAS, S., WELTON, N., SUTTON, A. and ADES, A., 2011. *NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials.*

DIAS, S., WELTON, N., SUTTON, A., CALDWELL, D., LU, G. and ADES, A., 2013a. Evidence Synthesis for Decision Making 4: Inconsistency in Networks of Evidence Based on Randomized Controlled Trials. *Medical Decision Making,* **33**(5), pp. 641-656.

DIAS, S., WELTON, N.J., CALDWELL, D.M. and ADES, A.E., 2010a. Checking consistency in mixed treatment comparison meta-analysis. *Statistics in medicine,* **29**(7-8), pp. 932-944.

DIAS, S., WELTON, N.J., MARINHO, V.C.C., SALANTI, G., HIGGINS, J.P.T. and ADES, A.E., 2010b. Estimation and adjustment of bias in randomized evidence by using mixed treatment comparison meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* **173**(3), pp. 613-629.

DIAS, S., SUTTON, A.J., WELTON, N.J. and ADES, A.E., 2013c. Evidence Synthesis for Decision Making 6: Embedding Evidence Synthesis in Probabilistic Cost-effectiveness Analysis. *Medical Decision Making,* **33**(5), pp. 671-678.

DIAS, S., WELTON, N.J., SUTTON, A.J. and ADES, A.E., 2013b. Evidence Synthesis for Decision Making 1: Introduction. *Medical Decision Making,* **33**(5), pp. 597-606.

DIGUISEPPI, C., SLATER, S., ROBERTS, I., ADAMS, L., SCULPHER, M., WADE, A. and MCCARTHY, M., 1999. The "Let's Get Alarmed!" initiative: a smoke alarm giveaway programme. *Injury prevention : journal of the International Society for Child and Adolescent Injury Prevention,* **5**(3), pp. 177-182.

DONNER, A. and KLAR, N., 2002. Issues in the meta-analysis of cluster randomized trials. *Statistics in medicine,* **21**(19), pp. 2971-2980.

DOWSWELL, T. and TOWNER, E., 2002. Social deprivation and the prevention of unintentional injury in childhood: a systematic review. *Health education research,* **17**(2), pp. 221-237.

DRUMMOND, M., WEATHERLY, H., CLAXTON, K., COOKSON, R., FERGUSON, B., GODFREY, C., RICE, N., SCULPHER, M. and SOWDEN, A., 2007. *Assessing the challenges of applying standard methods of economic evaluation to public health interventions.*

DRUMMOND, M., 2005. *Methods for the economic evaluation of health care programmes.* Oxford: Oxford University Press.

DRUMMOND, M., 2001. *Economic evaluation in health care : merging theory with practice.* Oxford: Oxford University Press.

EDLIN, R., MCCABE, C., HULME, C., HALL, P. and WRIGHT, J., 2015. *Cost effectiveness modelling for health technology assessment.* Springer.

EDWARDS, P., ROBERTS, I., GREEN, J. and LUTCHMUN, S., 2006. Deaths from injury in children and employment status in family: analysis of trends in class specific death rates. *BMJ (Clinical research ed.),* **333**(7559), pp. 119.

EGGER, M., DAVEY SMITH, G., SCHNEIDER, M. and MINDER, C., 1997. *Bias in meta-analysis detected by a simple, graphical test.* London : .

EUROPEAN CHILD SAFETY ALLIANCE AND EUROSAFE, 2012-last update, Child Safety Report Card: England. Available: www.childsafetyeurope.org/reportcards/info /england-report-card.pdf [January/19, 2018].

FENWICK, E., CLAXTON, K. and SCULPHER, M., 2001. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics,* **10**(8), pp. 779-787.

GELMAN, A. and RUBIN, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science,* **7**(4), pp. 457-472.

GIELEN, A.C., MCDONALD, E.M., WILSON, M.E., HWANG, W.T., SERWINT, J.R., ANDREWS, J.S. and WANG, M.C., 2002. Effects of improved access to safety counseling, products, and home visits on parents' safety practices: results of a randomized trial. *Archives of Pediatrics & Adolescent Medicine,* **156**(1), pp. 33-40.

GILKS, W.R., RICHARDSON, S. and SPIEGELHALTER, D.J., 1996. *Markov chain Monte Carlo in practice.* London: Chapman and Hall.

GINNELLY, L., SCULPHER, M., BOJKE, C., ROBERTS, I., WADE, A. and DIGUSEPPI, C., 2005. Determining the cost effectiveness of a smoke alarm give- away program using data from a randomized controlled trial. *The European Journal of Public Health,* **15**(5), pp. 448-453.

GRAY, A.M., 2011. *Applied methods of cost-effectiveness analysis in health care.* Oxford: Oxford University Press.

GREENLAND, S. and MORGENSTERN, H., 1989. Ecological Bias, Confounding, and Effect Modification. *International journal of epidemiology,* **18**(1), pp. 269-274.

HARBORD, R.M., EGGER, M. and STERNE, J.A.C., 2006. A modified test for small- study effects in meta- analyses of controlled trials with binary endpoints. *Statistics in medicine,* **25**(20), pp. 3443-3457.

HAYES, M., KENDRICK, D. and DEAVE, T., 2014. *Injury Prevention Briefing. Preventing unintentional injuries to the under fives: a guide for practitioners.*

HEALTH AND SOCIAL CARE INFORMATION CENTRE, 2013-last update, Hospital Episode Statistics, Admitted Patient Care, England – 2012–13: External Causes. Available: www .hscic.gov.uk/catalogue/PUB12566 [January, 2018].

HELEN SHAW, R., 2014. *Health Accident Surveillance Systems (HASS) 2002.*

HIGGINS, J. and GREEN, S., 2011. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* Available from www.cochrane-handbook.org: The Cochrane Collaboration.

HIGGINS, J.P. and WHITEHEAD, A., 1996. Borrowing strength from external trials in a meta-analysis. *Statistics in medicine,* **15**(24), pp. 2733-2749.

HIGGINS, J.P.T. and THOMPSON, S.G., 2004. Controlling the risk of spurious findings from meta- regression. *Statistics in medicine,* **23**(11), pp. 1663-1682.

HIGGINS, J.P.T. and THOMPSON, S.G., 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in medicine,* **21**(11), pp. 1539-1558.

HIGGINS, J.P.T., THOMPSON, S.G. and SPIEGELHALTER, D.J., 2009. A re- evaluation of random- effects meta- analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* **172**(1), pp. 137-159.

HM TREASURY, 2011. *Green Book: Appraisal and Evaluation in Central Government.* Stationery Office, London.

HUBBARD, S., COOPER, N., KENDRICK, D., YOUNG, B., WYNN, P.M., HE, Z., MILLER, P., ACHANA, F. and SUTTON, A., 2015. Network meta-analysis to evaluate the effectiveness of interventions to prevent falls in children under age 5 years. *Injury Prevention,* **21**, pp. 98-108.

HUSEREAU, D., DRUMMOND, M., PETROU, S., CARSWELL, C., MOHER, D., GREENBERG, D., AUGUSTOVSKI, F., BRIGGS, A.H., MAUSKOPF, J., LODER, E. and CHEERS TASK FORCE, 2013. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ (Clinical research ed.),* **346**, pp. f1049.

IBRAHIM, J.G. and CHEN, M., 2000. Power Prior Distributions for Regression Models. *Statistical Science,* **15**(1), pp. 46-60.

JACKSON, D., BOWDEN, J. and BAKER, R., 2010a. How does the DerSimonian and Laird procedure for random effects meta- analysis compare with its more efficient but harder to compute counterparts? *Journal of Statistical Planning and Inference,* **140**(4), pp. 961-970.

JACKSON, D., BOWDEN, J. and BAKER, R., 2010b. How does the DerSimonian and Laird procedure for random effects meta- analysis compare with its more efficient but harder to compute counterparts? *Journal of Statistical Planning and Inference,* **140**(4), pp. 961-970.

JAMES, A., YAVCHITZ, A., RAVAUD, P. and BOUTRON, I., 2018. Node-making process in network meta-analysis of nonpharmacological treatment are poorly reported. *Journal of clinical epidemiology; Journal of clinical epidemiology,* **97**, pp. 95-102.

KENDRICK, D., ABLEWHITE, J., ACHANA, F., BENFORD, P., CLACY, R., COFFEY, F., COOPER, N., COUPLAND, C., DEAVE, T., GOODENOUGH, T., HAWKINS, A., HAYES, M., HINDMARCH, P., HUBBARD, S., KAY, B., KUMAR, A., MAJSAK-NEWMAN, G., MCCOLL, E., MCDAID, L., MILLER, P., MULVANEY, C., PEEL, I., PITCHFORTH, E., READING, R., SARAMAGO, P., STEWART, J., SUTTON, A., TIMBLIN, C., TOWNER, E., WATSON, M., WYNN, P., YOUNG, B. and ZOU, K., 2017. Keeping Children Safe: a multicentre programme of research to increase the evidence base for preventing unintentional injuries in the home in the under-fives. *Programme Grants for Applied Research, Programme Grants Appl Res,* **5**(14),.

KENDRICK, D., COUPLAND, C., MULVANEY, C., SIMPSON, J., SMITH, S.J., SUTTON, A., WATSON, M. and WOODS, A., 2007a. Home safety education and provision of safety equipment for injury prevention. *The Cochrane database of systematic reviews,* (1), pp. CD005014.

KENDRICK, D., ZOU, K., ABLEWHITE, J., WATSON, M., COUPLAND, C., KAY, B., HAWKINS, A. and READING, R., 2015. Risk and protective factors for falls on stairs in young children: multicentre case– control study. *Archives of Disease in Childhood,* .

KENDRICK, D., COUPLAND, C., MASON-JONES AMANDA, J., MULVANEY, C., SIMPSON, J., SMITH, S., SUTTON, A. and WATSON, M., 2007b. *Home safety education and provision of safety equipment for injury prevention.* John Wiley & Sons, Ltd.

KENDRICK, D., ILLINGWORTH, R., WOODS, A., WATTS, K., COLLIER, J., DEWEY, M., HAPGOOD, R. and CHEN, C., 2005. Promoting child safety in primary care: a cluster randomised controlled trial to reduce baby walker use. *The British journal of general practice : the journal of the Royal College of General Practitioners,* **55**(517), pp. 582-582.

KENDRICK, D., MARSH, P., FIELDING, K. and MILLER, P., 1999. Preventing Injuries in Children: Cluster Randomised Controlled Trial in Primary Care. *BMJ: British Medical Journal,* **318**(7189), pp. 980-983.

KENDRICK, D., WYNN, P., MULVANEY, C., WATSON, M.C., COUPLAND, C., YOUNG, B., MASON-JONES, A.J., ILYAS, N., ACHANA, F.A., COOPER, N.J., HUBBARD, S.J., SUTTON, A.J. and SMITH, S., 2013. Home safety education and provision of safety equipment for injury prevention (Review). *Evidence-based child health : a Cochrane review journal,* **8**(3), pp. 761.

KENDRICK, D., YOUNG, B., MASON-JONES, A.J., ILYAS, N., ACHANA, F.A., COOPER, N.J., HUBBARD, S.J., SUTTON, A.J., SMITH, S., WYNN, P., MULVANEY, C., WATSON, M.C. and COUPLAND, C., 2012. Home safety education and provision of safety equipment for injury prevention. *The Cochrane database of systematic reviews,* **9**, pp. CD005014.

KIBRET, T., RICHER, D. and BEYENE, J., 2014. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clinical Epidemiology,* **2014**, pp. 451-460.

KIDSCREEN, . Available: www.kidscreen.de/cms/kidscreen10 [06/30, 2016].

KIND, P., HARDMAN, G. and MACRAN, S., 1999. *UK Population Norms for EQ-5D.*

KIND, P., DOLAN, P., GUDEX, C. and WILLIAMS, A., 1998. Variations in population health status: results from a United Kingdom national questionnaire survey. *British medical journal,* **316**(7133), pp. 736.

KING, W.J., KLASSEN, T.P., LEBLANC, J., BERNARD-BONNIN, A.C., ROBITAILLE, Y., PHAM, B., COYLE, D., TENENBEIN, M. and PLESS, I.B., 2001. The effectiveness of a home visit to prevent childhood injury. *Pediatrics,* **108**(2), pp. 382-388.

LAMBERT, P.C., SUTTON, A.J., ABRAMS, K.R. and JONES, D.R., 2002. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of clinical epidemiology,* **55**(1), pp. 86-94.

LAMBERT, P.C., SUTTON, A.J., BURTON, P.R., ABRAMS, K.R. and JONES, D.R., 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in medicine,* **24**(15), pp. 2401-2428.

LU, G. and ADES, A.E., 2004. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in medicine,* **23**(20), pp. 3105-3124.

LU, G. and ADES, A.E., 2006. Assessing Evidence Inconsistency in Mixed Treatment Comparisons. *Journal of the American Statistical Association,* **101**(474), pp. 447-459.

LUMLEY, T., 2002. Network meta-analysis for indirect treatment comparisons. *Statistics in medicine,* **21**(16), pp. 2313-2324.

LUNN, D., 2013. *The BUGS book : a practical introduction to Bayesian analysis.* Boca Raton, Fla. ; London: CRC Press.

LUNN, D., THOMAS, A., BEST, N. and SPIEGELHALTER, D., 2000. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing,* **10**(4), pp. 325-337.

MCCULLAGH, P. and NELDER, J.A., 1989. *Generalized Linear Models.* second edn. London, UK.: Chapman & Hall.

MCDONALD, E.M., SOLOMON, B., SHIELDS, W., SERWINT, J.R., JACOBSEN, H., WEAVER, N.L., KREUTER, M. and GIELEN, A.C., 2005. Evaluation of kiosk-based tailoring to promote household safety behaviors in an urban pediatric primary care practice. *Patient education and counseling,* **58**(2), pp. 168-181.

MILLER, T.R., ROMANO, E.O. and SPICER, R.S., 2000. The Cost of Childhood Unintentional Injuries and the Value of Prevention. *The Future of Children,* **10**(1), pp. 137-163.

MOHER, D., LIBERATI, A., TETZLAFF, J. and ALTMAN, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ,* **339**.

MORRISON, A., STONE, D.H., DORAISWAMY, N. and RAMSAY, L., 1999. Injury surveillance in an accident and emergency department: a year in the life of CHIRPP. *Archives of Disease in Childhood,* **80**(6), pp. 533-536.

MORRONGIELLO, B.A., HOWARD, A.W., ROTHMAN, L. and SANDOMIERSKI, M., 2009. Once bitten, twice shy? Medically- attended injuries can sensitise parents to children's risk of injuries on playgrounds. *Injury Prevention,* **15**(1), pp. 50.

MORRONGIELLO, B.A. and MATHEIS, S., 2007. Addressing the issue of falls off playground equipment: an empirically-based intervention to reduce fall-risk behaviors on playgrounds. *Journal of pediatric psychology,* **32**(7), pp. 819-830.

MRC BIOSTATISTICS UNIT, 2015b-last update, DIC: Deviance Information Criteria. Available: http://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/ [May/02, 2018].

NANSEL, T.R., WEAVER, N.L., JACOBSEN, H.A., GLASHEEN, C. and KREUTER, M.W., 2008. Preventing unintentional pediatric injuries: a tailored intervention for parents and providers. *Health education research,* **23**(4), pp. 656-669.

NANSEL, T.R., WEAVER, N., DONLIN, M., JACOBSEN, H., KREUTER, M.W. and SIMONS-MORTON, B., 2002. Baby, Be Safe: the effect of tailored communications for pediatric injury prevention provided in a primary care setting. *Patient education and counseling,* **46**(3), pp. 175-190.

NHS CHOICES, 2016-last update, Baby and toddler safety. Available: http://www.nhs.uk/conditions/pregnancy-and-baby/pages/baby-safety-tips.aspx#babies.

NICE, April 2013. *National Institute for Health and Care Excellence: Guide to the methods of technology appraisal 2013.*

NICE, 2014. *Developing NICE guidelines: the manual.* London: NICE.

NICE, 2013-last update, How NICE measures value for money in relation to public health interventions. Available: nice.org.uk/guidance/lgb10 [05/23, 2018].

NICE, 2012. *National Institute for Health and Care Excellence: Methods for the Development of NICE Public Health Guidance.* 3rd Edition.

NICE, 2010a. *Preventing unintentional injuries among under-15s:* **Costing template**. *NICE Public health guidance 29, 30 and 31.*

NICE, 2010b. *Unintentional injuries in the home: interventions for under 15s.*

NICHOLL, J., TURNER, J., YOUNG, T., FREEMAN, J., CHANAKIRA, E., DIXON, S., SANTARELLI, M., OHN, T., CROSS, S. and MASON, S., 2009. *The Long Term Health and Healthcare Outcomes of Accidental Injury (The HALO Study).* 0010017. University of Sheffield: .

OFFICE FOR NATIONAL STATISTICS, 2013a-last update, 2011 Census: Population estimates. Available: https://www.ons.gov.uk/peoplepopulationandcommunity [July, 2014].

OFFICE FOR NATIONAL STATISTICS, 2013b-last update, Mortality Statistics: deaths registered in England, 2008, 2009, 2010, 2011, 2012. Crown copyright reserved:. Available: www.ons.gov.uk/ons/rel/vsob1/mortality-statistics--deaths-registered-in-england-and-wales--series-dr-/index.html [July, 2014].

OFFICE FOR NATIONAL STATISTICS, , Census 2001, First Results on Population in England and Wales. Available: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-225134 [July 17th, 2014].

PEDEN, M., MCGEE, K., KARA, S., KRUG, E. and WHO, 2002. *Injury: A leading cause of the global burden of disease: 2000. Geneva : World Health Organization.*

PEDEN, M., OYEGBITE, K., OZANNE-SMITH, J., HYDER, A., BRANCHE, C. and RAHMAN, F., 2008. *World report on child injury prevention.*

PETERS, J.L., SUTTON, A.J., JONES, D.R., ABRAMS, K.R. and RUSHTON, L., 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of clinical epidemiology,* **61**(10), pp. 991-996.

PETRIDOU, E., TOLMA, E., DESSYPRIS, N. and TRICHOPOULOS, D., 1997. A controlled evaluation of a community injury prevention project in two Greek islands. *International journal of epidemiology,* **26**(1), pp. 173-179.

PETROU, S. and GRAY, A., 2011. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ: British Medical Journal,* **342**(7808), pp. 1195-1198.

PHELAN, K.J., KHOURY, J., XU, Y., LIDDY, S., HORNUNG, R. and LANPHEAR, B.P., 2010. A randomised, controlled trial of home injury hazard reduction: The HOME injury study. *Injury Prevention,* **16**(Supplement 1), pp. A171-A171.

PHILIPS, Z., BOJKE, L., SCULPHER, M., CLAXTON, K. and GOLDER, S., 2006. *Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment: A Review and Consolidation of Quality Assessment.* Cham: Adis International.

PHILLIPS, C., 2005. *Health economics : an introduction for health professionals.* Malden, Mass. ; Oxford: Blackwell.

PITT, M., ANDERSON, R. and MOXHAM, T., 2009-last update, Preventing unintentional injuries among under-15s in the home. Report 3: Cost-effectiveness modelling of home based interventions aimed at reducing unintentional injuries in children. Available: https://www.nice.org.uk/guidance/ph30/evidence/costeffectiveness-modelling-report-428721949; [11/15, 2013].

POSNER, J.C., HAWKINS, L.A., GARCIA-ESPANA, F. and DURBIN, D.R., 2004. A randomized, clinical trial of a home safety intervention based in an emergency department setting. *Pediatrics,* **113**(6), pp. 1603-1608.

PUBLIC HEALTH ENGLAND, February 2017-last update, Preventing Unintentional Injuries: A guide to staff working with children under five years, Crown Copyright 2017. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/595017/Preventing_unintentional_injuries_guide.pdf2018].

PUBLIC HEALTH ENGLAND, 2017-last update, Public Health Outcomes Framework. Available: www.phoutcomes.info/ [January/19, 2018].

R CORE TEAM (2013), *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL* http://www.R-project.org/.

R DEVELOPMENT CORE TEAM, 2012. *R: A language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

RAVENS-SIEBERER, U., WILLE, N., BADIA, X., BONSEL, G., BURSTRÖM, K., CAVRINI, G., DEVLIN, N., EGMAR, A., GUSI, N., HERDMAN, M., JELSMA, J., KIND, P., OLIVARES, P., SCALONE, L. and GREINER, W., 2010. Feasibility, reliability, and validity of the EQ- 5D-Y: results from a multinational study. *Quality of Life Research; An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation - Official Journal of the International Society of Quality of Life Research,* **19**(6), pp. 887-897.

REHFUESS, E. and AKL, E., 2013. Current experience with applying the GRADE approach to public health interventions: an empirical study. *BMC Public Health,* **13**(1), pp. 9.

RILEY, R.D., LAMBERT, P.C. and ABO-ZAID, G., 2010. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ (Clinical research ed.),* **340**, pp. c221.

RILEY, R.D., SIMMONDS, M.C. and LOOK, M.P., 2007. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current

practice and possible methods. *Journal of clinical epidemiology,* **60**(5), pp. 431.e1-431.e12.

RILEY, R.D. and STEYERBERG, E.W., 2010. Meta-analysis of a binary outcome using individual participant data and aggregate data. *Research Synthesis Methods,* **1**(1), pp. 2-19.

ROGOZIŃSKA, E., MARLIN, N., THANGARATINAM, S., KHAN, K.S. and ZAMORA, J., 2017. Meta- analysis using individual participant data from randomised trials: opportunities and limitations created by access to raw data. *Evidence Based Medicine,* **22**(5), pp. 157.

RÜCKER, G. and SCHWARZER, G., 2016. *Automated drawing of network plots in network meta- analysis.*

RUCKER, G., SCHWARZER, G. and CARPENTER, J., 2008a. Arcsine test for publication bias in meta-analyses with binary outcomes RID B-2965-2009 RID A-8168-2009. *Statistics in medicine,* **27**(5), pp. 746-763.

RUCKER, G., SCHWARZER, G., CARPENTER, J.R. and SCHUMACHER, M., 2008b. Undue reliance on I .sup.2 .sup.in assessing heterogeneity may mislead.(Research article)(Report). *BMC Medical Research Methodology,* **8**(79), pp. 79.

RYCHETNIK, L., FROMMER, M., HAWE, P. and SHIELL, A., 2002. Criteria for evaluating evidence on public health interventions. *Journal of epidemiology and community health,* **56**(2), pp. 119-127.

SALANTI, G., ADES, A.E. and IOANNIDIS, J.P.A., 2011. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of clinical epidemiology,* **64**(2), pp. 163-171.

SARAMAGO, P., COOPER, N.J., SUTTON, A.J., HAYES, M., DUNN, K., MANCA, A., KENDRICK, D. and KEEPING CHILDREN SAFE AT HOME STUDY, 2014. Cost-effectiveness of interventions for increasing the possession of functioning smoke alarms in households with pre-school children: a modelling study. *BMC public health,* **14**(1), pp. 459-459.

SARAMAGO, P., SUTTON, A.J., COOPER, N.J. and MANCA, A., 2012. Mixed treatment comparisons using aggregate and individual participant level data. *Statistics in medicine,* **31**(28), pp. 3516-3536.

SCULPHER, M. and CLAXTON, K., 2005. Establishing the Cost-Effectiveness of New Pharmaceuticals under Conditions of Uncertainty—When Is There Sufficient Evidence? *Value in Health,* **8**(4), pp. 433-446.

SELLAR, C., FERGUSON, J.A. and GOLDACRE, M.J., 1991. Occurrence and repetition of hospital admissions for accidents in preschool children. *BMJ (Clinical research ed.),* **302**(6767), pp. 16-19.

SHIELL, A., HAWE, P. and GOLD, L., 2008. Complex Interventions or Complex Systems? Implications for Health Economic Evaluation. *BMJ: British Medical Journal,* **336**(7656), pp. 1281-1283.

SIMMONDS, M.C. and HIGGINS, J.P.T., 2016. A general framework for the use of logistic regression models in meta-analysis. *Statistical methods in medical research,* **25**(6), pp. 2858-2877.

SIMMONDS, M.C., HIGGINS, J.P.T., STEWART, L.A., TIERNEY, J.F., CLARKE, M.J. and THOMPSON, S.G., 2005. Meta- analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical trials (London, England),* **2**(3), pp. 209.

SMARTRISK, 2009-last update,
The Economic Burden of Injury in Canada. SMARTRISK: Toronto, ON. Available: http://www.parachutecanada.org/downloads/research/reports/EBI2009-Eng-Final.pdf [04/02, 2014].

SMITH, T.C., SPIEGELHALTER, D.J. and THOMAS, A., 1995. Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in medicine,* **14**(24), pp. 2685-2699.

SONG, F., ALTMAN, D.G., GLENNY, A. and DEEKS, J.J., 2003. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. (Papers). *British medical journal,* **326**(7387), pp. 472.

SONG, F., XIONG, T., PAREKH-BHURKE, S., LOKE, Y.K., SUTTON, A.J., EASTWOOD, A.J., HOLLAND, R., CHEN, Y., GLENNY, A., DEEKS, J.J. and ALTMAN, D.G., 2011. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ,* **343**.

SPADY, D.W., SAUNDERS, D.L., SCHOPFLOCHER, D.P. and SVENSON, L.W., 2004. Patterns of injury in children: a population-based approach. *Pediatrics,* **113**(3 Pt 1), pp. 522-529.

SPIEGELHALTER, D., THOMAS, A., BEST, N. and LUNN, D., 2003-last update, WinBUGS user manual v1.4. MRC Biostatistics Unit. . Available: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf.

SPIEGELHALTER, D.J., ABRAMS, K.R. and MYLES, J.P., 2004. *Bayesian approaches to clinical trials and health-care evaluation.* Chichester: John Wiley & Sons.

SPIEGELHALTER, D.J., BEST, N.G., CARLIN, B.P. and VAN DER LINDE, A., 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **64**(4), pp. 583-639.

STATACORP. 2017, *Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC.*

STERNE, J.A.C., SUTTON, A.J., IOANNIDIS, J.P.A., TERRIN, N., JONES, D.R., LAU, J., CARPENTER, J., RÜCKER, G., HARBORD, R.M., SCHMID, C.H., TETZLAFF, J., DEEKS, J.J., PETERS, J., MACASKILL, P., SCHWARZER, G., DUVAL, S., ALTMAN, D.G., MOHER, D. and HIGGINS, J.P.T., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ,* **343**.

STEVENS, K. and RATCLIFFE, J., 2012. Measuring and Valuing Health Benefits for Economic Evaluation in Adolescence: An Assessment of the Practicality and Validity of the Child Health Utility 9D in the Australian Adolescent Population. *Value in Health,* **15**(8), pp. 1092-1099.

STEWART, L.A., 1995. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Statistics in medicine,* **14**(19), pp. 2057-2079.

STEWART, L.A., CLARKE, M., ROVERS, M., RILEY, R.D., SIMMONDS, M., STEWART, G. and TIERNEY, J.F., 2015. Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data: The PRISMA-IPD Statement. *JAMA,* **313**(16), pp. 1657-1665.

STINNETT, A.A., MULLAHY, J. and NATIONAL BUREAU OF ECONOMIC RESEARCH, 2013. Net Health Benefits : A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis. *Net Health Benefits : A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis,* .

SUTTON, A.J., KENDRICK, D. and COUPLAND, C.A., 2008. Meta-analysis of individual- and aggregate-level data. *Statistics in medicine,* **27**(5), pp. 651-669.

SUTTON, A.J. and ABRAMS, K.R., 2001. Bayesian methods in meta-analysis and evidence synthesis. *Statistical methods in medical research,* **10**(4), pp. 277-303.

SWEETING, M.J., SUTTON, A.J. and LAMBERT, P.C., 2007. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data.

SZNAJDER, M., JANVRIN, M.P., ALBONICO, V., BONNIN, M.H., BAUDIER, F. and CHEVALLIER, B., 2003. Evaluation of the effectiveness of an injury prevention kit delivery for toddlers in four French cities. *Archives de pédiatrie : organe officiel de la Sociéte française de pédiatrie,* **10**(6), pp. 510.

TAN, S.H., COOPER, N.J., BUJKIEWICZ, S., WELTON, N.J., CALDWELL, D.M. and SUTTON, A.J., 2015. Novel presentational approaches were developed for reporting network meta-analysis.

THE COCHRANE COLLABORATION, 2018-last update, Cochrane Public Health. Available: http://ph.cochrane.org/ [03/08, 2018].

THOMPSON, S.G. and HIGGINS, J.P.T., 2002. How should meta- regression analyses be undertaken and interpreted? *Statistics in medicine,* **21**(11), pp. 1559-1573.

THOMPSON, S.G., SMITH, T.C. and SHARP, S.J., 1997. Investigating underlying risk as a source of heterogeneity in meta-analysis. *Statistics in medicine,* **16**(23), pp. 2741-2758.

TOWNER, E., DOWSWELL, T., ERRINGTON, G., BURKES, M. and TOWNER, J., 2005. *Injuries in children aged 0-14 years and inequalities.*

TURNER, R.M., OMAR, R.Z., YANG, M., GOLDSTEIN, H. and THOMPSON, S.G., 2000. A multilevel model framework for meta- analysis of clinical trials with binary outcomes. *Statistics in medicine,* **19**(24), pp. 3417-3432.

TURNER, R.M., SPIEGELHALTER, D.J., SMITH, G.C.S. and THOMPSON, S.G., 2009. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* **172**(1), pp. 21-47.

UNICEF, 2001. *A league table of child deaths by injury in rich nations.* 2012.

VALKENHOEF, G., LU, G., BROCK, B., HILLEGE, H., ADES, A.E. and WELTON, N.J., 2012. Automating network meta- analysis. *Research Synthesis Methods,* **3**(4),.

VAN DER SLUIS, C.,K., STEWART, R.E., GROOTHOFF, J.W., TEN DUIS, H.J. and EISMA, W.H., 2005. A prospective study on paediatric traffic injuries: health- related quality of life and post- traumatic stress. *Clinical rehabilitation,* **19**(3), pp. 312-322.

WALKER, L., 2010. *Re-evaluation of home accidents.* PPR483.

WANLESS, D., 2004-last update, Securing good health for the whole population: Final report. Available: http://webarchive.nationalarchives.gov.uk/20111010152710/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4074426 [03/08, 2018].

WATSON, M., KENDRICK, D., COUPLAND, C., WOODS, A., FUTERS, D. and ROBINSON, J., 2005. Providing child safety equipment to prevent injuries: randomised controlled trial. *BMJ (Clinical research ed.),* **330**(7484), pp. 178-181.

WEATHERLY, H., DRUMMOND, M., CLAXTON, K., COOKSON, R., FERGUSON, B., GODFREY, C., RICE, N., SCULPHER, M. and SOWDEN, A., 2009. Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health policy,* **93**(2), pp. 85-92.

WELTON, N.J., ADES, A.E., CARLIN, J.B., ALTMAN, D.G. and STERNE, J.A.C., 2009. Models for potentially biased evidence in meta-analysis using empirically based priors. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* **172**(1), pp. 119-136.

WELTON, N.J., 2012. *Evidence synthesis for decision making in healthcare.* Chichester: John Wiley & Sons.

WELTON, N.J., CALDWELL, D.M., ADAMOPOULOS, E. and VEDHARA, K., 2009. Mixed Treatment Comparison Meta-Analysis of Complex Interventions: Psychological Interventions in Coronary Heart Disease. *American Journal of Epidemiology,* **169**(9), pp. 1158-1165.

WHITE, I.R., 2009. Multivariate random-effects meta-analysis. *Stata Journal,* **9**(1), pp. 40-56.

WHITEHEAD, S.J. and ALI, S., 2010. Health outcomes in economic evaluation: the QALY and utilities. *British medical bulletin,* **96**(1), pp. 5-21.

YOUNG, B., WYNN, P.M., HE, Z. and KENDRICK, D., 2013. Preventing childhood falls within the home: Overview of systematic reviews and a systematic review of primary studies. *Accident Analysis and Prevention,* **60**, pp. 158-171.