The effect of social networks on the participation by those with parental responsibility in the baby immunisation programme in the UK

Diane Mary Pople

A thesis submitted for the Degree of
Doctor of Philosophy (PhD)

Department of Infectious Disease Epidemiology
School of Public Health

Imperial College London
2017

Abstract

The majority of UK parents participate in the recommended baby vaccination programme, but some vaccines (notably MMR) have uptake below levels recommended to control outbreaks of vaccine-preventable disease, and this risk increases if the unvaccinated children are clustered. We explored the hypothesis that vaccination decisions made by parents based on information shared peer-to - peer could create clusters of opinions, and contribute to local vaccine uptake variations.

Ecological analysis of MMR uptake on a small spatial scale confirmed uneven coverage and a while a regression model showed uptake was associated with ethnicity and extremes of education, overall the observations were poorly explained by demographic factors. Mathematical modelling of decisions influenced by sharing information confirmed this process is theoretically able to create opinion clusters and changes in the proportions intending to vaccinate, but that results are qualitatively and quantitatively sensitive to network structure and decision representation. This uncertainty could not be resolved for UK baby vaccinations with existing data, so a survey was undertaken to address the knowledge gaps. Data were gathered on parents' networks of vaccine-information providers and on other variables within the MMR-measles decision-infection system, including social contacts for preschool children (with a larger sample than provided by all-age studies). The survey provided evidence of individual-level vaccination-behaviour clustering and informed revised mathematical models using empirically-supported network structures and decision representation. These simulations showed the UK conditions could enable information-sharing to create increased opinion clustering and to shift population-level vaccination sentiment (increasing those supporting schedule adherence).

Through an integrated programme of statistical analysis, data collection and mathematical modelling this thesis provides evidence to confirm the presence of clusters of vaccine opinion and to support the hypothesis that an information-sharing process is able to increase opinion clustering, albeit in a manner requiring further investigation to ascertain the associated relative outbreak risk.

Declaration of originality

I declare that this work is my own, completed under the Supervision of Professor Christophe Fraser and Professor Neil Ferguson. I conducted the analyses and simulations, designed and implemented the data collection, and drafted this thesis. Contributions of others have been appropriately referenced and acknowledged.

Copyright declaration

Acknowledgements

I would like to acknowledge and thank the following:

- my supervisors, Christophe Fraser and Neil Ferguson, for the opportunity to undertake this research, for their advice and guidance, both technical and non-technical, and support enabling the completion of this thesis.
- my assessors, Nick Grassly, Steven Riley and John Edmunds, for feedback which enhanced the final project.
- vaccination specialists at the Department of Health (David Salisbury, Joanne Yarwood and colleagues) and National Health Service (Emma Raworth, Great Yarmouth & Waveney PCT) for advice on the implementation of the UK vaccination programme
- my logistical support, Wes Hinsley and the DIDE IT team (particularly for salvage work when blue screen and ransomware attacks jeopardised the project) and the DIDE admin team.
- my colleagues, past and present, who helped with day-to-day queries (notably Anne Cori, Jeff Eaton, Jocelyn Elmes and room-mates Nick Beckley, Natsuko Imai, Pete Winskill, Hannah Clapham, Emily Pothin, Marcus Shepherd and Lucy Li) and provided continued encouragement (Branwen Owen, Amy Pinsent, Francesca Gauntlett and Miriam Tecle).

For specific elements of the work I acknowledge the following:

- Wes Hinsley (Imperial College) wrote the code that ran the online survey in chapter 4.
- Emma Raworth (then Great Yarmouth & Waveney PCT) provided additional vaccination data used in Chapter 2.
- Louisa McGrath (Imperial College) parallel-coded a sample of survey responses.

Special thanks go to all the staff and parents at the childcare settings involved in the survey, and survey pilot. I appreciate your time and co-operation.

Finally, I acknowledge and thank the Economic and Social Research Council for the funding that allowed me to embark on the whole process.

Table of contents

1. General Introduction 16
1.1. Childhood vaccination in the United Kingdom (UK) 16
1.1.1. Routine vaccinations 16
1.1.2. Surveillance instruments 18
1.1.3. Overview of vaccination behaviour and infection incidence 21
1.1.4. Understanding the sub optimal uptake of MMR(1) 29
1.2. Mathematical modelling of the vaccine decision process and resultant infection dynamics 34
1.2.1. Infection transmission models 35
1.2.2. Network models 35
1.2.3. Decision models 35
1.2.4. Mathematical models of information-behaviour-infection systems 39
2. Small Area Statistical Analysis 43
2.1. Motivation 43
2.2. Methods 45
2.2.1. Vaccine uptake data 45
2.2.2. Geographical units and ethical considerations 46
2.2.3. Variables considered for inclusion 47
2.2.4. Data sources 48
2.2.5. Data preparation for comparison at ward level 52
2.2.6. Data analysis 52
2.3. Results 55
2.3.1. Variation in MMR1 uptake 55
2.3.2. Exploratory analysis of ward-characteristics 57
2.3.3. Regression 69
2.4. Discussion 72
3. Initial Modelling 77
3.1. Motivation 77
3.2. Methods 79
3.2.1. Building the mathematical model 79
3.2.2. Applying the decision algorithms in a randomly-mixed population without underlying contact network structure 83

3.3. Results 84
3.3.1. Numbers of vaccine-supporters 84
3.3.2. Clustering of vaccination opinion 88
3.4. Discussion 92
4. Survey 95
4.1. Motivation 95
4.2. Initial design and pilot 100
4.2.1. Methods 101
4.2.2. Results 105
4.2.3. Discussion 107
4.3. Full survey 108
4.3.1. Methods 108
4.3.1.1. Variables measured 108
4.3.1.2. Survey format 108
4.3.1.3. Participants (recruitment) 109
4.3.1.4. Power calculation 112
4.3.1.5. Instrument development 113
4.3.1.6. Ethical considerations 115
4.3.1.7. Data analysis 115
4.3.2. Results 118
4.3.2.1. Implementation 118
4.3.2.2. Inferred opinions and vaccination status 124
4.3.2.3. Full networks structure 126
4.3.2.4. ‘Information’ networks 128
4.3.2.5. ‘Potential infection’ networks 132
4.3.2.6. Network overlap 136
4.3.2.7. Vaccination information received 136
4.3.2.8. Vaccination status 139
4.3.2.9. Intradyad agreement 141
4.3.3 Discussion 145
5. Revisiting the MMR1 Decision Model 151
5.1. Introduction 151
5.2. Parameter fitting 153
5.2.1. Methods 153

5.2.1.1. Network structure parameters 153
5.2.1.2 Decision model: Information network data 154
5.2.1.3 Infection model: Potential infection network data 156
5.2.1.4 Decision mathematical representation 156
5.2.2. Results 158
5.2.2.1 Decision model network structure 158
5.2.2.2 Infection model network structure 162
5.2.2.3. Decision model response function 165
5.3. Model building 168
5.3.1. Methods 168
5.3.2. Results 172
5.4. Discussion 178
6. Concluding Remarks 182
7. References 186
8. Appendices 211
8.1. Relating to Chapter 1: General Introduction 211
8.1.1. HPA Q11-1 request parameters 212
8.2. Relating to Chapter 3: Initial Modelling 213
8.2.1. Results for N=400 213
8.2.2. Results for count decision-algorithm, sensitivity to β 217
8.2.3. Results for threshold decision-algorithm, sensitivity to α 221
8.3. Relating to Chapter 4: Survey 225
8.3.1. Survey ethical procedures and safeguards 225
8.3.2. Survey materials 227
8.3.3. EAL in pre-school households in the shortlisted PCTs 234
8.3.4. Reported MMR uptake for survey sample 237
8.4. Relating to Chapter 5: Revisiting the MMR1 Decision Model 238
8.4.1. Initial visual inspection of decision-process data 238
8.5. Permissions 239

List of figures

Figure 1-1	UK routine vaccination schedule for under 5s	17
Figure 1-2	Timeline of events 1996-2016	21
Figure 1-3	UK vaccination uptake, reported vs WHO target	22
Figure 1-4	Spatial analysis of routine childhood vaccination coverage	23
Figure 1-5	Annual reported uptake of MMR1 in England, 1990-91 to 2015-16	24
Figure 1-6	Parents who consider MMR a greater risk than the disease it protects against	24
Figure 1-7	UK spontaneous ADR reports received for MMR vaccine per year (1998-2014)	25
Figure 1-8	Measles cases in England, 1996-2016	25
Figure 1-9	Confirmed measles cases in England by region and by age 2008-15	26
Figure 1-10	Measles cases in Wales outbreak 2012-13	28
Figure 1-11	Component concepts within the research question, with relevant potential interactions	34
Figure 1-12	Models of vaccine behaviour	37
Figure 1-13	Theory of Culture "grid-group" mapping of social relations	38
Figure 2-1	Geographic distribution of cases in the Merseyside and Manchester outbreaks	44
Figure 2-2	Map of PCTs of greatest epidemiological interest	47
Figure 2-3	Example geography hierarchy within one ward	52
Figure 2-4	Map of MMR1 uptake by ward in Great Yarmouth & Waveney PCT 2011-12	56
Figure 2-5	Box plots of MMR1 uptake at different nested spatial granularities	56
Figure 2-6	MMR1 uptake vs single-variable ward-characteristics	58
Figure 2-7	Principal components of compositional data ward-characteristics	59
Figure 2-8	MMR1 uptake vs selected principal components of compositional data ward characteristics	60
Figure 2-9	MMR1 uptake vs selected measures of deprivation	64
Figure 2-10	CATPCA for deprivation domains incl. income deprivation	65
Figure 2-11	CATPCA for deprivation domains incl. IDACI	67

Figure 2-12	Observed values of MMR1 (mean and 95% CI) vs predicted values from the model	72
Figure 3-1	Model stages for one simulation within a specified scenario	79
Figure 3-2	Distribution of vaccine-support (%) post decision process	86
Figure 3-3	A single decision in a population without network structure – expected outcomes	87
Figure 3-4	A single decision in a population without network structure - expected outcome compared with opinions of those canvassed	88
Figure 3-5	Intra-dyad agreement (%) post decision process	90
Figure 3-6	Intra-dyad agreement post decision process, observed vs expected value	91
Figure 4-1	Networks to be discovered	96
Figure 4-2	Networks' overlap	99
Figure 4-3	Survey presentation routes	111
Figure 4-4	Fieldwork period in context	112
Figure 4-5	Inference of MMR opinion and MMR1 status	117
Figure 4-6	Recruitment funnel	119
Figure 4-7	Roll sizes of recruitment centres	120
Figure 4-8	Response rate by centre by presentation	122
Figure 4-9	Example reconstructed linked networks from one centre	127
Figure 4-10	Advisors sample by relationship with respondent	128
Figure 4-11	Advisor distribution	130
Figure 4-12	Estimates of advisor categories accessed	131
Figure 4-13	Contacts sample by context	133
Figure 4-14	Contact distribution	134
Figure 4-15	Proportion of children with any contacts in specified context	135
Figure 4-16	Sentiment communicated	137
Figure 5-1	Extended model overview	152
Figure 5-2	Information network – MVD estimation	159
Figure 5-3	Best fits for Information network candidate degree distributions	161
Figure 5-4	Potential Infection network – MVD estimation (excluding childcare)	162
Figure 5-5	Best fits for Potential Infection network candidate degree distributions	163
Figure 5-6	Childcare room contacts' distribution	164

Figure 5-7 Best fits for response function (expected outcome vs observations) 167
Figure 5-8 Model stages for one simulation 169
Figure 5-9 Example rewiring to tune the overlap between networks 170
Figure 5-10 Summary of post decision process outcomes: vaccine-support (%), intradyad agreement and the index of intradyad agreement vs expected 173
Figure 5-11 Summary of post decision process outcomes – unstratified data, sensitivity to MVD parameter 174
Figure 5-12 Summary of post decision process outcomes – stratified data, sensitivity to MVD in both degree distributions 175
Figure 5-13 Summary of post decision process outcomes – stratified data, under varying proportions of lone and couple parents 176
Figure 5-14 Summary of post decision-making outcomes – stratified data, using the “second best fitting” candidates for network structure and for decision algorithm 177
Figure 8-1 Distribution of vaccine-support (%) post decision process – N=400 214
Figure 8-2 Intra-dyad agreement (%) post decision process – N=400 215
Figure 8-3 Intra-dyad agreement post decision process, observed vs expected value – N=400 216
Figure 8-4 Distribution of vaccine-support (%) post decision process – vary β 218
Figure 8-5 Intra-dyad agreement (%) post decision process – vary β 219
Figure 8-6 Intra-dyad agreement post decision process, observed vs expected value – vary β 220
Figure 8-7 Distribution of vaccine-support (%) post decision process – vary α 222
Figure 8-8 Intra-dyad agreement (%) post decision process – vary α 223
Figure 8-9 Intra-dyad agreement post decision process, observed vs expected value – vary α 224
Figure 8-10 EAL vs MMR1 uptake 236
Figure 8-11 Observed opinion status vs alters with same status 238

List of tables

Table 2-1	Summary of ward-characteristics included in the analysis	51
Table 2-2	Comparison of domain contribution to deprivation measures derived from the 2-dimensional CATPCA solution	68
Table 2-3	Measures of deprivation	68
Table 2-4	Summary of ward-characteristic variables	69
Table 2-5	Univariate analysis	70
Table 2-6	Model parameters	71
Table 3-1	Decision algorithm formulations	82
Table 3-2	Post decision process vaccine-support (%) – mean	84
Table 4-1	Questionnaire specification	103
Table 4-2	Pilot survey respondent characteristics	106
Table 4-3	Shortlisted PCT	110
Table 4-4	Survey formats offered by recruitment centres to parents	120
Table 4-5	Sample characteristics – adults	123
Table 4-6	Respondent characteristics – children	124
Table 4-7	Categories of advisors accessed by respondent	131
Table 4-8	Mean weekly contacts by context	136
Table 4-9	Children's MMR status	139
Table 4-10	Parent's MMR participation	140
Table 4-11	Ego-alter ties within 'information' network	142
Table 4-12	Ego-alter ties within 'potential infection' network	144
Table 5-1	Summary recap of candidate responses functions	158
Table 5-2	Network structure: best fit degree distributions and parameters	165
Table 5-3	List of model parameters	171
Table 8-1	EAL in shortlisted PCT	235
Table 8-2	Stratified MMR uptake for survey sample	237

List of other numbered items

Box 4-1	Survey questions	114
Box 4-2	Indications of scale	138
Equation 5-1	Watts-Strogatz algorithm degree distribution	153
Equation 5-2	Barabási Albert algorithm degree distribution (non-limit)	154
Equation 5-3	Expected proportion of post-decision adhere nodes	158

Abbreviations

ABMHB	Abertawe Bro Morgannwg Health Board
ADR	Adverse Drug Reaction
BSCS	British Social Contact Survey
CATPCA	Categorical Principal Component Analysis
CCG	Clinical Care Commissioning Group
CHIS	Child Health Information System
CI	Confidence Interval
COI	Central Office of Information
COVER	Cover of Vaccination Evaluated Rapidly
df	Degrees of Freedom
DH	Department of Health
DH/COI CITS	DDH/COI Childhood Immunisation Tracking Survey
EAL	English as Additional Language
ESEN2	European Sero-Epidemiology Network 2
FRP	Family Reference Person
GB	Great Britain
GCSE	General Certificate of Secondary Education
GLM	Generalised Linear Model
GP	General Practitioner
GY&W	Great Yarmouth & Waveney (PCT)
HCP	Healthcare Professional
Hib	haeomophilius influenza type b
HPA	Health Protection Agency
IDA	Intra-dyad Agreement
IDACI	Income Deprivation Affecting Children Index
IMD	Index of Multiple Deprivation
IOD	English Indices of Deprivation
LSOA	Lower Super Output Areas
MHRA	Medicines & Healthcare products Regulatory Agency
MLE	Maximum Likelihood Estimation

MMR	Measles Mumps Rubella (3-in-1 vaccine)
MMR1	1st dose of MMR vaccine
MMR2	2nd dose of MMR vaccine
NAFIS	National Association of Family Information Services
NeSS	Neighbourhood Statistics
NHS	National Health Service
ONS	Office for National Statistics
PCT	Primary Care Trust
PHE	Public Health England
Pre-c/u	Pre (MMR) catch-up campaign (measured at start of 2013)
QAS	Questionnaire Appraisal System
QOF	Quality and Outcomes Framework
SEIR	Susceptible Exposed Infectious Recovered
SHA	Strategic Health Authority
SIR	Susceptible Infected Recovered
SNA	Social Network Analysis
UK	United Kingdom
USA	United States of America
VAF	Variance Accounted For
WHO	World Health Organisation

General note on terminology

This thesis makes reference to the administrative regions and bodies used by the National Health Service (NHS) and Department of Health (DH) to organise primary care and public health services in England. During the time-period referenced within the thesis, there has been a significant restructuring as a result of the Health and Social Care Act 2012 [1].

Of greatest relevance for this research, on 1st April 2013 Primary Care Trusts (PCT) were abolished, with Clinical Commissioning Groups (CCG) taking-over most of their primary care service organisation responsibilities, and Public Health England (PHE) was established, taking-over the publication of several surveillance datasets (including routine immunisation uptake and notifiable disease incidence) from the simultaneously disbanded Health Protection Agency (HPA) [2] [3]. The national reports on coverage of the vaccinations included in the routine schedule continued to use PCT definitions and terminology until 31st March 2016 [4], so we follow this precedent and refer to PCT throughout.

1. General Introduction

1.1. Childhood vaccination in the United Kingdom (UK)

Vaccination is a cornerstone of public health, protecting individuals from the corresponding infectious disease through induced immunity, and performing a key role in programmes for the eradication or elimination of diseases such as polio and measles.

The UK routine baby immunisation programme [5] is designed to protect children from dangerous vaccine-preventable diseases and to deliver the levels of population immunity required to control the disease. But consent, given by those with parental responsibility, is required for the child's vaccination. Securing parents' support for vaccination is therefore vital to achieve the desired programme participation.

This thesis explores the patterns of routine childhood vaccination coverage (focussing primarily on the MMR vaccination, which has sub-optimal uptake in England) and how information-exchange across social networks might influence these vaccination decisions, so contributing to local variations in uptake.

1.1.1. Routine vaccinations

The Department of Health (DH) recommends a programme of routine childhood vaccinations to protect against twelve childhood infections: diphtheria, tetanus, pertussis, haeomophilius influenza type b (Hib), polio, meningococcal serogroup C and serogroup B, rotavirus, measles, mumps, rubella and pneumococcal [5]. The recommended schedule (as at the end of 2016) for children under 5 years old is given in Figure 1-1 .

These vaccinations are voluntary, but parents are encouraged to participate fully in this programme to protect their child. Healthcare professionals (HCPs) such as General Practitioners (GPs) and Health Visitors are instructed that ‘Every effort should be made to ensure that all children are immunised’ Public Health England [5] p81). The injections are usually administered via GP surgeries and the standard recommended vaccines are available free-of-charge through the National Health Service (NHS).

Figure 1-1 UK routine vaccination schedule for under 5s

Age of child	Vaccine	Infection(s) against which vaccine provides protection
8 weeks	DTaP/IPV/Hib	Diphtheria, tetanus, pertussis, polio, Hib
	PCV	Pneumococcal
	MenB [(b)]	Meningococcal B
	Rotavirus [(a)]	Rotavirus
12 weeks	DTaP/IPV/Hib	Diphtheria, tetanus, pertussis, polio, Hib
	Rotavirus [(a)]	Rotavirus
16 weeks	DTaP/IPV/Hib	Diphtheria, tetanus, pertussis, polio, Hib
	MenB [(b)]	Meningococcal B
	PCV	Pneumococcal
12 months [(c)]	Hib/MenC	Hib, meningococcal C
	PCV	Pneumococcal
	MMR	Measles, mumps, rubella
	MenB	Meningococcal B
40 months [(c)]	DTaP/IPV or dTaP/IPV	Diphtheria, tetanus, pertussis, polio
	MMR	Measles, mumps, rubella
From 2 years (annually)	LAIV [(a)]	Influenza

Children now under 5 includes those subject to previous schedules:

(a) Rotavirus and (annual) influenza were introduced in 2013.

(b) MenC doses removed in 2013 and 2016, and MenB was introduced in 2015

(c) Wording of 12 month and 40 month vaccinations' window tightened in 2012

Department of Health [5]

Other vaccines to protect against childhood infections are marketed, but are not offered via the NHS although they may be obtained via private practice [6, 7]. These include the single-antigen vaccines for measles, for mumps and for rubella (popularly referred to as "singles") which may be administered in place of MMR [8].

Reduced vaccine uptake has public health implications, due to reduced herd immunity, in addition to the disease risk for the unvaccinated children themselves.

1.1.2. Surveillance instruments

1.1.2.1. Percentage uptake of routine childhood vaccinations

Uptake of routine childhood vaccinations is monitored via the COVER programme (Cover of vaccination evaluated rapidly). The percentage of children who are up-to-date for each age-appropriate vaccination is measured at one, two and five years old.

In England, COVER is administered by Public Health England (PHE) [9], and previously the Health Protection Agency (HPA) [10]. The vaccine uptake statistic was reported at Primary Care Trust (PCT), Strategic Health Authority (SHA) (or Region) and nation levels from 1 April 2003 to 31 March 2016. Since the NHS restructuring in 2013 (see General Note on Terminology, above), COVER reports for Local Authority areas (with different inclusion criteria) and a parallel programme via the UNIFY system [11] have been introduced in England (the latter is described as an operational management tool, not surveillance, and is subject to less stringent data assurance). All these data are ultimately sourced from GPs: GP surgeries input vaccination uptake into their PCT's computerised Child Health Information System (CHIS) [9]. The PCT collates the required statistics quarterly and annually, and forwards them to PHE – the form completed by PCTs details the methodology used, including numerator and denominator definitions, see Appendix for example form showing PCT denominator definitions [12].

Similar programmes monitor uptake in the other nations within the UK: via Health Protection Scotland [13], National Public Health Service for Wales [14] and Communicable Disease Surveillance Centre Northern Ireland [15]. Annual reports combine data from all four nations to report a UK statistic.

COVER is the sole surveillance system of all routine childhood immunisations' uptake in NHS surgeries; it provides summary epidemiological data with minimal delay (quarterly reports are published three months after quarter-end) [16] and is backed by government mandate. However there are some omissions and uncertainties.

Data is rarely complete for all fields for every PCT, particularly for quarterly reports, with a number of PCTs experiencing technical or other reporting issues. The separate call for annualised information direct from PCTs enables fuller data-reporting than summing quarterly reports, so is preferred for analysing trends in uptake. There is a data quality issue, data on birthdates (hence correct evaluation quarter for each child) may be inaccurate or missing and the calculation of the PCT denominator is not straightforward, it is not the

GP-registered population. Furthermore GP-registration data are subject to quality concerns, for example duplicate registrations (which could be inter- or intra- PCT) [17] and a quality assurance report by Office of National Statistics (ONS) [18] found both under- and over-reporting concerns, which would result in artificially inflated or depressed COVER figures respectively, and GP funding schemes with potentially distortive effects.

Also, it is noted that COVER excludes non-routine immunisations, notably single-antigen vaccines for measles, mumps and rubella. Therefore there is no surveillance data which enables an estimate of protection obtained from, say, all measles-containing vaccines. Lack of measurement of "singles"-derived protection remains a weakness in the surveillance data available for analysis relating to the potential for infection outbreaks.

An indication of "singles" vaccines uptake can be gained from one-off studies (although these data are not compatible with COVER, so cannot be combined to give an overall protection statistic). In 2001-2002 records obtained from clinics and details of vaccine imports lead to an estimation of the absolute contribution of these vaccines as relatively small ("single" measles doses were equivalent to 1.7% and 2.1% of the 2-years-old cohort in each year in England & Wales) [8]. A large-scale cohort study of children reaching 2 years old in 2002-2004, found 5.2% had received at least one "single" vaccine [19].

1.1.2.2. Parental opinions, attitudes and behaviour

From 1991 the DH commissioned, in conjunction with the now defunct Central Office of Information (COI), market research on the immunisation knowledge, attitudes and behaviour, and types of information sources, of parents of children aged 0-2 years [20]. This DH/COI childhood immunisation tracking survey (DH/COI CITS) was conducted annually from 2005 and the sample expanded to include parents of 3-4 year olds in 2010 [21] . The final wave of fieldwork was conducted in January-February 2010 and consisted of in-home interviews with 1730 parents of pre-school children, 1142 of whom had children aged 0-2 years [21]. The sampling strategy was designed to deliver respondents that are demographically and geographically representative of the UK. PHE have commissioned 3 equivalent annual studies (fieldwork in 2015-2017); it is anticipated that data from the first of these will be published in 2017 [22].

This body of evidence enables tracking of the relative attitudes to routine vaccinations for each population of parents passing through the period when their children are offered routine vaccinations. Papers summarising several years' data from DH/COI CITS have been

published in peer-reviewed journals [20, 23] Within these articles, the potential for availability bias in the sample, due to quota-based sampling is acknowledged.

1.1.2.3. Serological status

There is no on-going surveillance of the serological status of the UK population with regards to the vaccine-preventable diseases targeted in the routine childhood immunisation programme. The most recent published sero-surveillance in England & Wales for measles used samples collected in 2000, and included analysis of serostatus vs measles and rubella, as part of the European Sero-Epidemiology Network 2 (ESEN2) [24]. The source of the sera (residual samples from routine laboratory testing) introduces bias.

1.1.2.4. Cases of vaccine-preventable diseases

The list of notifiable diseases includes vaccine-preventable childhood diseases: acute meningitis, acute poliomyelitis, diphtheria, measles, mumps, pertussis, rubella and tetanus [25]. For these infections, doctors have a statutory duty to report suspected cases, and laboratory reports are also used to collate incidence data. Incidence data are supplied to the PHE who publish the weekly Notifications of Infectious Diseases report. Incidence is reported at national, regional and local authority level [26].

The use of multiple sources is believed to give a fuller picture, but can also produce double-counting of cases.

1.1.2.5. Adverse reactions to vaccines

As part of their remit for post-market monitoring of pharmaceutical product safety, Medicines & Healthcare products Regulatory Agency (MHRA) operates the Yellow Card scheme for the reporting of adverse drug reactions (ADR) to medicine (including vaccines) [27]. ADR may be reported voluntarily by healthcare professionals (HCP) and by patients; data from the scheme cannot be used to calculate the incidence of ADR nor the proportion of ADR reported to the authorities. The association of the reported symptom(s) with the vaccine need only be suspected, proven causality is not required, and the assessment of a suspected association may not be consistently applied by the range of possible reporters.

1.1.3. Overview of vaccination behaviour and infection incidence

In Figure 1-2 we list national events relating to routine childhood vaccinations (for under 5s), with a focus on MMR, to provide a context for surveillance data interpretation.

Figure 1-2 Timeline of events 1996-2016

Year	Changes to routine schedule	MMR-specific events
1996		MMR2 introduced
1998		MMR-autism link published (The Lancet)
1999	MenC introduced	
2001-2		Media coverage of subsequent MMR-autism papers
2004	DTaP/IPV/Hib '5-in-1' vaccine introduced	
2006	PCV introduced	
2008		MMR catch-up campaign
2009	Pandemic influenza vaccine offered	
2010		MMR-autism lead author struck-off
2013	Rotavirus introduced, seasonal influenza introduced	MMR catch-up campaign
2016	MenB introduced, MenC phased out	

Department of Health, British Broadcasting Corporation [28-32]

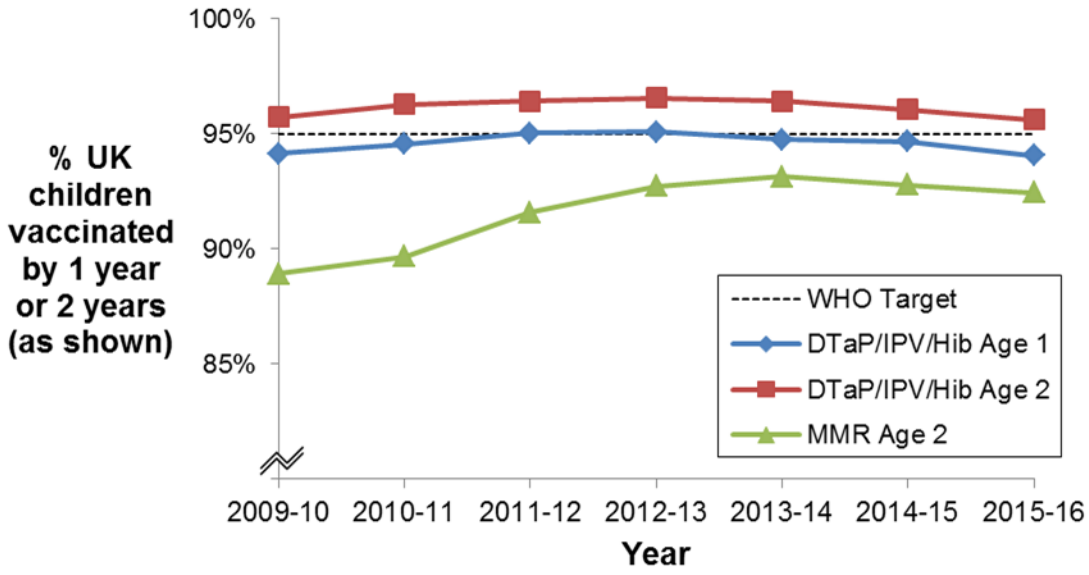
The World Health Organisation (WHO) sets targets of 95% national vaccination coverage by two years old for protection against measles, rubella, mumps, diphtheria, polio and tetanus and "in infancy" for pertussis [33]. This corresponds to the upper range of the estimated critical proportion for pertussis and measles (i.e. estimated population vaccinated proportion required for infection elimination via herd immunity effect) [34]. Suboptimal coverage exposes the population to the risk of disease outbreaks, from endemic or imported infection.

1.1.3.1. UK routine childhood vaccination participation

The majority of UK children participate in the routine childhood vaccination programme [35-42].

For the UK, WHO targets for diphtheria, polio and tetanus have been met every year since 2009-10 (Figure 1-3), but those for the MMR vaccine have not been achieved. Within the UK, England has the lowest cover of the four constituent countries (all strata of vaccine, age and year).

Figure 1-3 UK vaccination uptake, reported vs WHO target



COVER [35-42]

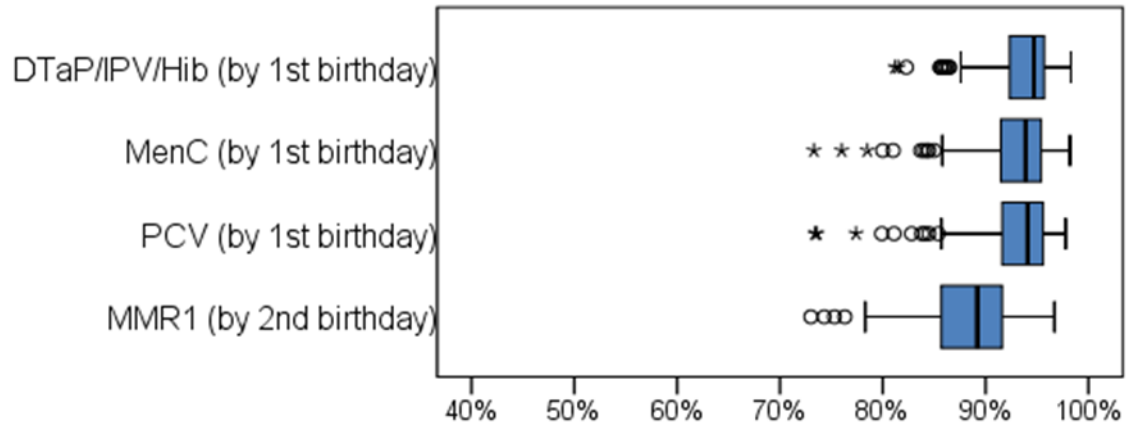
The majority of parents claim to "automatically have their child's (pre-school) immunisations done when they were due" (75% of those with 0-2 year olds) [21]. However, the awareness of the constituent elements of the programme varies considerably (from 73% spontaneous awareness for MMR to 12% for the pneumococcal vaccine) and the vaccines themselves are viewed as "completely safe" by no more than 58% of parents (of 0-4 year olds).

Within England there is considerable spatial variation in uptake of primary vaccination courses (Figure 1-4, measured at PCT or Local Authority level), and London consistently has the lowest regional uptake across this period. Considering the WHO district vaccination targets [33], at this granularity the majority of districts meet those for diphtheria, polio and tetanus vaccination (90% at two years), and the majority do not meet those for measles, mumps, and rubella vaccination (95% at two years). Less is published at geography more

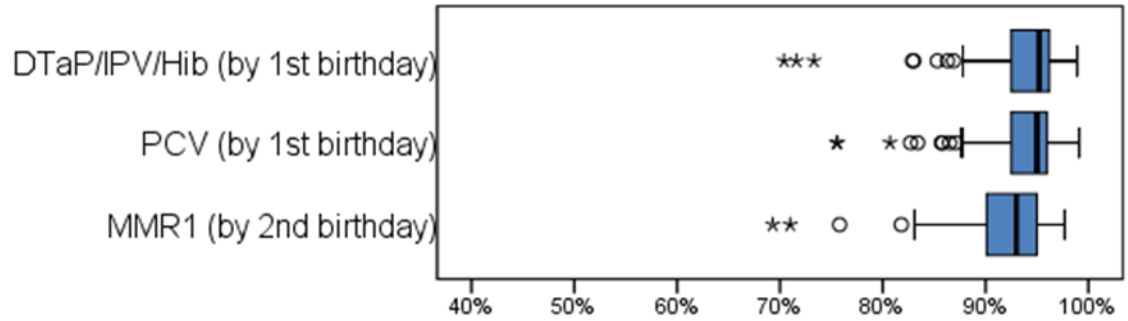
granular than by PCT, but management audit data [43, 44] indicate that this variability persists at smaller scales.

Figure 1-4 Spatial analysis of routine childhood vaccination coverage

a] 2009-10 primary course uptake by PCT



b] 2015-16 primary course uptake by Local Authority



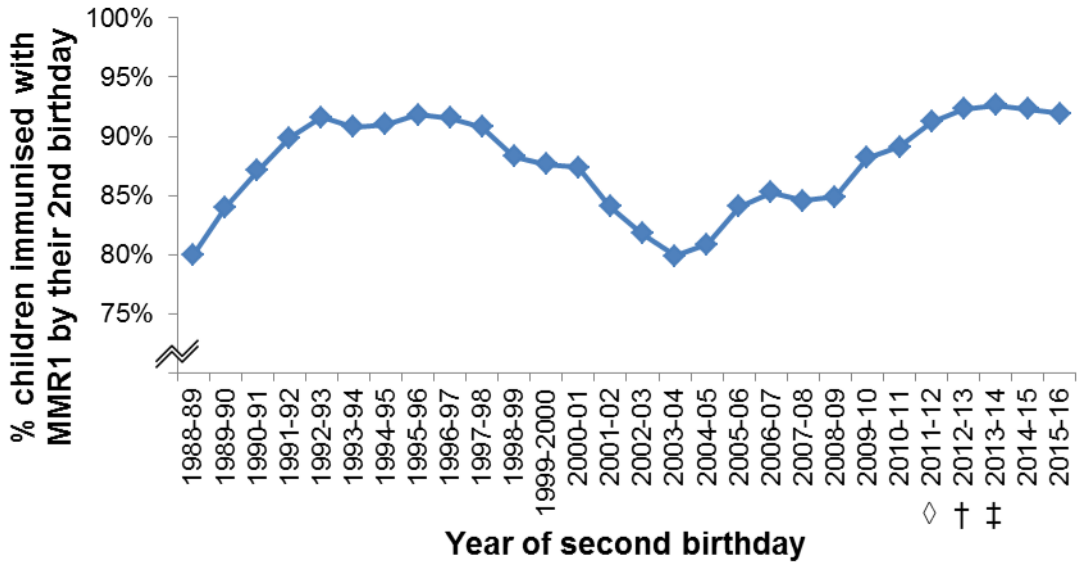
NHS Information Centre, Screening and Immunisations Team, NHS Digital [36, 42]

1.1.3.2. MMR1 and measles

There is a clear temporal pattern for MMR1 uptake in England, since its introduction in 1988 (Figure 1-5). After a period of stability following its establishment, there was a sustained drop in uptake, from the late 1990s to the early 2010s, with the lowest coverage in 2003-04. We are not aware of any concurrent stock or access issues (MMR nor childhood vaccinations generally e.g. diphtheria vaccine coverage remains stable at over 90% at 12 months [42]).

The available cross-sectional serological data was collected prior to this trough. The ESEN2 analysis of samples collected in 2000 (18.9% of 2-4yr olds were sero-negative vs measles [24]), is not inconsistent with the estimated 90% sero-conversion [28] from the relevant years' MMR1 uptake data (87%-91%).

Figure 1-5 Annual reported uptake of MMR1 in England, 1990-91 to 2015-16

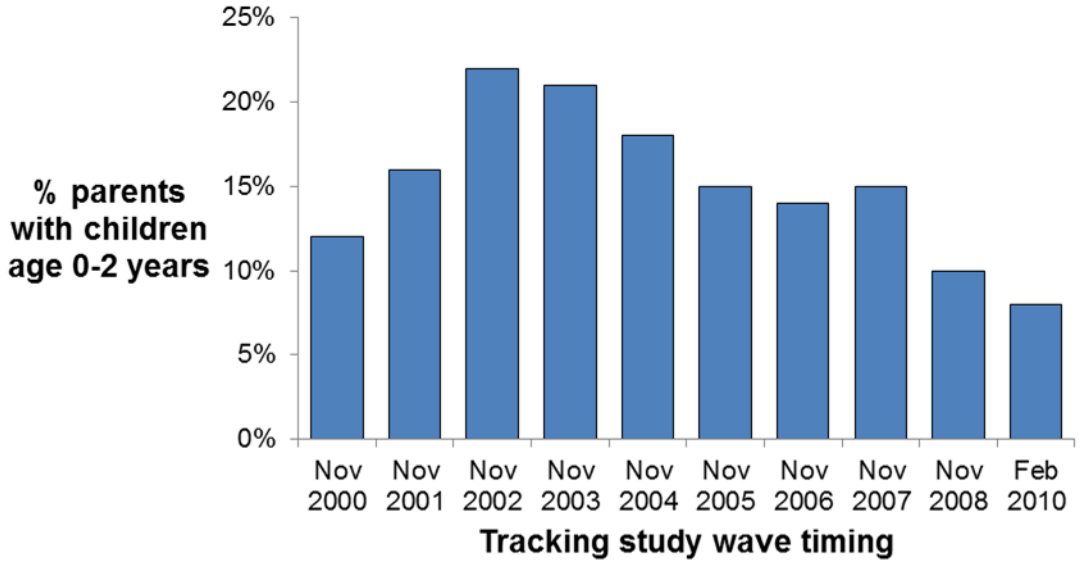


Comparison with research timing
MMR1 data analysed in Chapter 2 (◊)*; Fieldwork in Chapter 4: pilot* (†),*full survey* (‡)

NHS Digital [42]

Data from the annual attitudinal surveys (Figure 1-6) show a peak in the proportion of parents who consider MMR a greater risk than the disease it protects against, which is near-synchronous with the uptake trough (peaking about a year earlier).

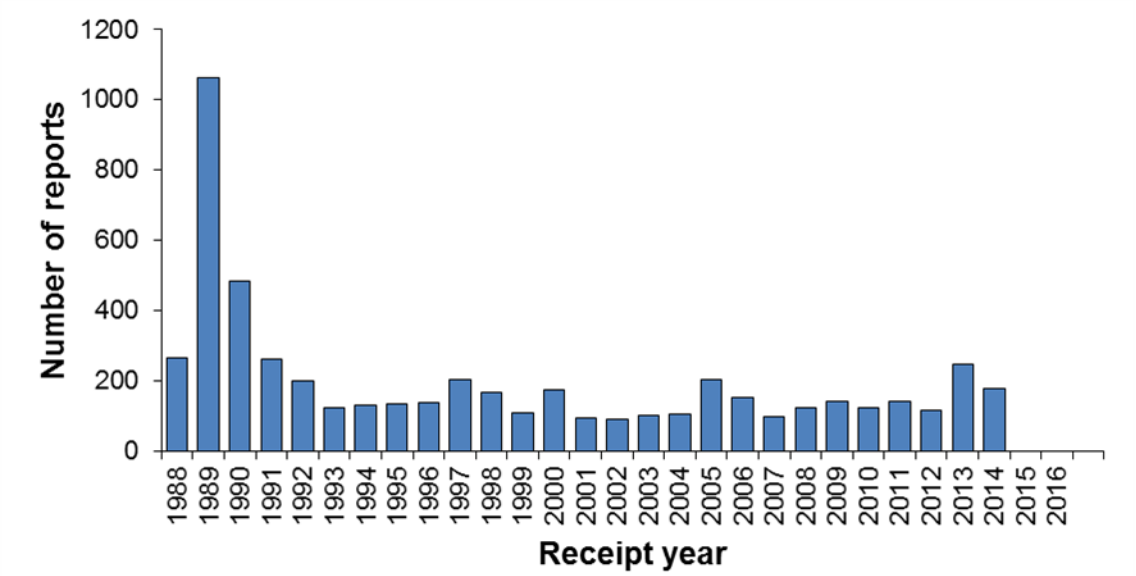
Figure 1-6 Parents who consider MMR a greater risk than the disease it protects against



DH/COI CITS [21]

However, there was no clear temporal association of higher ADR reporting (Figure 1-7).with neither the dip in MMR coverage nor the peak in relative perceived risk noted above.

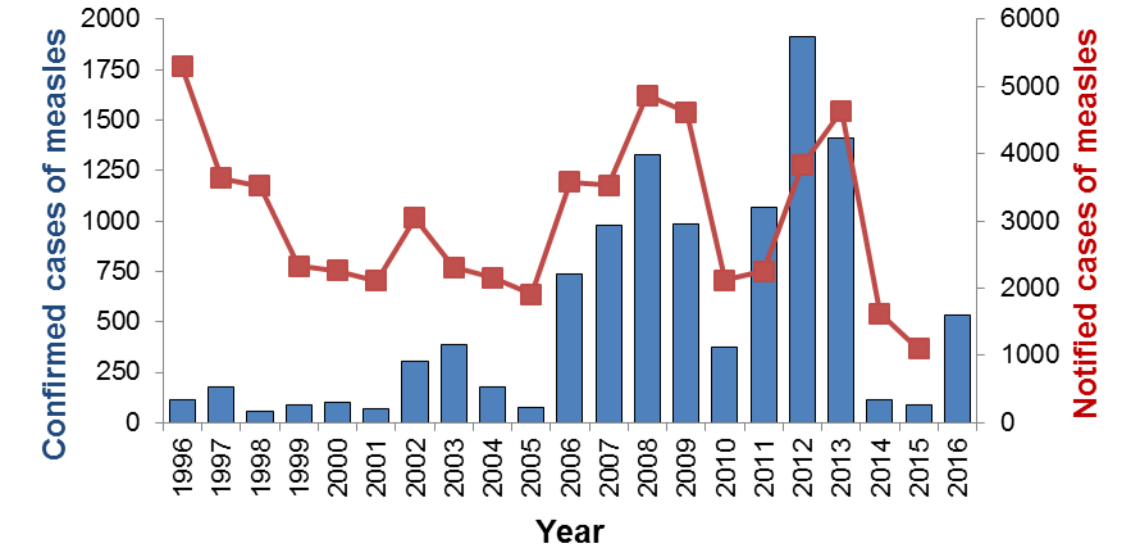
Figure 1-7 UK spontaneous ADR reports received for MMR vaccine per year (1998-2014)



MHRA [45]

After being restricted to localised outbreaks in the last 1990s, measles was declared endemic again in the UK in 2008 [46] . Since 2000, there has been an overall increasing trend in confirmed measles cases in England (Figure 1-8). An association between the trough uptake and this increased incidence is hypothesised; earlier analysis of the cases across 1995-2002 associated the decline in MMR uptake with an increase in outbreak size [47].

Figure 1-8 Measles cases in England, 1996-2016

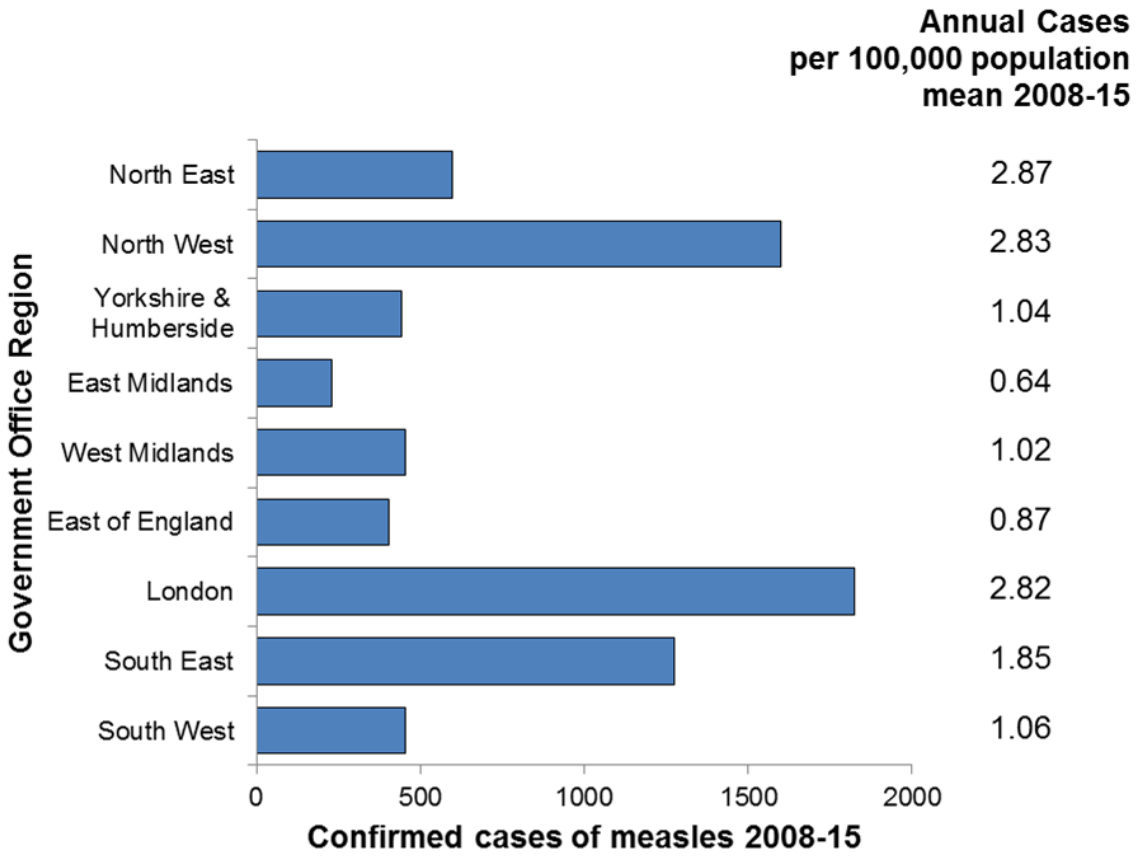


Confirmed cases include saliva IgM positives and/or PCR and laboratory reports

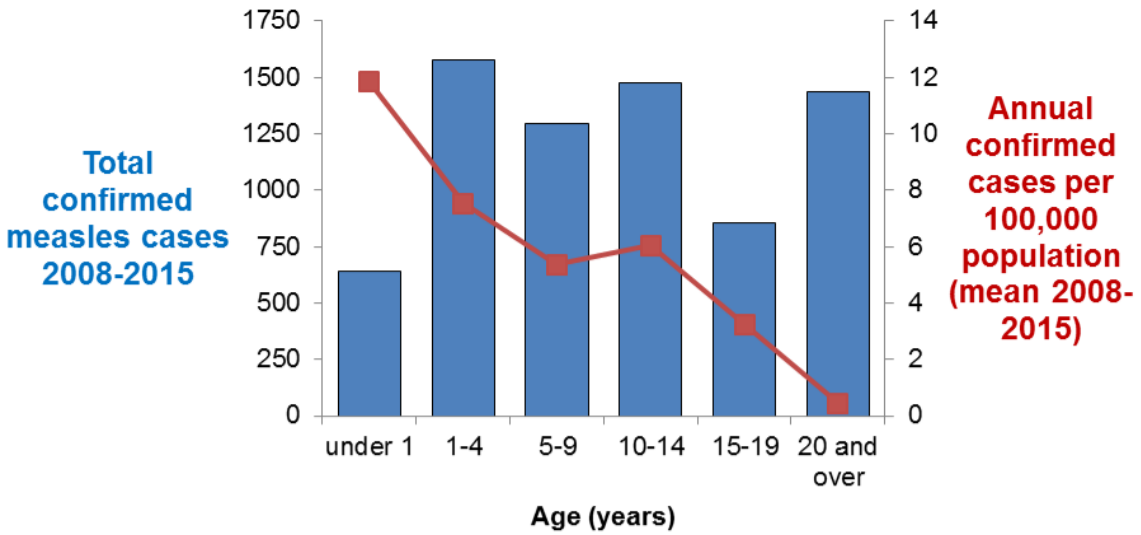
PHE HPA [48-51]

Figure 1-9 Confirmed measles cases in England by region and by age 2008-15

a] confirmed cases by region



b] confirmed cases by age



Confirmed cases include saliva IgM positives and/or PCR and laboratory reports

HPA, PHE, ONS, [48, 49, 52]

Considering the endemic period (2008 onwards), the incidence of measles is also geographically heterogeneous within England (Figure 1-9a), with the highest cases per

capita in North West (peak in 2012), North East (peak in 2013) and London. 28% of confirmed cases are in children age 0-4, with the highest cases per capita in under1 year olds (before routine vaccination age) (Figure 1-9b).

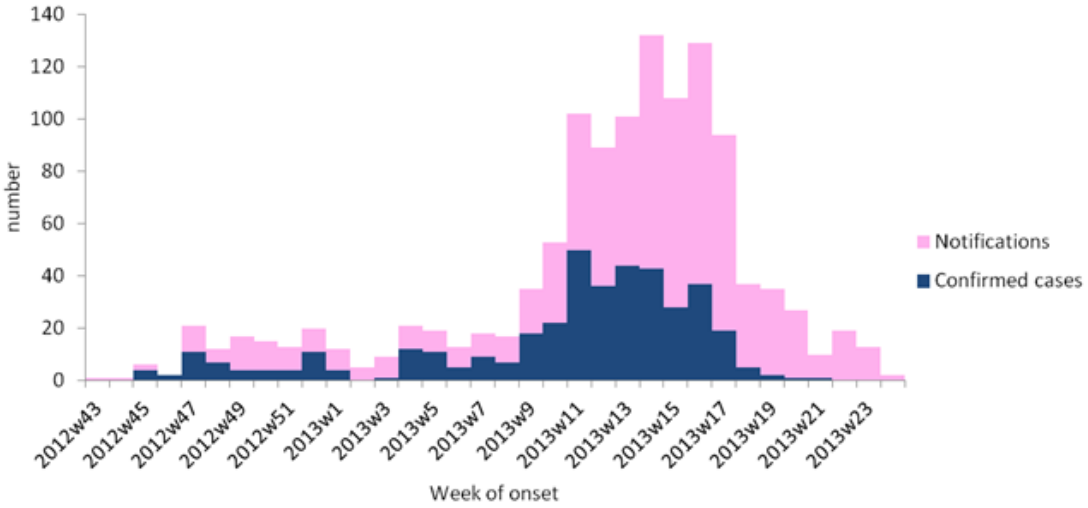
1.1.3.3. 2013 measles outbreak and MMR catch-up campaign

A large measles outbreak occurred in Wales in 2012-13, specifically in the Health Board areas of Abertawe Bro Morgannwg (ABMHB), Hywel Dda and Powys (with the index cases in Swansea) [53]. There were 1,202 notified cases, 88 hospital admissions and one death. Incidence peaked in March-April 2013 (Figure 1-10a). The majority of notified cases were in children (Figure 1-10b), with the highest cases per capita in those aged under 1 year (>350 notified cases per 100,000 population). The outbreak received substantial coverage in local and national media [54, 55].

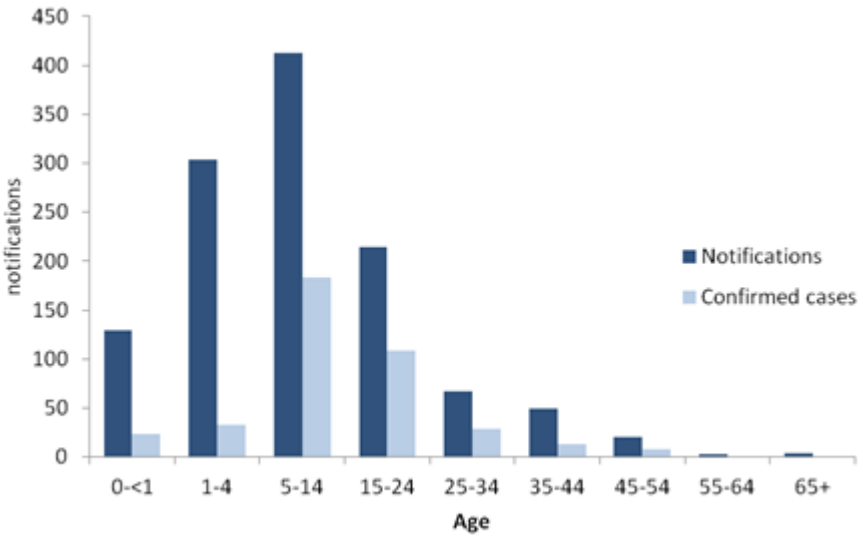
From 1998-2009, MMR uptake in ABMHB was consistently below the Welsh average (with lowest routine MMR1 uptake occurring in 2002-04) which was associated with a 1997 campaign by the main local paper [56] amplified by the UK-wide vaccine scare (§1.1.4.1). Outbreak control efforts addressed this vaccination gap (temporal and spatial) via additional vaccination opportunities (routine and catch-up) such as drop-in clinics and school activities. At least 77,805 catch-up doses of MMR were delivered before the outbreak ended.

Figure 1-10 Measles cases in Wales outbreak 2012 13

a] week of onset



b] age distribution



Source: Public Health Wales. (2013). Outbreak of measles in Wales Nov 2012 – July 2013. Report of the agencies which responded to the outbreak, dated October 2013. Cardiff: Public Health Wales NHS Trust

In response to the increase in measles incidence (including the Swansea outbreak), in April 2013 the DH announced a national MMR catch-up campaign [30]. The focus of the campaign was 10-16 year olds with sub-schedule vaccination history. This cohort would have been due for routine MMR vaccination in the late 1990s and early 2000 (§1.1.4.1). The objective of the campaign was to achieve 95% uptake of 1+ doses of MMR in this age-group

[57] and all GP practices were expected to proactively search for un-/under-vaccinated individuals and offer the catch-up vaccination(s).

An evaluation of the campaign [58], using a weighted sample of target-age children with no MMR vaccination recorded in CHIS, showed 11% were vaccinated during the catch-up. The 95% coverage (in 10-16 year olds, based on GP records) was achieved by the end of August 2013. MMR doses ordered by GPs indicate that levels of MMR2 and MMR vaccination of other age-groups may also have been positively affected by the campaign and/or measles outbreak publicity. A longitudinal comparison of GP records for the target cohort showed a 1.8% decrease in unvaccinated children year-on-year [59]. Both studies highlighted difficulties with vaccination data associated with patient mobility and software issues.

1.1.4. Understanding the sub-optimal uptake of MMR(1)

In the UK, more children are either not included, or sub-optimally included, in the programme for MMR than for other routine vaccinations and the UK (as a whole) has yet to meet WHO's national vaccination target for the associated pathogens . Although national levels of MMR uptake recovered to levels seen before the trough of the early 2000s, there is substantial geographic variation and the disease threat is very real in some areas. Indeed during this period of recovery, Wales experienced the largest measles outbreak since MMR's introduction (§1.1.3.3). In terms of offering an opportunity to study a vaccination-decision process, MMR is the vaccination for which the pros and cons are consciously weighed-up by more parents than any other routine childhood vaccination [21].

Hence, it is proposed to explore MMR1uptake as a means to protect against measles: there are patterns in uptake that are of interest, and as it has the greatest potential for parents to be able to provide information regarding the vaccination-decision process.

1.1.4.1. MMR safety scare in the late 1990s and early 2000s

The temporal pattern in MMR uptake is associated with a well-documented vaccine-safety scare [60, 61]. Research published in 'The Lancet' in 1998 connected MMR to autism and bowel disorders [62] and received mainstream media coverage [63]. Subsequent papers by the same lead author sparked further public interest in 2001, and coverage was widespread in print, broadcast and online media, peaking in winter 2002 when it was fuelled by high-profile speculation over Leo Blair's vaccination status (baby son of the then Prime Minister)

[64, 65]. Subsequent research has refuted these adverse event claims [66] and the original paper has been fully retracted [67] and the lead author struck-off [32].

At the height of the controversy, 24% of parents (with children aged 0-2 years) believed that MMR posed 'a greater risk than the diseases it protects against' [23]. Whilst the prevalence of this concern has decreased slowly across the intervening years, risk of autism was still specifically cited by 20% of MMR-rejecters in 2008, and in 2010 MMR remained the routine vaccination with the lowest "completely safe" rating from parents (46%) [21]. Declines in uptake were greatest in affluent areas [68], areas with high population density [69] and in children of highly-educated parents [68-70].

The downturn in MMR coverage after 1998 (Figure 1-5) was most dramatic in the UK, although decreases were reported in the Republic of Ireland [71] and other parts of the English-speaking world [72], and autism as an adverse event associated with MMR is a concern for 30% of parents in Sweden [73]. .

1.1.4.2. Minority cultures

It is known that some cultures' beliefs result in unvaccinated clusters of that community's children. These cultural beliefs include the refusal of vaccination as espoused by the anthroposophic community (believing disease benefits the child) [74] and orthodox Calvinists (avoiding interference with divine providence) [75]. In contrast, the British ultra-orthodox Jewish community does not reject vaccination per se, but believes that relative cultural isolation reduces their risk for many diseases [76]. Another minority culture with low vaccination levels is the Traveller community; where access to healthcare poses an additional challenge to achieving high uptake. These example communities have also been associated with measles outbreaks in the UK [74, 77] or as sources of imported infections in similar communities overseas [78, 79] . However these identified communities do not account for all under-immunisation nor for the temporal pattern.

1.1.4.3. Evidence from quantitative and qualitative studies into MMR(1) uptake

A number of studies have been identified which have investigated MMR uptake, using quantitative, qualitative and mixed methods to investigate the relationships with demographic characteristics and personal beliefs of parents, and parent-community interactions. Given the MMR vaccine-scare outlined in §1.1.4.1, this review includes only studies with fieldwork

in 1999 onwards, in the UK, and including children without MMR vaccination (or their parents).

Personal beliefs regarding vaccination and vaccine-preventable disease

Reduced immunisation levels are associated with general concerns about side-effects and vaccine safety [80-86, 88-90] with personal awareness of incidents of serious reactions to vaccination or vaccine-attributed adverse events reported in some studies [83, 87, 88, 90, 92]. Qualitative research has suggested that parents frame these risks against the perceived vulnerability of their own child [94-96]

Some studies have found an association between low uptake and a perception of that MMR was not supported by sufficient medical research [89, 93], but perceived vaccine efficacy is uncorrelated to uptake [86, 93], which would indicate that it is safety research that is thought lacking. Reduced uptake is also associated with parental belief that combination vaccines, such as MMR, overload the child's immune system unlike 'natural' infection [81, 84, 89, 95].

Whilst the perception of vaccine disease-prevention efficacy is unrelated to uptake, there is an association between low uptake and lower levels of both the perceived seriousness of measles [81, 89, 93] (including qualitative citations of personal experience of serious vaccine-preventable disease morbidity [92]), and of the child's risk of exposure to the virus [88, 97]. Qualitative studies observe that some parents believe they can reduce the child's risk of exposure to vaccine-preventable diseases [91], unlike the child's reaction to vaccination which is out of their control [82].

Personal demographic characteristics

Associations with decreased MMR uptake have been observed with parents who are older [19, 89] and more highly qualified [19, 70, 89] There is no consensus on an association between MMR uptake and working status [83, 85, 93] and living in areas of deprivation [69, 98]. The largest quantitative survey analysed [19] did not provide an analysis of factors associated with MMR-rejection per se, instead including multivariate analyses of two MMR-rejecter subgroups ("singles"-users, and those with no vaccination) and the results presented for these two groups are inconsistent for work and income factors.

Similarly the impact of total family size is unclear [85, 86], including contradictory results for the two MMR-rejecter subgroups in Pearce et al [19]; although there is some evidence of higher uptake levels for first-borns [89, 93], and no associations were found with the number of parents present in the household [83, 85, 86] (again unclear for Pearce et al [19]).

The effect of ethnicity is most clearly observed from studies focussing on this factor [61, 99], with more sensitive categorisation, which have found highly significant, but not consistent, links between ethnicity and MMR1 uptake, and have advanced qualitatively-driven explanations such as senior family members' cultural role, family-ties in countries with higher vaccine-preventable morbidity, and limited use of English-language media.

Parent–Community Interactions

Several studies report a parental perception of self-interest in HCPs' recommendation of immunisation (or more general distrust or cynicism of government advice) which is then associated with reduced uptake [83, 89, 91-94, 97, 100, 101]. Advice is also sought from family, friends and the media [21, 80, 82] and other parents' opinions are trusted more than official advice [94, 95, 100, 102]:, with qualitative studies indicating that HCPs can cross this divide when they give advice drawing on their own experiences as a parent [88].

Lack of peer support for vaccination is associated with vaccine refusal [83, 89] and the perception of vaccination as a social responsibility is related to increased uptake [87, 89, 103] with a suggestion from qualitative studies that this is related to the risk of being seen as a "bad parent" [87, 97]. There is limited evidence of the impact of organised anti-vaccination groups [85, 93].

1.1.4.4. Emergence of the research question

With the exception of highly-educated parents, older parents and ethnicity (the latter sensitive to the measurement tool), these data indicate that MMR uptake is not well-predicted by demographic factors. Evidence is inconsistent regarding other demographic factors that have been found as correlating with overall routine immunisation programme participation (e.g. Samad et al [104] found deprivation, lone parenting, presence of siblings, high parental education and ethnicity as all being associated with low uptake).

Factors emerging from the synthesis of evidence relating to personal beliefs and interactions (with HCPs and others) may derive from individual differences in psychology and experience, but it is also possible that social or community-level processes contribute to their development. There is a lack of quantitative evidence clearly considering how these beliefs relate at different social scales – the analysis in the studies above considered differences at the individual level only (with one, ecological study, exception [105]).

These studies also point to active engagement with the question of whether to vaccinate with MMR, by both eventual vaccinators and non-vaccinators [91, 97, 103]. MMR vaccination

was not “automatic” for 25% of parents in 2010 [21]); and Pearce et al [19] found 75% of parents with non-MMR vaccinated children had made a “conscious decision” not to vaccinate.

Social contagion theory suggests that attitudes and behaviours may spread via interpersonal relationships and result in clustering of similarly-minded or similarly-acting individuals. Clustering of susceptible individuals is of epidemiological interest as it has been proposed as an explanation for outbreaks in otherwise well-protected populations, [106], e.g. in Switzerland and USA [107, 108]. Outbreaks in susceptible clusters may spread into the wider population, even in the presence of high immunisation levels. For example, the very large measles outbreak (>22,000 cases) in France during 2008-2011 (with 87%-90% MMR coverage nationally) originated in a group of religiously-motivated vaccine-rejecters [109] [111]. In addition to outbreaks seeded through local transmission, measles outbreaks in clusters of susceptibles embedded in the general population have been recorded, where the index case was an imported infection (e.g. in Denmark [110]). There is evidence that clustering of unvaccinated individuals can lead to major outbreaks at higher population immunisation coverage than if vaccination behaviour is homogeneous [112], hence disproportionately hindering attempts to eliminate vaccine-preventable diseases.

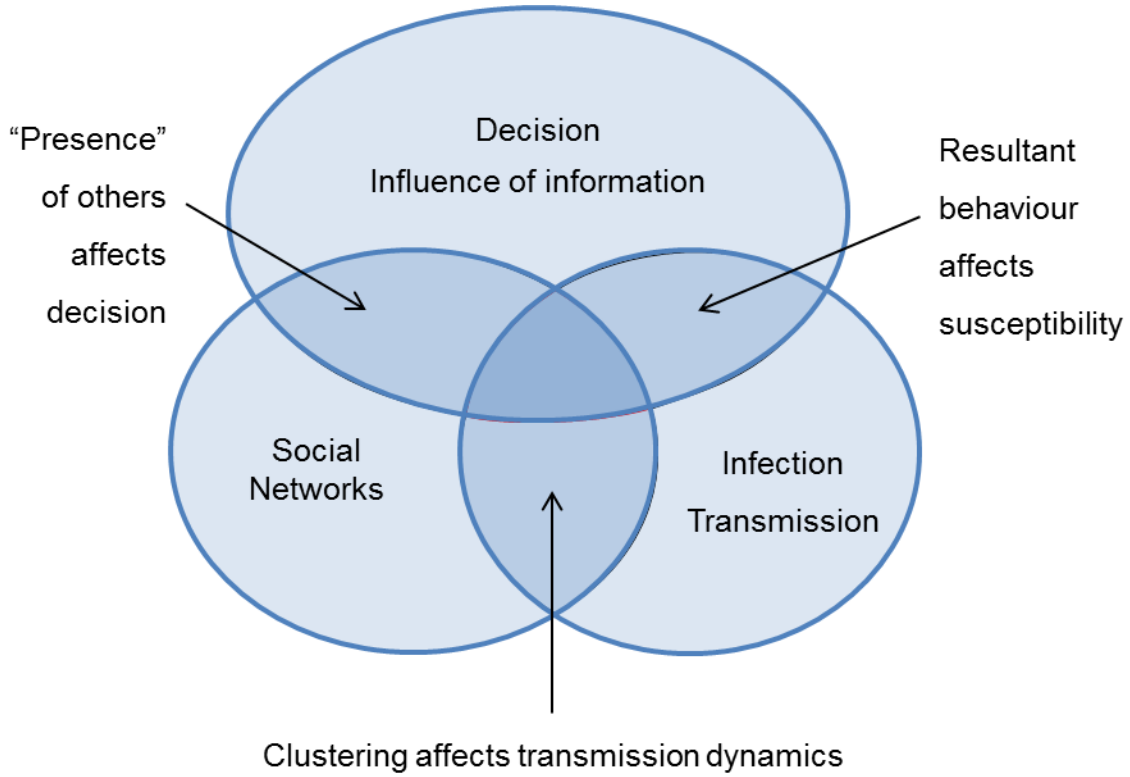
We hypothesise that vaccination-decision influences that act across social networks may generate clusters of individuals with similar vaccination opinions (specifically regarding MMR), and act as a mechanism that contributes to small-scale geographic variation in MMR1 uptake and so affect the potential for measles outbreaks

1.2. Mathematical modelling of the vaccine-decision process and resultant infection dynamics

As a part of the aim to understand the hypothesised processes, it is intended to build a mathematical model of the parental MMR vaccine-decision process and the measles infection dynamics in the corresponding child population.

A selection of models relevant to the component concepts within the research question (Figure 1-11) is given here. These models are of different types (not only mathematical) and from a range of original provenance (including epidemiology, psychology, economics, anthropology). This selection is not comprehensive in all areas, and is not intended to offer fully-detailed expositions or informed critiques of each model. However it is presented to provide a background to existing formulations that have been used to explore questions related to this research and to provide inspiration for possible alternative formulations.

Figure 1-11 Component concepts within the research question, with relevant potential interactions



1.2.1. Infection transmission models

There is a long history of mathematical models of measles transmission. Simple Susceptible-Infected-Recovered (SIR) and Susceptible-Exposed-Infectious-Recovered (SEIR) models have been used successfully and explored temporal patterns of measles cases observed prior to vaccination introduction in the UK, e.g. Fine & Clarkson [113] including the potential for chaotic dynamics e.g. Olsen & Schaffer [114] More sophisticated models have also been developed, notably the Realistic Age Structured model by Schenzle [115].

1.2.2. Network models

Social network analysis (SNA), in which individuals are represented by nodes, joined to each other by edges representing social connections, has been used to model a variety of processes in a range of fields, including health and specifically infectious disease dynamics [116, 117]. When compared with equivalent models using mean-field representations, these models (where potential transmission events only occur between directly-connected individuals) predict different infection dynamics, with the exact properties also dependent on the structure of the network itself. Network structure can be characterised by measures such as the node degree distribution and measures of clustering such as transitivity (presence of closed triads). The relationship between degree distribution and infection transmission across network is better understood that the influence of clustering [118]. Examples exist where the network structure increases the potential for infection outbreaks, e.g. power law degree distributions. The majority of the literature considers static networks, although some more recent models investigate dynamic networks. For information networks, there are two hypothesised processes of social contagion: “simple contagion” and “complex contagion” [119] in the latter, clustering may improve the successful adoption of innovations, through increased peer-reinforcement [120]

1.2.3. Decision models

Models of decisions have been developed in many social science disciplines, and these may have relevance for parents’ vaccination decisions. We indicate here some pertinent models, and inputs to these models (such as perceived risk); this treatment is necessarily brief.

Game theory has been applied to many decisions of strategies to be adopted from a discrete number of choices, initially in the field of economics [121]. It is assumed that individuals

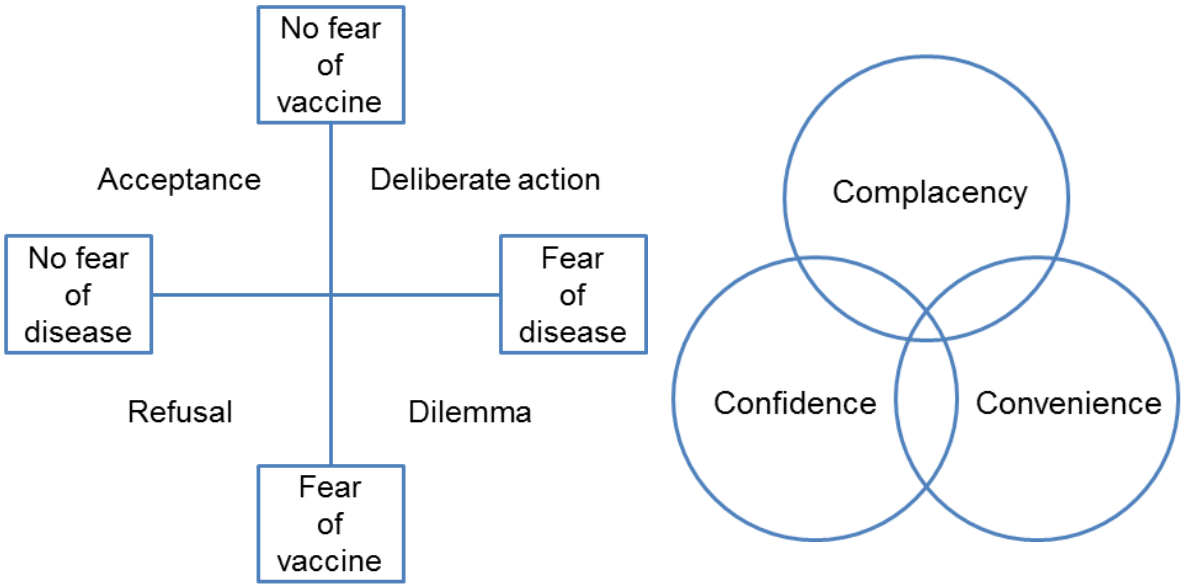
calculate the ‘payoff’ each strategy would achieve against the strategy adopted by others, and then adopt the strategy that maximises this individual value (which may not be the same strategy that gives the optimal utility calculated at a population level). When applied to decisions made by individual humans, this model raises the question as to what measures are used to inform the ‘payoff’; classical game theory assumes individuals have perfect knowledge of the strategies and their associated risks, costs and benefits. This may be an unrealistic assumption and some applications use perceived values or sample-based values in place of population statistics. Individual’s judgements of the payoff may be subject to biases, such as omission bias [122] and delay discounting [122], and if strategies are updated during the individual’s lifetime, there is the question of whether perceptions of payoffs are dependent on the existing strategy state.

Psychologists have developed of models to specifically explain adoption of health behaviours, including the Health Behaviour Model [123], Theory of Reasoned Action [124] and Theory of Planned Behaviour [125]. These have been employed to produce quantitative explanations of variation in behaviour, including MMR2 uptake (Ticker [103] using the Theory of Planned Behaviour).

Finally, we note two specific models for vaccination behaviour. Salisbury (cited in Yarwood et al [20]) proposed categorising the reaction as the result of interplay of two fears, that of the vaccine and the vaccine-preventable disease, (Figure 1-12a); this reaction may be a dilemma resolving to either outcome. The ‘SAGE Working Group’ on vaccine hesitancy [126] uses a framework incorporating the concepts in the 3 C’s model (first proposed by WHO EURO Vaccine Communications Working Group), which encompasses a wider context: ‘confidence’ not only in the vaccine but in health professionals and policymakers, ’complacency’ and the ‘convenience’ of vaccination access (Figure 1-12).

Figure 1-12 Models of vaccine behaviour

a] Fear of vaccine vs fear of disease b] Three C's



Yarwood et al, [20] *MacDonald et al* [126]

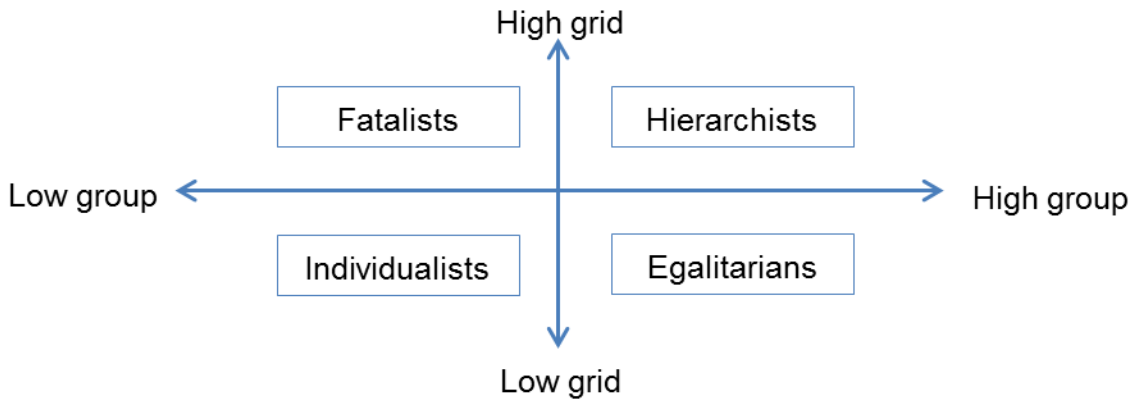
1.2.3.1. Information and evidence used to inform decisions

All these models rely on information and evidence as inputs. Cognitive psychology includes investigation on how external stimuli are processed and the biases that may be introduced (at that stage or during later recall). Psychologists have proposed that individuals employ heuristics in obtaining and assessing evidence, e.g. availability, anchoring and representative heuristics [127]. These can result in biases, as subsets of evidence are used, and the proposal that humans have only bounded rationality [128] . Information from different sources may be regarded as having different quality (e.g. Casiday [87]) and the current status of the individual can affect how evidence is processed (e.g. confirmation bias [129], backfire effect). These processes may thus affect the inputs that are used within an individual's decision process, and correspondingly the choices made in modelling this process.

Risk perception has received considerable attention from researchers; two high-profile models are summarised here. Originally developed by anthropologists, the Theory of Culture [130] allocates individuals to categories arising from a two-dimensional (grid and group) mapping of ways of living (Figure 1-13). The Theory proposes that individual's perception of a specific risk (e.g. vaccination [131]) is framed by the category to which they belong, as their societal context creates "cultural bias". The Psychometric Paradigm of Risk [132]

proposes that perceived risk is largely determined by certain attributes of the hazard itself: novelty and dread (some versions of the model also include social trust i.e. the degree to which the lay population trust the experts and authorities who control the hazard). These attributes have been incorporated into tools that can be used to analyse a given hazard's potential to create "community outrage" due to high perceived risk e.g. Sandman's model of public risk perception (applied to the MMR scare by Burgess et al [72])

Figure 1-13 Theory of Culture "grid-group" mapping of social relations



'Grid' axis: extent to which societal stratification constrains an individual's way of life

'Group' axis: extent to which the individual lives as part of a bonded group with shared choices

Douglas & Widavsky [130]

1.2.3.2. Influence of others on individuals' decisions

Some decision models assume that the decision will be based on a measure of the decisions already taken within the population, without separate consideration of reasoning behind that decision. These models include 'imitation' and 'majority rules', economists have noted that these sequential application of the latter creates a "rational herding" effect [133]. Psychological experiments have shown that the presence of others can affect the decisions that individuals take, with informational and normative influences affecting decisions [134]. SNA has explored the contribution that social network membership can have on influencing individual decisions. This influence may be through the conscious sourcing of information to be used in decision-making (such advice networks have been studied within organisations [135, 136]). Social Contagion Theory proposes that the network ties themselves create communities with shared beliefs and behaviours. This process has been explored for some health behaviours [137-139] and for risk perception [140].

1.2.4. Mathematical models of information-behaviour-infection systems

1.2.4.1. Review of the most relevant mathematical models

The epidemiological literature contains a number of mathematical models considering the interplay of infection dynamics and information (derived directly or indirectly from infection prevalence) that is assumed to influence individuals to adopt specific infection-transmission-related behaviour. A subset of these models are reviewed here, focussing on those whose assumptions are closest to those appropriate for modelling the hypothesised mechanism of the research question.

Thus details are not given for models that do not consider (some) decision-process inputs and the infection process as being transmitted along social network edges. This excludes several models of vaccination where the decision process is modelled by using game theory in which payoffs are calculated based on population level prevalence or on a random sample of population member(s) (using either "perceived" probabilities, functionally linked to actual prevalence, or perfect knowledge of the modelled values) e.g. Bauch et al 2003, Bauch & Earn 2004, Bauch 2005, Bhattacharyya & Bauch 2011, Reluga et al 2006 [141-145]. Some models adjust the evaluated payoff values for perceptions e.g. Voinson et al [146] (adjusts by the agent's current strategy - "confirmation bias"), Oraby et al [147] (weight by strategy's population prevalence – "social norm") . Several game theoretic models, with various assumptions, have found population optimal vaccination uptake is not achieved through individuals maximising their own strategies, and the potential for oscillations about equilibrium points. Other models in this category: Del Valle et al 2005, House 2011, Fu et al 2011, Codeco et al 2007, Shim et al 2012 [148-151].

Similarly, details are not given for models which represent behaviour-change as rewiring of the host social network (inappropriate for representing vaccination) e.g. Gross et al 2006, Epstein et al 2008, Shaw & Schwartz 2008, Zanette & Risau-Gusman 2008,Van Segbroek et al 2010 [152-156].

For simplicity, it is assumed that vaccination is fully protective, conferring lifelong immunity (acknowledging MMR does not take perfectly, with approximately 90% of individuals seroconverting after [28]). Hence models with no immune class are not detailed nor those where the behaviour reduces susceptibility (or infectiousness) rather than conferring immunity (although these could be adapted by considering the extreme case where the

reduction proportion parameter is set to 1, the general results discussed in the papers assume some intermediate value) e.g. Funk et al 2009, Funk et al 2010, Kiss et al 2009, Hatzopoulos et al 2011, Bagnoli et al 2007 [157-161].

For specific relevance to the routine immunisation, details are not given of models which meet the above criteria and use vaccination-choice as their behaviour dynamics, but which act on a timescale closer to reactive vaccination in the face of an outbreak or for reaction to repeated infection seasons. Such models include those by Perisic & Bauch [162, 163] (game theory models using perceived risks based on infection status of immediate social network alters, and find a node degree threshold for infection to escape this 'ring vaccination' process) and seasonal influenza models [164, 165].

Although these excluded models are not detailed here, they provide a wider pool of functional forms which can be adopted should observations indicate they offer a reasonable representation of the decision process. We note some recent game theory models include two functional forms, agents proportionally selecting the optimum self-interest payoff or a strategy calculated from another functional form e.g. Ndeffo Mbah et al 2012 (alternative is imitating a social network alter), Xia & Liu 2013 (a weighted average of all alters' strategies – local "social norm"), Shim et al 2012 (payoff includes incremental total payoff for the population – "altruism") [151, 166, 167].

We also note that the majority of the models assume that the population (and any associated structure) that form the information source is the same as that which is relevant to the spread of infection. Eames [168] (parent and child) and Fukuda et al [165] (assumes a duplex structure on a single population with different networks for payoffs evaluation and infection transmission) are exceptions.

This process has highlighted five models of particular relevance to the research question, all consider the active vaccination 'decision' process and infection dynamics occur sequentially, i.e. all vaccination decisions are made (and nodes are vaccinated according to the final opinion) before the infection is introduced to the network (which is static) and outbreak dynamics investigated. They differ in the decision formulation, how the final decision is identified and the underlying network(s).

These include the models of Ndeffo Mbah et al, and Xia & Liu discussed above [166, 167]. In both models, initial vaccination opinions (pro- or anti-) are randomly allocated and then all nodes update their opinion in parallel according to the formulations described above – a proportion selecting the game theoretic optimum, others adopting the strategy of their alter(s). Updates are repeated until a steady state is reached. Ndeffo Mbah et al compare

results on three networks, Xia and Liu use a network based on empirical data from a high school. Both find that greater use of the imitation-style formulation increases uptake of low-cost vaccines and decreases uptake of high-cost vaccines. The models find results also depend on the network used (Mbah et al) or initial conditions (Xia Liu).

Campbell & Salathé 2013 [169] treat the decision process as the spread of anti-vaccination opinion within an initially fully supportive population, represented by small world networks with varying degrees of rewiring. Nodes are exposed to quanta of anti-vaccination sentiment from alters and from an external information source, they adopt that opinion once a threshold of cumulative exposure is reached. Vaccination occurs once a fixed level of anti-vaccination has been generated. They find that clusters of anti-vaccination form and that there are fewer, larger such clusters when the exposure threshold for adoption is greater than one; these clusters then support larger outbreaks.

Models by Salathé & Bonhoeffer [170] and Eames [168] are related, but use different networks. They each consider families of networks generated from a single network-generation algorithm, Eames using separate parent and child networks. Vaccination opinions are initially allocated randomly, but nodes are randomly selected and change their opinion with probability proportional to the percentage of alters with that opposing view, this selection repeats a fixed number of times before vaccination and infection occurs. Both models find vaccine-rejecters clusters are formed, and that the probability of outbreaks increases as the constant of proportionality in the decision-changing probability increases, and that this effect can produce major outbreaks even in highly-vaccinated populations. Eames finds the further result that the strength of the effect on infection dynamics (from vaccination-decision cluster-formation) is moderated by the amount of overlap between the adult and child networks.

1.2.4.2. Implications for exploration of the research question

Most models reviewed here are treated from a theoretical perspective only, without empirical data to inform functional forms, parameter values or network structure, nor comparison of the predicted results with observations (of infection incidence or behaviour/opinion prevalence).

It is unclear if the included network-generation algorithms create a realistic social network structure, and in the case of childhood vaccination whether a similar algorithm is appropriate for both parents (making the decision) and children (exposed to infection). The assumptions and results of the highlighted models indicate that network structure parameters for both decision-making and infection network should be considered, such as node degree,

measures associated with network algorithm (e.g. clustering), and overlap between decision-making and infection networks. Specifically in relation to the realism of social structure for childhood vaccination and infection, in Eames [168] (the only childhood vaccination model highlighted) the adult-child networks are in 1-1 node correspondence; excluding situations where parents have more than one child or sample opinions from adults without a child in the infection network (e.g. their child's grandparents).

The criteria used to identify the 'final' opinion (fixed number of sequential decisions [168, 170], fixed vaccination proportions [169] or steady state ([166, 167]) may create artefacts in the level or distribution of vaccination-acceptance. Campbell & Salathé acknowledge that the fixed vaccination end-point creates a necessary relationship between the size and numbers of clusters formed. Specific to Eames there is the consideration of whether artefacts are formed by the forced "balancing" of opinion-changes (to maintain constant vaccination-support within the population), especially if this contributes to the opinion-clustering by increasing the proportion of intra-dyad agreement.

It may be possible to incorporate some aspects of other decision-process models, observations from the existing body of MMR research, or empirical data into the selection of the function form(s) used for the decision-process. In terms of incorporating the existing social science decision-models, well-established, validated non-vaccine-specific measurement tools exist for a number of the factors or vaccine-specific tools have been piloted in other studies (e.g. "Immunisation Beliefs and Intentions Measure" [171]). However, it is acknowledged that several of these have a substantial respondent burden (e.g. 58 measures in Tickner's tool [171])

Finally, in addition to the investigation of the theoretical assumptions included in the construction of the mathematical model to be used in the exploration of the research question, it requires parameterisation specific to the MMR vaccination-decision and measles infection dynamics within the UK.

2. Small Area Statistical Analysis

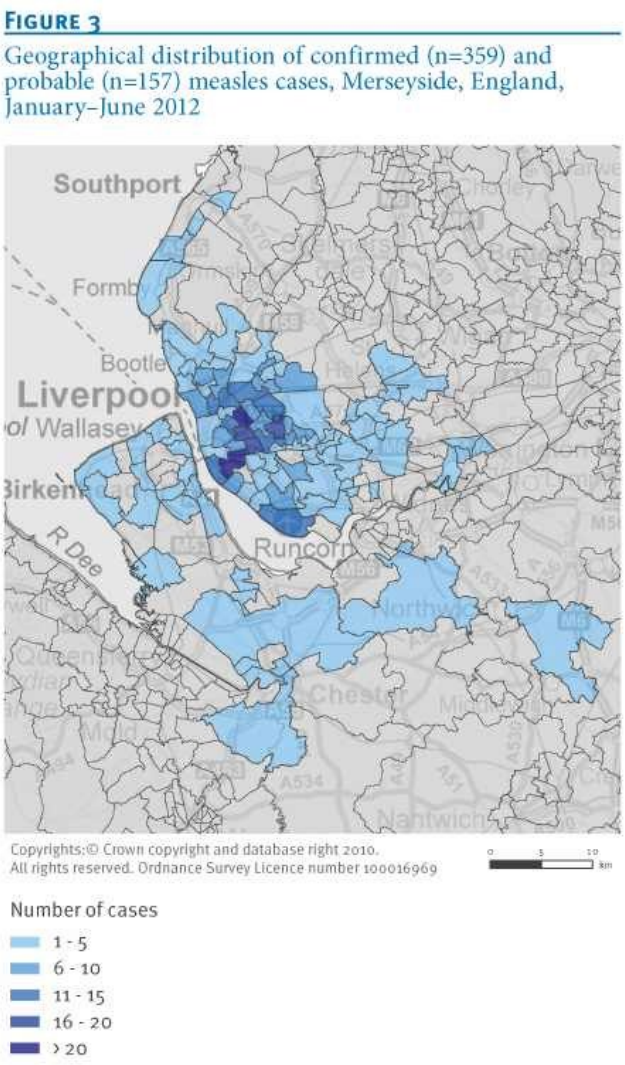
2.1. Motivation

The spatial analysis of surveillance data (Chapter 1) reveals variation in MMR1 uptake at a PCT scale, which is the smallest geographic scale routinely reported in surveillance data. From these data it is not possible to determine if the uptake is also non-homogenous at geographical scales below this. The level of vaccine uptake acts both to determine the supply of susceptible individuals in a locality, and the extent of protection provided by the herd immunity effect. The ability for herd immunity to contain the potential spread from an index case to an "outbreak" is a process which acts at a spatial level below PCT regions. Hence an examination of vaccine uptake on a small geographical scale is of interest to gain a deeper understanding of the outbreak risk.

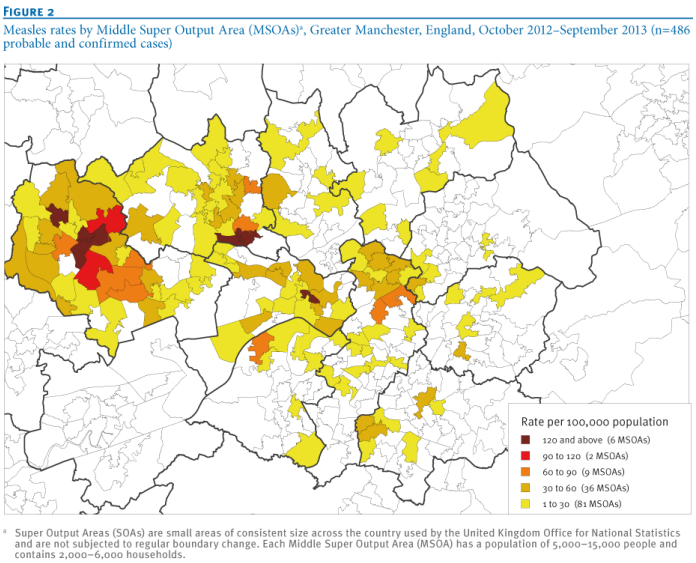
Mathematically, uneven uptake distribution within a PCT may create local geographies with uptake below the PCT's mean. Such regions of under-vaccination have an increased risk that the presence of a measles case results in onward transmission in this local area (compared with an otherwise identical region). Conversely other local geographies may have above-average uptake and are associated with a correspondingly lower risk of onward transmission. It is noted that it is on this spatial scale (below PCT) that several measles outbreaks in the UK have operated (e.g. [172, 173]). Furthermore geographical analysis of two more-widespread outbreaks (Merseyside 2012 and Manchester 2012-3) also reveals uneven case distribution. Both these outbreaks numbered hundreds of infections, spread across more than one PCT, but the incidence measured on smaller geographies within the defined outbreak area - middle super-output area for Manchester [174], ward for Merseyside [175] - demonstrate this non-homogeneity (Figure 2-1).

Figure 2-1 Geographic distribution of cases in the Merseyside and Manchester outbreaks

a] Geographical distribution of confirmed (n=359) and probable (n=157) measles cases, Merseyside, England, January-June 2012



b] Measles rates by Middle Super Output Area (MSOAs)a, Greater Manchester, England, October 2012–September 2013 (n=486 probable and confirmed cases)



a] SOURCE: Vivancos et al figure 3 [175]
b] SOURCE Pegorie et al figure 2 [174] reproduced under Creative Commons Attribution (CC BY)

Such incidence patterns will be affected by local variation in factors affecting transmission (e.g. contact rates), case reporting, treatment and outbreak response activities, in addition to any underlying under-vaccination spatial heterogeneity. Thus, local variation in incidence is insufficient to confirm that non-homogeneous vaccine uptake is present on spatial scales not reported in surveillance data. So, in order to determine if small scale variation in vaccine uptake is observed in practice, with associated outbreak risk implications, an analysis of uptake data with suitable, additional spatial granularity is required.

Some of the factors identified in qualitative and quantitative studies (Chapter 1) as having possible association with MMR1 uptake vary on small spatial scales. Studies have investigated a potential correlation between MMR vaccination status and an area-defined characteristic, deprivation, albeit with inconsistent conclusions [19, 69, 98, 176]. PCT regions can exhibit social diversity with population profiles of demographic characteristics varying on smaller spatial scales, including those considered in these studies, such as ethnicity [19, 61, 99, 176], working status [19, 83, 176] and income [85, 93]. We therefore secondly consider the relationships between these factors and MMR1 uptake, as measured at a population level for sub-PCT areas, and the extent to which any variation in these factors can explain any observed spatial variation in MMR1 uptake.

2.2. Methods

Data for statistical analysis were sourced from existing surveillance and census sources. The period of MMR1 uptake analysed was April 2011- March 2012. The time period used for data collection for other variables was selected to be as contemporaneous as available for the preferred source (as detailed below). All data used was supplied at a spatial granularity at least as great as the geographical unit used for analysis (Ward, see below).

2.2.1. Vaccine uptake data

Surveillance data for MMR1 uptake in England are prepared via the COVER system [10] . The published information is the uptake for the specified period, i.e. proportion of children who had their 2nd birthday during the period who were vaccinated, by PCT, region and nation. Data for smaller geographic units are not held by Public Health England (PHE), so are not available for analysis.

The COVER data is derived from information supplied by PCTs taken from their Child Health Information System (CHIS), which in turn is supplied by GP practices; this detailed

information is retained by the PCT and not supplied to PHE [177] Hence we examined the MMR1 uptake data from CHIS as held by PCTs, as this is directly compatible with COVER surveillance data, but can be summarised by smaller geographical units than the figures published by PHE.

2.2.2. Geographical units and ethical considerations

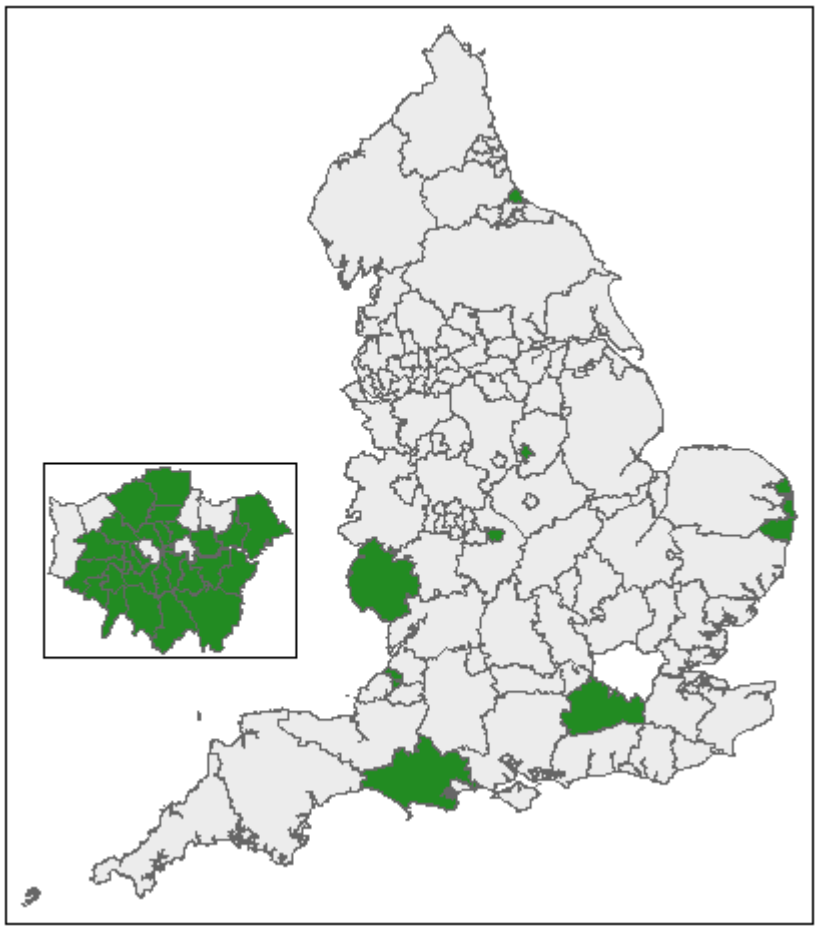
The geographic unit used for the small scale analysis is a Ward, defining this as Administrative Ward with 2011 boundary definitions. The boundaries for these regions are chosen by ONS to be temporally stable (Ward populations are not equal, and will fluctuate with time) [178]. They form part of the nested hierarchy of statistical measurement geography used by ONS. Specifically 98% of wards are coterminous with Lower Layer Super Output Areas (LSOA) which are elements of the Super Output Area geography used in datasets from Census 2011. Super Output Areas and Wards all nest within Local Authority areas, the majority of which are themselves conterminous with PCTs (as defined in March 2011).

Ward-level data sourced from ONS may be subject to disclosure controls, due to identification issues inherent with combinations of geographical units and variables which produce small cell numbers [179]. Similarly, we followed guidelines produced by the Association of Public Health Observatories [180] regarding cell threshold numbers that would trigger requesting disclosure measures to be applied to CHIS-sourced information from PCTs.

The request for MMR1 CHIS data summarised by Ward was conducted under NHS REC 11/EE/0343 ethical approval. The PCTs approached to supply MMR1 uptake figures by Ward were those of particular epidemiological interest. They were defined as those PCTs which reported MMR1 uptake figures in the lowest decile for England in any of the following COVER reports (the most recent published prior to Ethical Approval submission): annual reports for 2008-9, 2009-2010 and quarterly reports for 2010-2011 [35, 36, 181-184]). (The PCT names and boundaries used are as were in operation in March 2011). 33 PCTs fall into this definition (listed in Chapter 4, Table 4-3) and were included under the REC 11/EE/0343 ethical approval. The majority of these PCT are located in Greater London (Figure 2-2).

A second tier of approval was required for data release itself, and was only secured for one PCT – Great Yarmouth & Waveney.

Figure 2-2 Map of PCTs of greatest epidemiological interest



2.2.3. Variables considered for inclusion

Factors considered for inclusion in the analysis, as having a possible association with MMR1 uptake, were selected from those used in previous qualitative and quantitative studies on MMR1 uptake in the UK (Chapter 1). Analysis of these studies grouped these into three areas as follows:

- personal vaccination/diseases beliefs: general concerns about side-effects and vaccine safety, personal awareness of serious reactions to attributed to vaccination, perceived vaccine efficacy, perceived seriousness of measles
- demographics of the parent: parent age, educational achievement, working status, income, deprivation, family structure (size and number of parents present), ethnicity
- parent-community interactions: trust in HCP, distrust of authority and government, non-professional advice (from friends, family and the media)

Additionally factors that have the potential to affect vaccination access logistics have been considered:

- registration with a GP, urban/rural location, demand on childhood vaccination services, appointment communications and proficiency in English. (Logistics has also been proposed as an underlying mechanism behind some previously identified variables e.g. parental working status, ethnicity)

In the following analysis, we will refer to each of these potentially associated factors, measured at a ward-level as a ward-characteristic.

2.2.4. Data sources

The 2011 UK Census is used as the data source for the following ward-characteristics (directly or by proxy):

- age of parents: using the 'Age of Family Reference Person (FRP)', obtained from the derived variables ''Age' (AGE) and 'Family Reference Person'(FRPPUK11) [185], cross-tabulated with 'Youngest dependent child in family'(DPCFAMUK11) [186], reported for all families.
- educational achievement: using the derived variable "Highest Level of Qualification" (HLQPUK11) [185, 187], reported for all usual residents aged 16 and over.
- working status: using the derived variable "Economic Activity" (ECOPUK11) [185, 188], reported for all usual residents aged 16 and over.
- family structure – family size using the derived variable 'Dependent children' (DCHPUK11) [185, 189], reported for all families.
- family structure – parents present: using the two 'Lone parent family' categories of the derived variable 'Family Type' (FMTFAMUK11) [185] cross-tabulated with 'Youngest dependent child in family'(DPCFAMUK11) [189] , reported for all families.
- ethnicity: using derived variable "Ethnic Group" (ETHNICID) [185, 190], reported for all usual residents.

- demand on childhood vaccination services : via a proxy variable, the percentage of the population aged 0-4, calculated using the derived variable “Age” (AGE) [185, 191] , reported for all usual residents
- proficiency in English: using the “English as Main Language” category of the “Proficiency in English” variable (MAINLANGPRF11) [192] , reported for all usual residents aged 3 and over.

The census offers the benefit of being near contemporaneous with the start of the CHIS data period (1 April 2011 – 31 March 2012), as the data was collected on 27 March 2011 [193] . The census methodology also provides population data (minimising sampling error), and data is published at Ward level for all the variables used.

The 2011 census is also the source for the population data used in the calculation of population density, which is used as a proxy for urban/rural location:

- urban/rural location: via the proxy variable “Population Density (people per hectare)” [194]

Again this is published at Ward level, using the Ward 2011 boundaries.

Deprivation data is taken from the English Indices of Deprivation (IOD), produced by ONS, described as “the official measure of relative deprivation for small areas” ([195] p1). The 2015 IOD is used, as most of the data used in generating the indices relates to the tax year 2012-13, which is closer to the CHIS data period than the data used in the previous release of the IOD (2010 IOD) [195, 196].

The IOD includes measures for each of seven domains which represent different aspects of deprivation: “Income Deprivation”, “Employment Deprivation”, “Education, Skills and Training Deprivation”, “Health Deprivation and Disability”, “Crime”, “Barriers to Housing and Services” and “Living Environment Deprivation”. These measures are each calculated using a basket of indicators. Given the multivariate nature of the deprivation data, a supplementary analysis is included to determine if dimension reduction can be usefully undertaken to produce a more parsimonious measure of deprivation (see §2.2.6.2). This supplementary analysis also considers the published overall Index of Multiple Deprivation (IMD), which combines values from all seven domains, and - given our objective of assessing the association with variation in MMR1 uptake - the Income Deprivation Affecting Children Index (IDACI), which is a supplementary IOD index.

The Quality and Outcomes Framework (QOF) was identified as a source of variables that might be used as proxy for the ward-characteristics relating to GP appointments and access. The QOF is reported by GP practice and QOF 2011-12 [197] is exactly contemporaneous with the CHIS data period. The following QOF 2011-12 indicators were selected as relevant proxies for this study:

- Child Health Surveillance (CH501) [198]
- Patient Experience (PE01) [199]

QOF 2011-12 for Great Yarmouth & Waveney PCT includes data from 27 practices. However, all 27 practices reported the same values for both the selected indicators. Variation by Ward in these indicators may exist by Ward, but any estimates of Ward values calculated from these practice-level data will not show variation. Given the objective of the current analysis is the examination of small spatial scale variation (using Wards as the spatial unit) these ward-characteristics have been excluded.

For the remaining ward-characteristics listed above (§2.2.3) no suitable data sources have been identified. None of the potential sources identified publish data on the required spatial scale, and secondary analyses to obtain Ward data are not feasible due to original survey methodology (e.g. DH/COI CITS [21], which reports on most of the factors grouped under Personal Beliefs and the use of non HCP information sources), issues of patient confidentiality (e.g. vaccine-attributed adverse events [27]) or the likely extent of disclosure measures.

Where possible the demographic data obtained from the census is extracted for the sub-group most relevant to routine childhood immunisation (parents of, or families containing, dependent preschool children) in preference to the full resident population. The ability to apply this restriction is limited by the sub-groups that are published at ward level for each measurement. Also where a restriction results in zero-value cells a less stringent restriction is applied (zeros may remain present in ‘all residents’ data). The sub-groups used for the base populations used are given in Table 2-1.

Table 2-1 Summary of ward-characteristics included in the analysis

Ward-characteristic	Measured variable (source)	Units: Base population used
Age of Parent	Age of FRP (Census 2011)	Family: with youngest dependent child age 0-4 years
Educational achievement	Highest Level of Qualification (Census 2011)	Persons: age 16 and over in family with dependent child
Working status	Economic Activity (Census 2011)	Persons: age 16 and over
Deprivation	IMD, IOD domains, IDACI subdomain (IOD 2015)	n/a
Family size	Dependent Children (Census 2011)	Family: all
Parents present	Family Type (Census 2011)	Family: with youngest dependent child age 0-9 years
Ethnicity	Ethnic Group (Census 2011)	Persons: all usual residents
Urban/rural location	Population Density (Census 2011)	n/a
Demand on childhood vaccination services	Age (Census 2011)	Persons: all usual residents
Proficiency in English	English as Main Language (Census 2011)	Persons: age 3 and over

2.2.5. Data preparation for comparison at ward level

The IOD data are published for LSOA, which are more spatially granular than Wards [195]. Thus, these data are first converted to Ward-level data using the process recommended by ONS [200] . In summary, this process first identifies the LSOAs contained in each Ward and then produces a population-weighted average score for the Ward. (No additional assumptions were required in this process as all of LSOA and Wards in the study region (Great Yarmouth & Waveney PCT) are coterminous).

To identify the LSOA for each Ward, tables allocating census Output Areas (OA) to LSOA [201] and allocating OA to [202] were used in a two-step allocation of LSOA to Wards (as no direct allocation of LSOA with Wards is published).

Figure 2-3 Example geography hierarchy within one ward (Pakefield E05007243)

LSAO Output Areas

Source: Edit from NeSS Map Viewer, used under Open Government Licence v3.0, http://neighbourhood.statistics.gov.uk/dissemination/LeadBoundaryViewer.do?xW=1280&xH=800

The population to be used in the weighted-average calculations is specified for each domain and sub-domain of the IOD 2015 [203]. It is defined as the "population at risk" for the specified measure of deprivation, e.g. a working age restriction is used in calculating populations for use with the Employment deprivation domain, and the majority are derived from ONS mid-2012 population data.

2.2.6. Data analysis

The data analysis was performed using SPSS software (version 22) [204] with one exception (detailed below §2.2.6.1). The variation in MMR1 uptake by Ward was quantified and compared with the variation observed at the larger spatial scale of PCTs (as used by COVER data).

For each of the ward-characteristics, an exploratory analysis was performed, including variation by Ward and its relationship with MMR1 uptake (categorisation of correlation strength taken from Evans [205] e.g “weak” for correlation coefficient in [0.2,0.4)).

These exploratory analyses are used to inform a regression analysis, with MMR1 uptake as the dependent variable. The regression analysis uses a GLM with a logit link function, with the dependent variable for each data-point regarded as the results of a Bernoulli trial for each child in the ward (trial success defined as being vaccinated). Both univariate and multivariable regression analyses are conducted.

2.2.6.1. Handling compositional datasets

It is noted that for several of the ward-characteristics, the data are derived from categorical variables measured by individual (person or family). So, when collated to form a Ward data-point, the factor is represented by a set of percentages for each level in the original categorical measurement, i.e. compositional data. The ward-characteristics affected are “Age of Parent”, “Educational Achievement”, “Working Status”, “Family Size” and “Ethnicity”.

The compositional datasets, with each Ward’s data-point of the form $x = (x_1, \dots, x_D)$ with $x_i > 0$ and $\sum_i^D x_D = 1$, present two complications in a regression: they contain a (summation) constraint (reducing the degrees of freedom), and the number of variables used to describe the factor is not a single categorical variable (with D levels) but D variables. Hence a dimension reduction analysis is performed on the set of compositional data for each of these ward-characteristics and (as with the original compositional data) an exploratory analysis completed prior to the regression.

Principal component analysis is a not suitable dimension reduction procedure for compositional data [206] . The compositional data dimension reduction was performed using procedure described by Filmoser et al [207] and codified in the robCompositions package (version 2.0, 2016) in R (version 3.2.4 Revised, 2016) [208, 209]. Two log ratio transformations of compositional data are used in this process: the centred log ratio, $x = (x_1, \dots, x_D) \rightarrow y = (y_1, \dots, y_D) \in \mathbb{R}^D$, and the isometric log ratio, $x = (x_1, \dots, x_D) \rightarrow z = (z_1, \dots, z_{D-1}) \in \mathbb{R}^{D-1}$

where $y_i = \ln\left(\frac{x_i}{\sqrt[D]{\prod_{j=1}^{D} x_j}}\right)$ and $z_i = \sqrt{\frac{i}{i+1}} \ln\left(\frac{\sqrt[i]{\prod_{j=1}^{i} x_j}}{x_{i+1}}\right)$

The compositional data undergo an isometric log ratio transformation to enable principal component analysis, producing (D-1) components, the results are then back-transformed to the centred log ratio space to facilitate interpretation in terms of the original variables [207].

It is noted that log ratio transformations fail when a zero-value cell is present. In such cases the zero values are replaced with imputed strictly positive values using an assumption that these values represent percentages below a “detection limit” of 1 person (although it is known that these are indeed records of zero people).

2.2.6.2. Supplementary analysis of the multivariate deprivation data

The IOD contains sets of data each measuring relative deprivation for seven domains. Given the number of data points in the main analysis (40 Wards), this multivariate data may not be appropriate; hence a more parsimonious measure of deprivation is considered.

Within the IOD data release a single measure of relative deprivation is included, the Index of Multiple Deprivation (IMD). The data for the seven domains are combined in a fixed manner (for all geographies within England), for the IOD 2015 this is as follows:

- Income Deprivation (22.5%)
- Employment Deprivation (22.5%)
- Education, Skills and Training Deprivation (13.5%)
- Health Deprivation and Disability (13.5%)
- Crime (9.3%)
- Barriers to Housing and Services (9.3%)
- Living Environment Deprivation (9.3%)

This fixed calculation may not deliver an appropriate set of values to represent the variation in deprivation across the study region. Additionally, it is thought that the child-related income deprivation measurement (IDACI) may be more appropriate to in a study of factors affecting child vaccination uptake than the (all) income deprivation measure included in the single score and standard set of domains.

Hence we produce two alternative deprivation measures generated from a dimension reduction analysis of domain data, one using the income deprivation domain and one where this is replaced with child-related income deprivation. As above, an exploratory analysis is

completed for the resultant components, and they are compared with the standard single deprivation score prior to the regression analysis.

The dimension reduction for the deprivation domains is performed using non-linear (categorical) principal component analysis (CATPCA) [297] as the domain scores include both scale and ordinal data variables.

2.3. Results

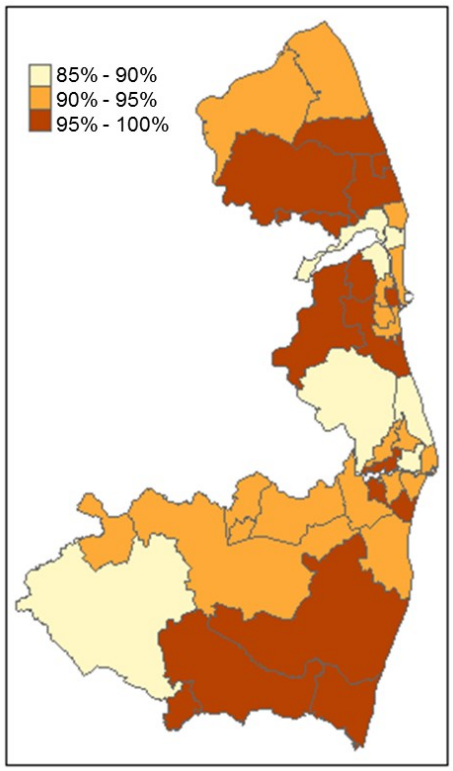
CHIS data for MMR1 were supplied by Great Yarmouth & Waveney PCT [177] and contained no merged wards.

Great Yarmouth & Waveney PCT (as defined in March 2011) is located in East Anglia in England. It is coterminous with the Local Authorities of Great Yarmouth (in Norfolk) and Waveney (in Suffolk). The total population is 212,531 [211]. It contains 40 Wards and 134 LSOA. Ward populations range from 2,150 to 8,681. The PCT-level MMR1 uptake in 2011-12 was 92.7% (of 2,410 children) [38].

2.3.1. Variation in MMR1 uptake

Within Great Yarmouth & Waveney PCT, ward MMR1 uptake ranges from 85.9% to 100.0%, with mean 93.4% and variance 15.8 (n=40). A hypothesis that these ward data-points are all drawn from a common distribution is not supported by the results of an appropriate test of proportions (Fisher Exact, p=0.40 : mean from 10 multiple Monte Carlo estimates each based on 100,000 sampled tables). The Marascuilo procedure [212] is performed on all pairwise combinations of wards, to check if this non-homogeneity is caused by specific ward(s) being significantly different to the others. No significant results are found (at 0.05).

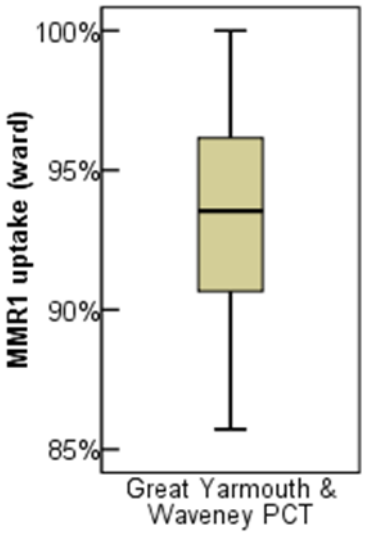
Figure 2-4 Map of MMR1 uptake by ward in Great Yarmouth & Waveney PCT 2011-12



MMR1 uptake data from Raworth [177]

Figure 2-5 Box plots of MMR1 uptake at different nested spatial granularities

a] Great Yarmouth & Waveney PCT by Ward

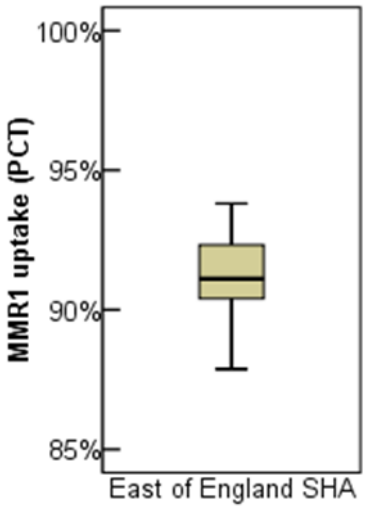


Mean = 93.4%

Variance = 15.8

(n=40)

b] East of England SHA by PCT

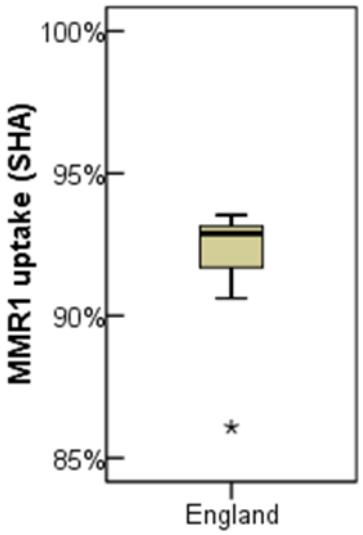


Mean = 91.4%

Variance = 2.8

(n=13)

c] England by SHA



Mean = 91.8%

Variance = 5.5

(n=9)

Raworth, NHS Info Centre [38, 177]

The variance observed within ward-level MMR1 uptake within Great Yarmouth & Waveney PCT is greater than that observed for PCT-level uptake within East of England SHA (which contains Great Yarmouth & Waveney PCT) and for SHA-level uptake within England (Figure 2-5).

2.3.2. Exploratory analysis of ward-characteristics

The association with MMR1 uptake for each ward-characteristic analysed uses a Spearman correlation throughout. It has been used due to the number of ward data-points available (n=40 for all analyses) and for consistency given several variables do not meet the conditions for use of Pearson correlation.

Each ward-characteristic is being considered individually at this stage (to inform a later univariate and multivariable regression analysis), so we do not adjust the p-values for multiple significance testing.

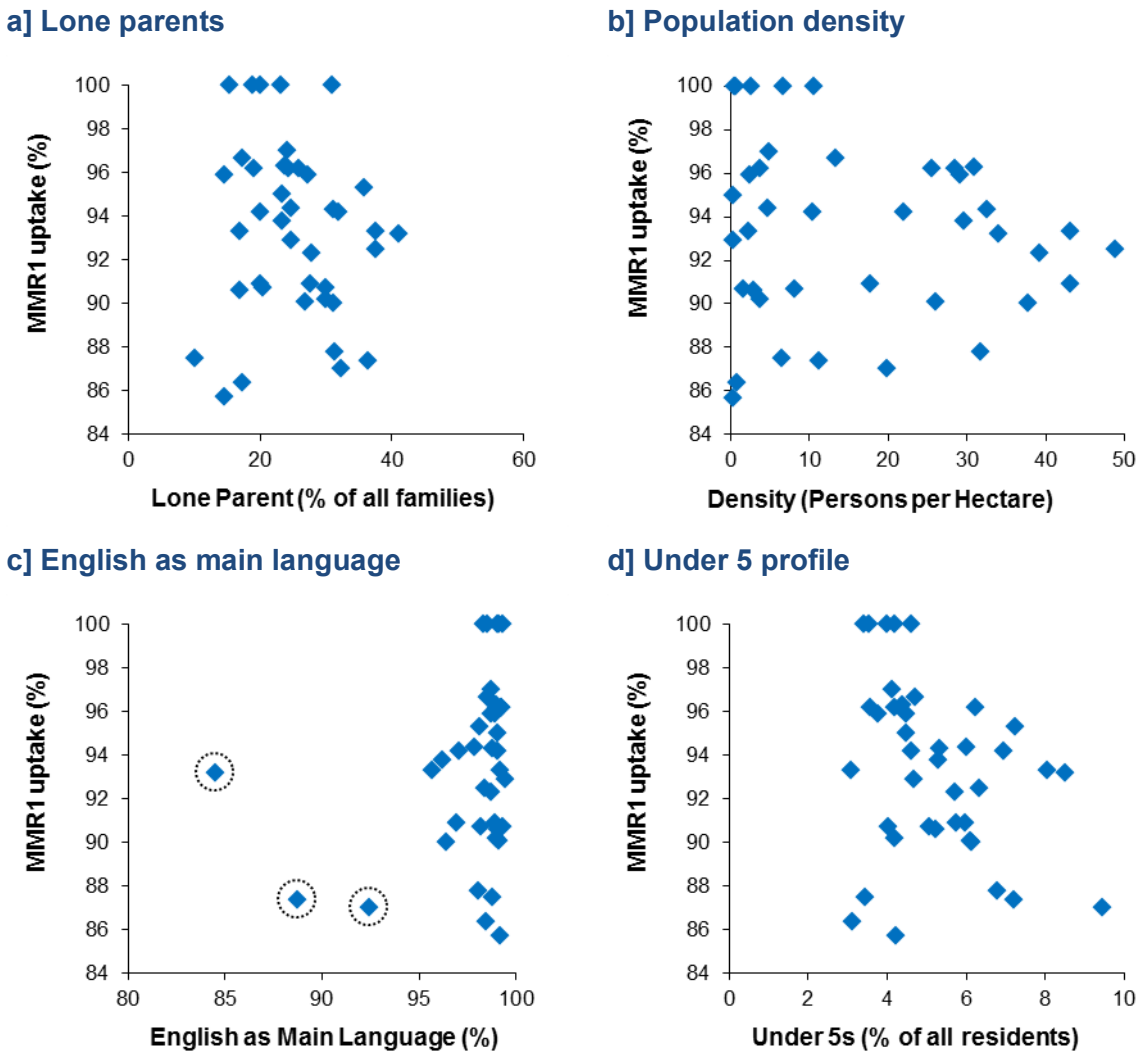
2.3.2.1. Lone parents, population density, English proficiency and population aged under 5 years

There is no evidence for a correlation between either the proportion of families with a lone parent or the population density and the uptake of MMR1 by ward (Spearman's rho = -0.178, p>0.25 and Spearman's rho = -0.077, p>0.6 respectively).

The proportion of people with English as their main language has a weak positive correlation with MMR1 uptake (Spearman's rho = 0.244, p>0.1). However three wards (circled in Figure 2-6) are outliers – these all have below average use of English as main language and below average MMR1 uptake; removing these data-points from the analysis removes the weak correlation (Spearman's rho = -0.033, p>0.9, n=37).

The proportion of the population aged under 5 has a weak negative correlation with the uptake of MMR1 (Spearman's rho = -0.319, p=0.045).

Figure 2-6 MMR1 uptake vs single-variable ward-characteristics



(points considered as outliers circled to avoid ambiguity)

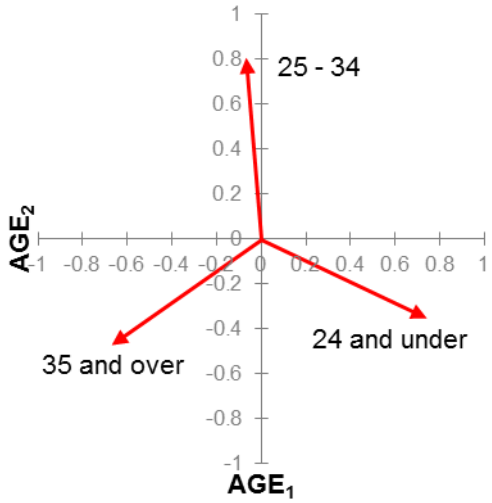
2.3.2.2. Ward-characteristics with compositional data

2.3.2.2.1. Age of parent

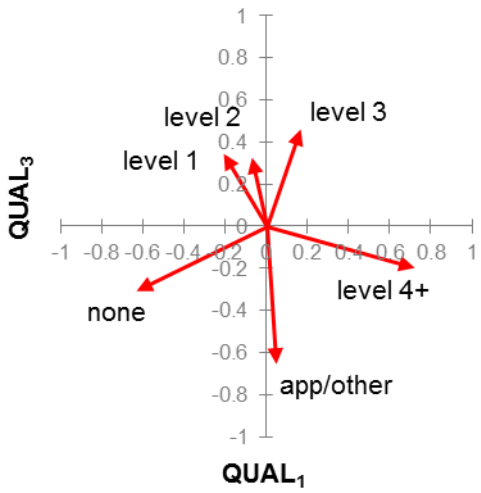
The data on the age of the FRP is compositional data with three categories: age 24 years and under, age 25-34 years, age 35 years and over. We perform a compositional data dimension reduction, obtaining two principal components. A principal component weighting plot is shown in Figure 2-7a; the first principal component can be interpreted as (approximately) higher values for wards with higher proportions of people in the youngest age-group (24 and under) and lower values for wards with higher proportions of people in

Figure 2-7 Principal components of compositional data ward-characteristics

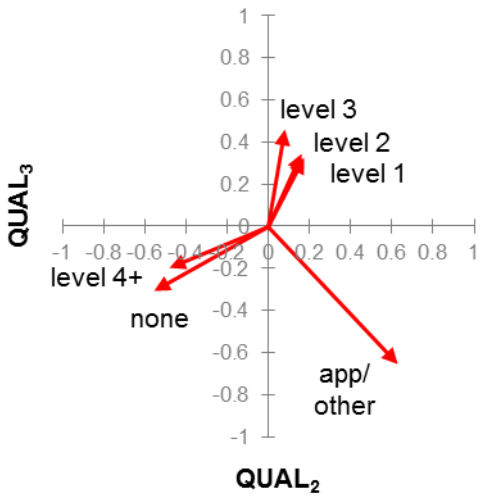
a] Age of parent



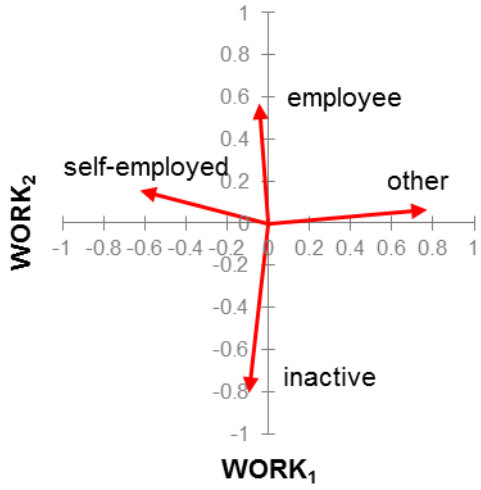
b] Highest qualification: 1st and 3rd



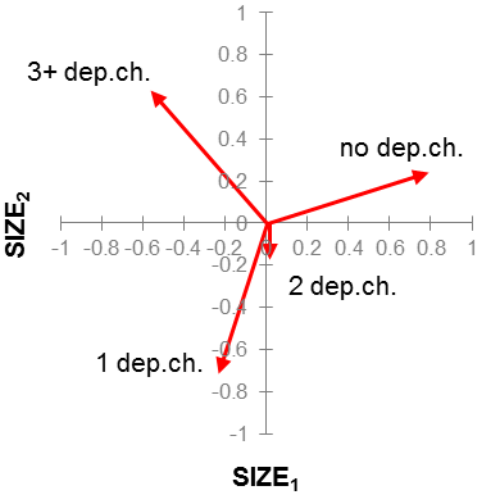
c] Highest qualification: 2nd and 3rd



d] Working status



e] Family size



f] Ethnicity

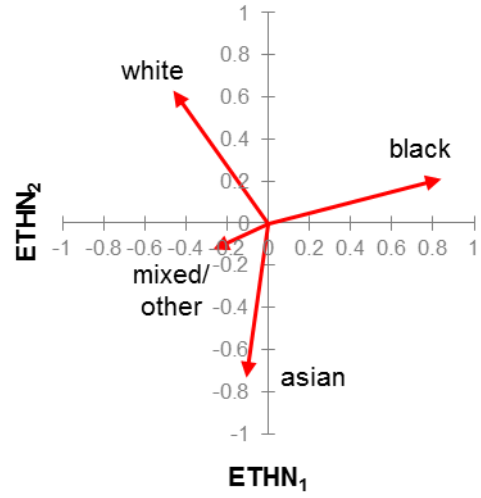
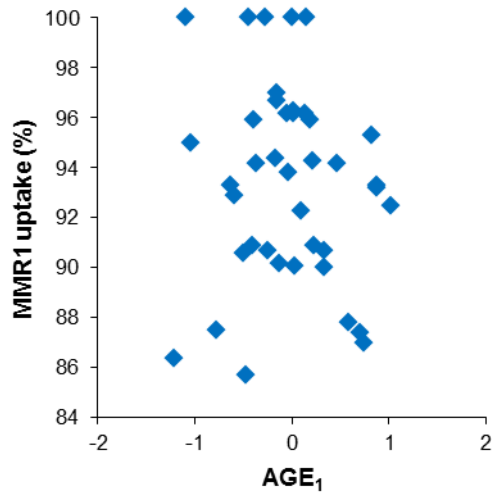
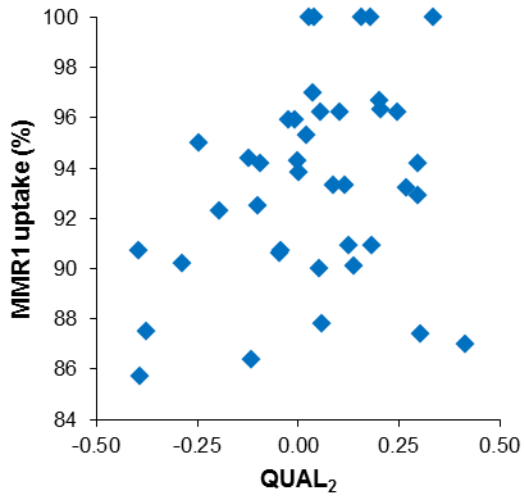


Figure 2-8 MMR1 uptake vs selected principal components of compositional data ward-characteristics

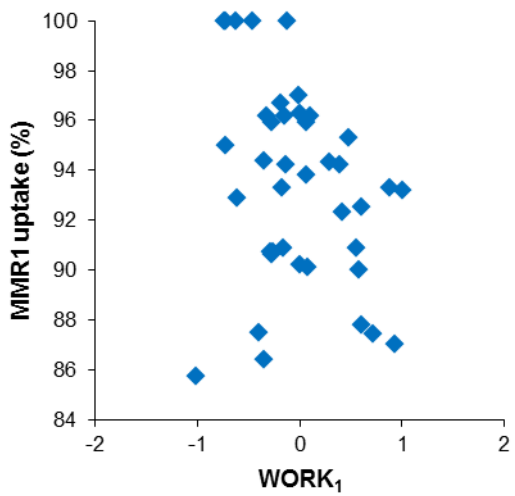
a] Age of parent 1st (AGE_1)



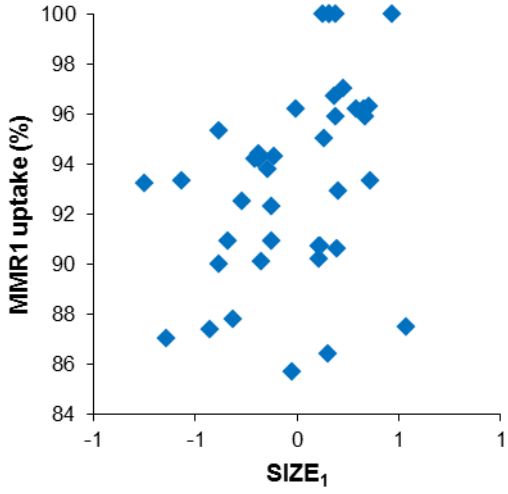
b] Highest qualification 2nd ($QUAL_2$)



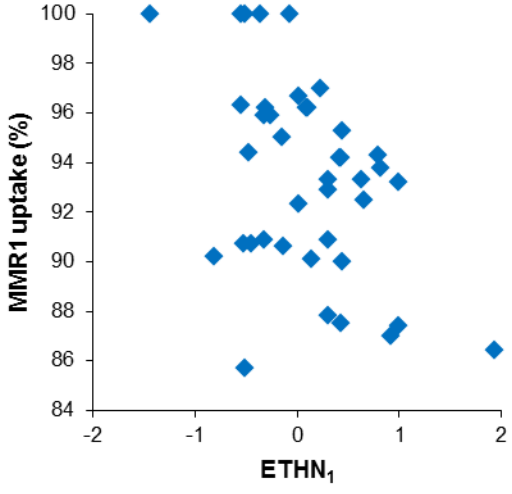
c] Working status 1st ($WORK_1$)



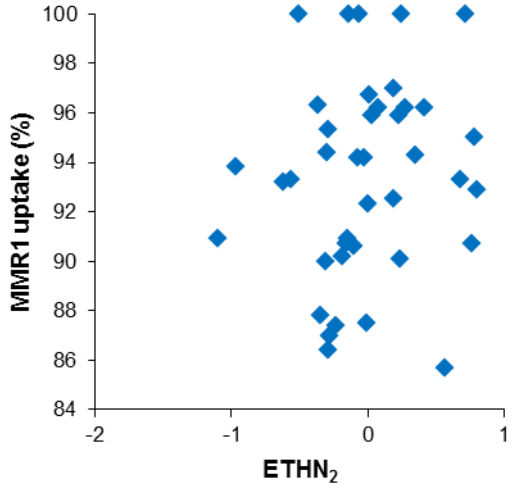
d] Family size 1st ($SIZE_1$)



e] Ethnicity 1st ($ETHN_1$)



f] Ethnicity 2nd ($ETHN_2$)



the oldest age-group (35 and older). It is appropriate to consider a 1 dimensional solution as the Variance Accounted For (VAF) for the first principal component (AGE_1) is 89.2%. There is no correlation observed between this component (AGE_1) and MMR1 uptake (Spearman's rho = -0.042, p>0.75).

2.3.2.2.1. Educational achievement: highest qualification

The data on the highest qualification achieved is compositional data with six categories: no qualifications, level 1 qualification, level 2 qualification, level 3 qualification, level 4 (or above) qualification, other qualifications including apprenticeship. The qualification levels are typified by the following examples: degree (level 4+), 2 or more A levels (level 3), 5 or more GCSE at grade A*-C (level 2), 1 or more GCSE (level 1). Five principal components are obtained from the compositional principal component analysis. Examination of the scree plot and VAF values indicates that a 3 dimensional solution is appropriate, which has a total VAF 95.7% (with individual components' VAF at 63.0%, 19.6%, 13.2%). The first three principal components are shown in Figure 2-7b&c. An approximate interpretation of these three components is:

$QUAL_1$: higher values = wards with higher proportions of level 4+ qualifications;
lower values = wards with higher proportions of those with no qualifications

$QUAL_2$: higher values = wards with higher proportions of "other" qualifications;
lower values = wards with higher proportions of people with either no qualifications or level 4+ qualifications

$QUAL_3$: higher values = wards with higher proportions of those with qualifications at levels 1, 2 or3;
lower values = wards with higher proportions of "other" qualifications.

$QUAL_2$ has the strongest evidence of these components for a linear relationship with MMR1 uptake (Figure 2-8b), but this is a weak positive correlation (Spearman's rho = 0.255, p>0.1). The other components have no evidence of a correlation with MMR1 uptake (absolute values of Spearman's rho <0.2 with p>0.3).

2.3.2.2.2. Working status

The data on working status is compositional data with four categories: employee, self-employed, other economic activity, no economic activity. (Splitting the employee

category into full-time and part-time was considered, but did not substantially affect the following analysis, so in the interests of parsimony the single “employee” variable is retained). The compositional principal component analysis yields three principal components, and the scree plot indicates that the 2-dimensional solution is appropriate (VAF = 94.0%). From Figure 2-7d we obtain a clear interpretation of the principal components:

WORK$_1$: higher values = wards with higher proportions of “other” economic activity;
lower values = wards with higher proportions of the self-employed

WORK$_2$: higher values = wards with higher proportions of employees;
lower values = wards with higher proportions of the economically inactive

There is no evidence that either component has a linear relationship with MMR1 uptake (absolute values of Spearman’s rho <0.2 with p>0.4)

2.3.2.2.3. Family size

The family size data is compositional data with four categories: no dependent children in family, one dependent child in family, two dependent children in family, three or more dependent children in family. The compositional principal component analysis yields three components, and the scree plot confirms that the 2-dimensional solution is appropriate (VAF= 90.2%). Interpretation of the components (Figure 2-7e) is less clear-cut than for most of the ward-characteristics:

SIZE$_1$: higher values = wards with higher proportions of families without dependent children;
lower values = primarily higher proportions of families with 3+ dependent children

SIZE$_2$: higher values = primarily wards with higher proportions of families with 3+ dependent children;
lower values = primarily wards with higher proportions of families with 1 dependent child

The first component (*SIZE*$_1$) has a weak correlation with MMR1 uptake (Spearman’s rho = 0.328 with p= 0.039), but there is no evidence the second component (*SIZE*$_2$) has a linear relationship with MMR1 uptake (absolute value of Spearman’s rho <0.2 with p>0.4).

2.3.2.2.4. Ethnicity

Ethnicity data is compositional data with 4 categories: White, Asian (includes British Asian), Black (includes British Black), Other/Mixed. Three principal components are obtained from the compositional principal analysis and the scree plot indicates that a 2-dimensional solution is appropriate (VAF = 92.4%). Using weighting plot (Figure 2-7f) an approximate interpretation of the retained principal components is:

$ETHN_1$: higher values = wards with higher proportions of black residents;
lower values = wards with higher proportions of non-black residents

$ETHN_2$: higher values = primarily wards with higher proportions of white residents;
lower values = primarily wards with higher proportions of Asian residents

There is evidence that both components (Figure 2-8e&f) have a weak correlation with MMR1 uptake ($ETHN_1$ Spearman's rho = -0.337 with p=0.033, $ETHN_2$ Spearman's rho = 0.280 with p>0.05)

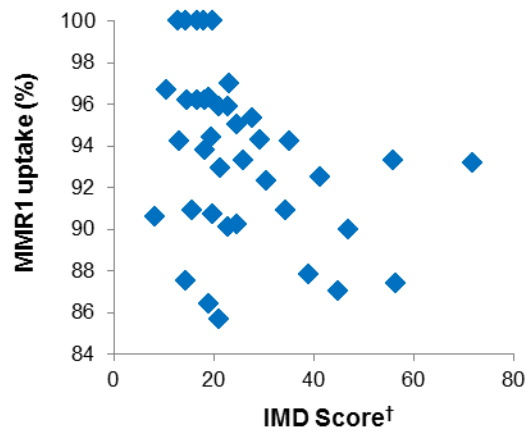
2.3.2.3. Deprivation

For both the IMD and the domains of deprivation, a higher score represents a higher degree of deprivation. The IMD is moderately correlated with MMR1 uptake (Spearman's rho = -0.416 with p=0.008).

There is evidence that four of the six non-income-related domains of deprivation have weak negative correlations with MMR1 uptake (Figure 2-9) Employment and Environment (Spearman's rho = -0.326 with p=0.40), Health (Spearman's rho = -0.322 with p=0.042), Crime (Spearman's rho = -0.244 with p>0.1).

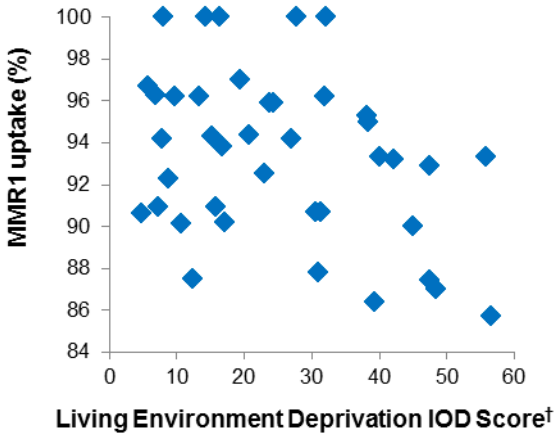
Figure 2-9 MMR1 uptake vs Selected measures of deprivation

a] IMD

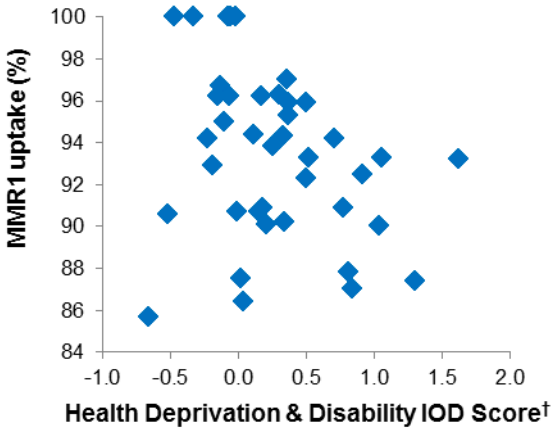


b] Employment domain

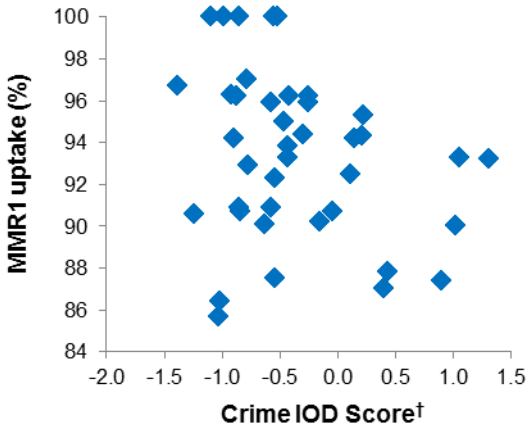
c] Environment domain



d] Health domain



e] Crime domain



† ordinal data

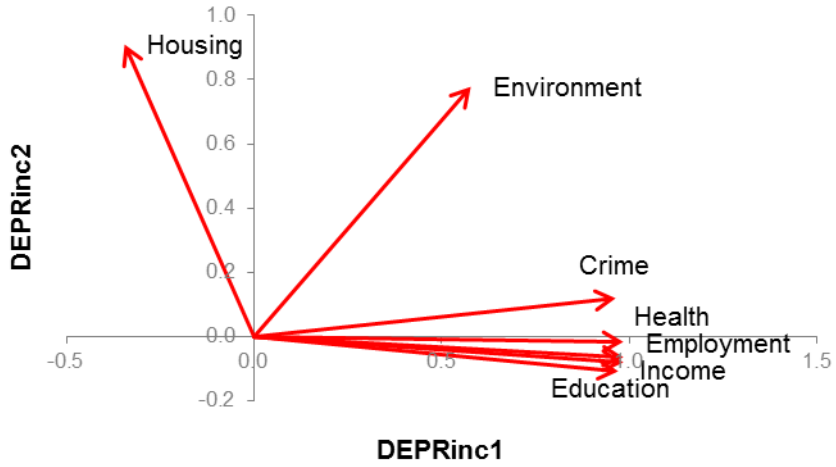
2.3.2.3.1. Dimension reduction including income deprivation

Income deprivation shows a weak negative correlation with MMR1 uptake (Spearman's rho = -0.250 with p>0.1). The initial CATPCA analysis specifies a full solution, with 7-dimensions (corresponding to the seven domains included), and its scree plot indicates that no more than 3-dimensions are appropriate. With the total VAF for a 2-dimensional CATPCA at 95.7%, a 3-dimensional CATPCA is rejected.

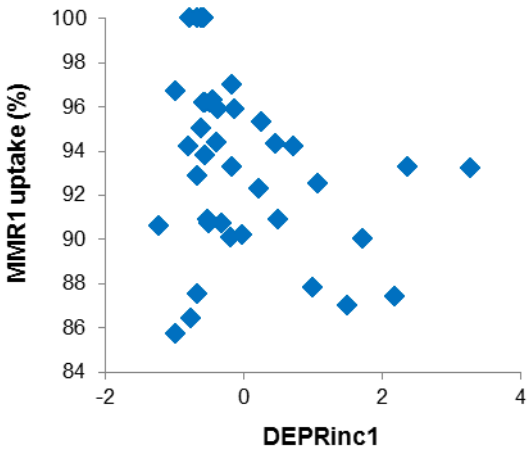
The component loadings and object scores for the 2-dimensional CATPCA are given in Figure 2-10, where it can be seen that the deprivation domains of income, employment, education, health and crime are treated similarly and contribute primarily to the first principal component (*DEPRincm*$_1$), with the second principal component (*DEPRincm*$_2$) is associated with housing and environment deprivation.

Figure 2-10 CATPCA for deprivation domains incl. income deprivation

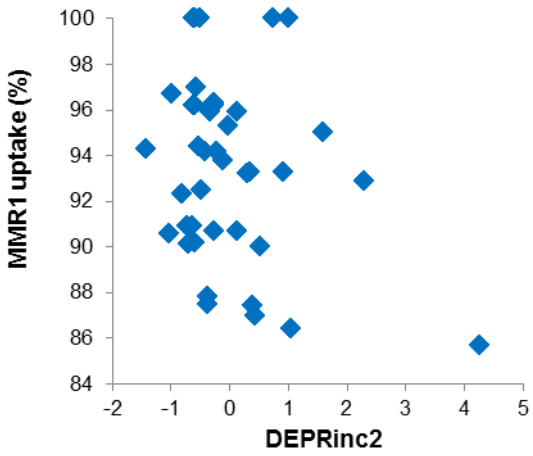
a] 2-dimensional CATPCA component loadings



b] MMR1 uptake vs first principal component (*DEPRincm*$_1$)



c] MMR1 uptake vs second principal component (*DEPRincm*$_2$)



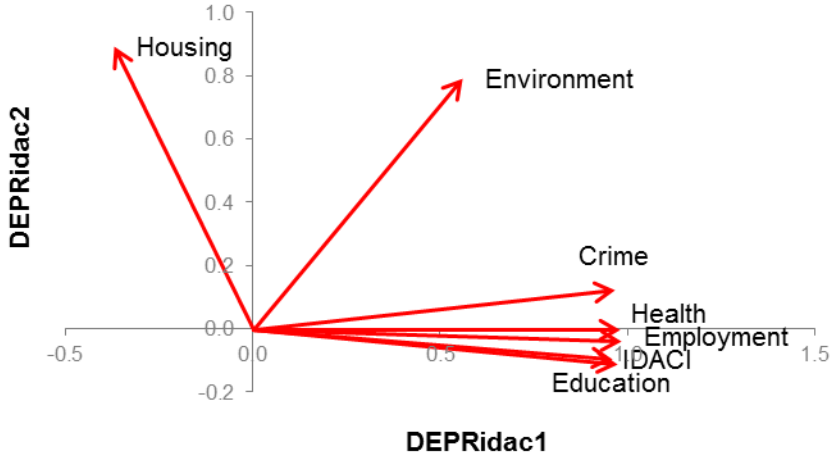
2.3.2.3.2. Deprivation reduction including IDACI

There is no evidence for a correlation between IDACI and MMR1 uptake (Spearman's rho = -0.199 with p>0.2). The initial CATPCA analysis specifies a full solution, with 7-dimensions (corresponding to the seven domains included), and its scree plot indicates that no more than 3-dimensions are appropriate. With the total VAF for a 2-dimensional CATPCA at 94.7% a 3-dimensional CATPCA is rejected.

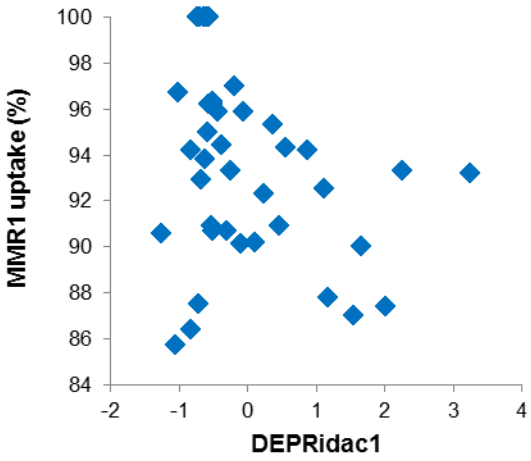
The component loadings and object scores for the 2-dimensional CATPCA are given in Figure 2-11. A similar pattern is obtained to that observed for the 2-dimensional principal components including income deprivation: income, employment, education, health and crime are treated similarly and contribute primarily to the first principal component (*DEPRidac*$_1$), with the second principal component (*DEPRidac*$_2$) is associated with housing and environment deprivation.

Figure 2-11 CATPCA for deprivation domains incl. IDACI

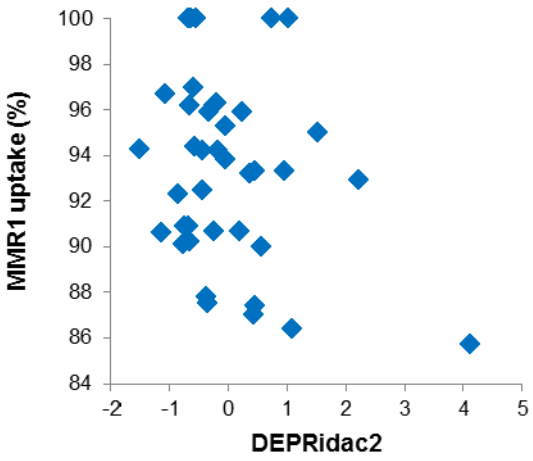
a] 2-dimensional CATPCA component loadings



b] MMR1 uptake vs first principal component (*DEPRidac$_1$*)



c] MMR1 uptake vs second principal component (*DEPRidac$_2$*)



2.3.2.3.3. Summary of supplementary analysis of deprivation

The objective of the supplementary analysis was to identify alternative univariate measures of deprivation for comparison with the IMD, which might be a more appropriate measure for the region and the subject of routine childhood immunisations.

The two sets of seven domains of deprivation both optimally reduce to a 2 dimensional solution under a CATPCA. Hence, for a univariate measure, we consider a weighted combination of these principal components. The weighting coefficients are optimised for best fit with MMR1 uptake, and the resulting measures weight the income-related and employment domains less heavily than the IMD and place more weight behind the crime and environment domains (Table 2-2).

Table 2-2 Comparison of domain contribution to deprivation measures derived from the 2-dimensional CATPCA solution

%	Income	IDACI	Empl't	Educ'n	Health	Crime	Hous'g	Envir't
IMD	22.5	-	22.5	13.5	13.5	9.3	9.3	9.3
Income	14.8	-	15.1	14.1	15.8	17.5	8.2	21.2
IDACI	-	13.9	15.2	13.9	15.7	17.4	7.6	21.0

We also consider univariate measures derived from the 1-dimensional CATPCA solutions and compare them with the 2-dimensional solution and the standard IMD (Table 2-3).

Table 2-3 Measures of deprivation

		Correlation with MMR1 uptake	
Variable	VAF	Spearman's rho	p-value
IMD	n/a	-0.416	0.008
Using income deprivation			
1-dimension CATPCA	78.1%	-0.256	0.110
2-dimension CATPCA	95.7%		
first component	74.6%	-0.263	0.101
second component	21.1%	-0.193	0.234
weighted combination	n/a	-0.421	0.007
Using IDACI			
1-dimension CATPCA	77.1%	-0.260	0.105
2-dimension CATPCA	94.7%		
first component	73.7%	-0.256	0.110
second component	21.0%	-0.171	0.291
weighted combination	n/a	-0.416	0.008

The 1-dimensional CATPCA solutions are rejected as accounting for insufficient variance within the dataset of deprivation domains. The weighted combinations of the principal components offer similar levels of correlation with MMR1 uptake as the IMD measure, so offer little advantage as an alternative univariate measure of deprivation. Hence IMD is the preferred measure of deprivation used in the regression analysis.

2.3.3. Regression

We summarise the variables resulting from the exploratory analysis to determine which are appropriate for consideration in the stepwise regression (Table 2-4). Following the dimension reduction, the 10 ward-characteristics considered are represented by 15 variables.

Table 2-4 Summary of ward-characteristic variables

		Correlation with MMR1 uptake		
Ward-characteristic	Variable	Spearman's rho		p-value
Lone parent	*LONE*	-0.178		0.272
Population density	*DENS*	-0.077		0.636
English proficiency	*ENGL*	0.244	†	0.129
Under 5 population	*PSCH*	-0.319	†	0.045
Age of parent	AGE_1	-0.042		0.799
Educational Achievement	$QUAL_1$	0.035		0.830
Educational Achievement	$QUAL_2$	0.255	†	0.112
Educational Achievement	$QUAL_3$	0.162		0.318
Working Status	$WORK_1$	-0.133		0.412
Working Status	$WORK_2$	-0.031		0.851
Family Size	$SIZE_1$	0.328	†	0.039
Family Size	$SIZE_2$	-0.133		0.415
Ethnicity	$ETHN_1$	-0.337	†	0.033
Ethnicity	$ETHN_2$	0.280	†	0.080
Deprivation	*IMD*	-0.416	‡	0.008

Categorisation of correlation strength (as defined in Evans [205]) † weak ‡ moderate

Seven of the identified variables have a monotonic correlation with MMR1 uptake categorised as weak or moderate; deprivation, ethnicity, family size and preschool population have the strongest correlation. Examination of the scatterplots for all these variables reveals no obvious non-monotonic relationships with MMR1 uptake.

From a univariate GLM analysis for MMR1 uptake (using a "logit" link function), MMR1 uptake is significantly associated with Ethnicity (first principal component), Family Size (first principal component) and families with a Lone Parent (Table 2-5). Examining the interpretations of the principal components for the associated variables reveals MMR1

uptake is positively associated with high proportions of non-Black residents and families with no dependent children; and negatively associated with high proportions of Black residents, families with 3+ dependent children and families with lone parents.

Table 2-5 Univariate analysis

Ward-characteristic	Variable	beta	p-value	
Lone parent	*LONE*	-0.022	0.043	*
Population density	*DENS*	-0.007	0.195	
English proficiency	*ENGL*	0.022	0.252	
Under 5 population	*PSCH*	-0.079	0.121	
Age of parent	AGE_1	-0.256	0.093	
Educational Achievement	$QUAL_1$	0.209	0.194	
Educational Achievement	$QUAL_2$	0.559	0.208	
Educational Achievement	$QUAL_3$	0.192	0.396	
Working Status	$WORK_1$	-0.165	0.338	
Working Status	$WORK_2$	0.157	0.794	
Family Size	$SIZE_1$	0.505	0.043	*
Family Size	$SIZE_2$	-0.338	0.564	
Ethnicity	$ETHN_1$	-0.312	0.036	*
Ethnicity	$ETHN_2$	0.259	0.194	
Deprivation	*IMD*	-0.012	0.089	

A stepwise procedure is used to develop a parsimonious multivariable model (0.05 entry criterion, 0.10 exit criterion). Only main effects are included in the model.

The resultant model contains two independent variable::

$$MMR1\ uptake\ in\ ward = (e^{2.599-0.409ETHN_1+0.954QUAL_2})/(1+e^{2.599-0.409ETHN_1+0.954QUAL_2}) \times 100\%$$

where

$ETHN_1$ is the first principal component of ethnicity (from §2.3.2.2.4)

$QUAL_2$ is the second principal component of highest qualification (from §2.3.2.2.1)

Table 2-6 Model parameters

Logistic model

Parameter	B	Lower CI	Upper CI	Wald χ^2	sig
Intercept	2.599	2.410	2.788	725.117	0.000
$ETHN_1$	-0.409	-0.714	-0.105	6.929	0.008
$QUAL_2$	0.954	0.037	1.871	4.163	0.041

Principal component variables

$ETHN_1$	Loading	$QUAL_2$	Loading
White	-0.383	None	-0.555
Asian	-0.052	Level 1	0.159
Black	0.832	Level 2	0.174
Other	-0.397	Level 3	0.077
		Level 4+	-0.484
		Other	0.629

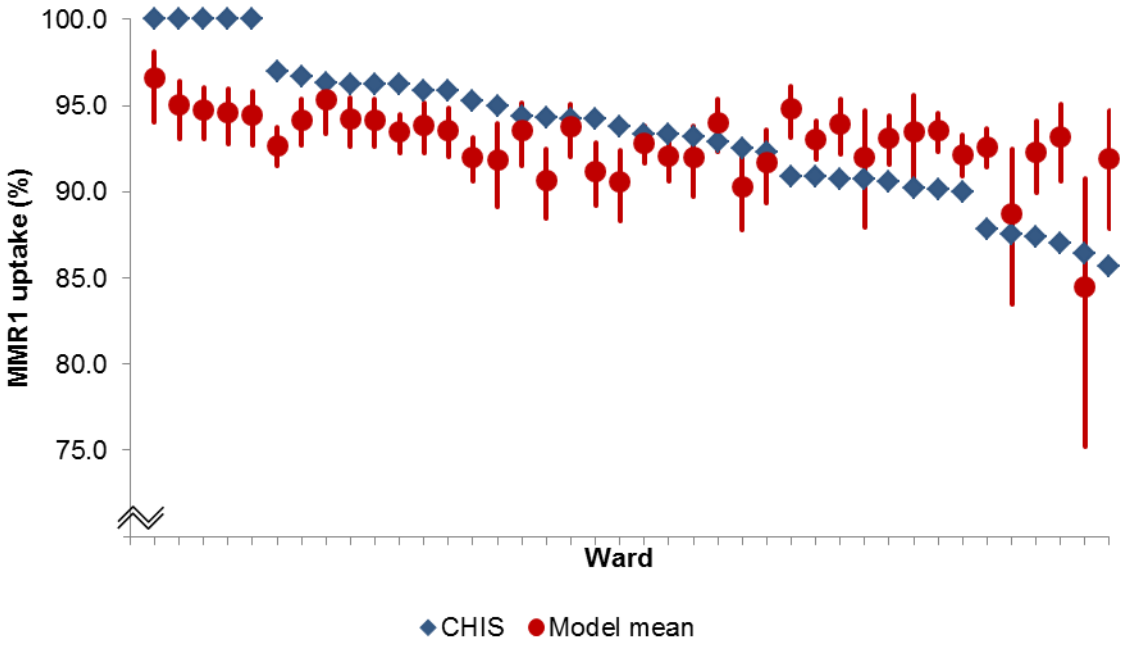
Hence this model indicates negative associations between MMR1 uptake in a ward and the proportion of the resident population with black ethnicity, and with the proportion of adults (in families with dependent children) who have no qualifications or qualifications at level 4 and above.

Given deprivation is not present in the final model, having noted it is represented by the variable with the strongest monotonic correlation with MMR1 uptake, we check the robustness of its exclusion, by substituting the alternative univariate measures of deprivation derived above (§2.3.2.3.3) for IMD and repeating the stepwise model construction; the same final model is obtained.

Model diagnostics are satisfactory. Specifically the Cook's distance values indicate that no points exert unacceptable influence on the model parameters (D<1 for all points), additionally the only point with D>0.1 is associated with the third most populous ward which is therefore one of the least likely to be affected by uncertainties due to "small numbers" effects.

The goodness-of-fit for the model is satisfactory (Hosmer and Lemeshow test: Chi-squared value = 8.937, 8 d.f., p = 0.348). There is a moderate correlation between the predicted mean value from the regression model and the observed values of MMR1 uptake (Spearman rho = 0.572, p=0.000) with an R^2 value of 0.365.

Figure 2-12 Observed values of MMR1 vs predicted values from the model (mean and 95% CI)



Comparison of the modelled values for MMR1 uptake with the values from CHIS ordered by CHIS value (Figure 2-12) reveals that high observed values are underestimated, low observed values are overestimated. A trend in errors suggests model assumptions may not be sufficiently valid (e.g. logistic regression's assumption that log odds are linearly related to the independent variables). It could also be indicative of how the non-measured variables act, i.e. amplifying the deviation from the global mean.

2.4. Discussion

We have not seen ward-level data used previously to investigate the relationships between demographic factors and MMR1 uptake in the UK – although a few studies have used ecological data from NHS primary care administrative units (GP practices [98], the forerunners of PCTs [69] and the precursors of SHA [70]). Additionally several of the demographic variables are categorical, and we made no a priori assumptions as to which category (or categories) of the candidate ward-characteristics should be included in binary measurements; we have not seen this approach in previous studies.

The analysis of MMR1 uptake at ward levels reveals, and quantifies, significantly non homogenous distribution across the PCT. In the example of Great Yarmouth and Waveney, 14 of the 40 wards meet the WHO district vaccination target for control of measles (95%

[33], set in line with critical vaccination threshold estimates), although the PCT-level uptake (92.7%) is below the WHO target, and 16 of the 40 wards are below the PCT's uptake level. It is also noted that the variation between wards in MMR1 uptake is greater than that observed at higher levels within the NHS geography hierarchy. Without further data from more PCTs, we do not know if this is unique to the specific branch of the hierarchical tree analysed or evidence of more general pattern.

The observation of spatial heterogeneity in vaccine uptake offers evidence of the presence of spatial pockets of under-vaccination within which outbreaks could occur within an otherwise well-protected population. Such circumstances indicate that public health officials in PCTs meeting WHO MMR1 uptake guidelines might not be able to afford complacency in terms of outbreak preparedness and that locally-acting interventions to address uptake shortfalls on small spatial scales should be considered to create a more robust protection.

The statistical model identifies ethnicity and educational achievement as characteristics of the ward population significantly associated with levels of MMR1 uptake, within a multivariable analysis. The educational association is not monotonic; with lower uptake for wards with a higher combined proportion of residents at the extremes of educational achievement. The presence of ethnicity as a significant factor is consistent with previous studies [61, 70, 99] (although the details of the association are difficult to compare across different categorisations employed). In addition to ethnicity, wards containing higher proportions of families with lone parents and families with 3+ dependent children are associated with lower MMR1 uptake when characteristics are considered in isolation.

Conversely, we found no significant relationship between MMR1 uptake and parental age, in turn suggesting personal memory of the onset of the MMR-autism scare is not a key factor in lower uptake (and mindful that the majority babies in the UK are being born to mothers aged over 30 [213], i.e. those who were reaching child-bearing age at the time of the MMR autism scare).

The CATPCA technique has enabled the identification of an association with education which is not based on the inherent categorical order – in contrast to the definitions used in previous studies [19, 69, 70, 89, 93, 98]. The statistical model retains an education variable representing the combined extremes (no qualifications and university-level qualifications), in preference to one contrasting the extremes. However whilst other studies provide evidence to inform hypotheses for a relationship between low MMR1 uptake with each educational extreme, they propose differing mechanics behind such associations. Two UK ecological data studies [69, 70] included cross-sectional data for periods either side of the peak of the MMR-autism controversy publicity. Although using different binary measurements of

education, both concluded that areas with better-educated populations were more affected by the scare, with greater reductions in vaccination uptake. Conversely a (causal) relationship between poor education and poor health outcomes / health behaviour is noted in health economics literature [214], and a pan-European meta-analysis of MMR1 [215] found lower education to be a significant factor in lower uptake. Neither set of literature offers a single mechanism as to why the combination, specifically, of the educational extremes is found here to be the key measurement. We hypothesise that there may be a locally acting dynamic in which an interaction between members of these two groupings drives the lowest levels of MMR1 uptake.

Qualitative studies have called attention to the need for accessible information on the risks/benefits of vaccination [84]. The nature of the educational achievement variable emphasises that careful consideration be given to the presentation of intervention content to parents, since the assumed prior knowledge, the style of language and the type of evidence (e.g. narrative or statistical) that is appropriate (and is felt appropriate by the message recipient) may not be the same for those at either end of the educational achievement spectrum. Hence multiple communication materials may be necessary to mirror the abilities and expectations of the very differently-educated constituent groups.

The demographic factors associated with MMR1 uptake (education, ethnicity, family composition) are subject to underlying population trends (e.g. the proportion of graduates in the UK population has doubled in the last 20 years [216]), if such trends act to increase the proportion of the population with the lower-uptake characteristics, we would expect the population immunity to be adversely-affected (amplified by the non-linear relationship between herd immunity and vaccine coverage [34]).

The health protection discourse [217], supported by qualitative studies (e.g. [21]), recognises that the non-presentation of children for timely vaccination may be due to a parent's disinclination to attend and/or an inability to attend. We hypothesise association of lower MMR1 uptake and the family composition demographics (lone parenting and 3+ children – which have also found to be factors associated with sub optimal participation in the UK routine baby vaccination programme [104] and for MMR1 uptake across Europe [215]) may, at least in part, relate to these parents finding vaccination appointments less accessible due to family logistics.

A strength of this analysis is the use of a single source (Census 2011) for the majority of the ward-characteristics data and the Census's near contemporaneity with the fieldwork for the other data sources (CHIS 2011-12 and IOD 2015) together with the exact coterminous nature of the geographic units used in all these sources. This removes uncertainties that

might otherwise be introduced into the analysis from disjointed temporal and spatial definitions for the range of variables considered or due to differing methodologies of different studies. This study also borrows the strengths of the Census data collection process, notably it uses population data, reducing sampling error. Sourcing MMR1 uptake data from the CHIS is in contrast to the parental recall method used by many studies examining uptake factors [19, 86, 104] as the primary or sole source of the child's vaccination status. Whilst there is evidence which demonstrates the accuracy of such data is acceptable [218], the CHIS system is the data origin for the accepted surveillance data on vaccine cover [10] and is not subject to recall bias. As a ward-level analysis, this study is inherently subject to ecological bias, however we would expect to be less susceptible to this bias than the previously published ecological studies which have larger populations units [69, 70].

The explanatory power of the model is necessarily limited by the variables considered and several candidate factors were not included due to lack of suitable data at the required level of spatial granularity. Data on GP access and appointment factors were not included. However QOF data was uninformative, and in a previous study of uptake by GP practice [98] no identifiable practice characteristics were significantly associated with MMR variation. These observations offer some evidence that the lack of (non-proxy) health service provision data is not a damaging omission from this analysis. None of the "parent-community interaction" category factors (significant in many studies [21, 80, 82, 83, 89, 91, 92, 94, 100, 101]) had suitable quantitative measures. Also, although deprivation variables were included, we note that for some domains (e.g. income) a lower score indicates a relative lack of deprivation [219] but gives little or no information on the presence of affluence. Future work could address the exclusion of these factors via data collection, however it is noted that incorporating further data sources will weaken the cohesiveness of the current datasets and may introduce sampling errors (depending on the data collection methodology).

Disclosure measures [179] may affect small cell number data from the census, so to minimise exposure to these potential effects, this analysis only uses data from census tables which have been published at ward level, and categories have been additionally collapsed to avoid low cell values (0, 1 or 2) where possible. Furthermore, Cook distance analysis of the model fitting shows that no single ward is unduly influential, and the most influential ward is relatively populous. However, in the case of ethnicity, collapsing from 19 census categories [185] to 4 broad categories may have blunted the ability of the study to capture more nuanced relationships, by hiding effects specific to narrowly-defined ethnic groups (e.g. travellers who form <0.1% of the PCT population [190]).

The generalizability of the study results is compromised by the single PCT region included; this could be addressed by extending the analysis to other geographical areas. This was the initial intention of the present study (with ward-level data sought from 32 other PCTs) but logistical and ethical approval procedural difficulties thwarted this intention, and would need to be overcome in order to produce a more generalizable set of results.

The majority of variation in MMR1 uptake at the ward-level of spatial granularity is not explained by the statistical model. It is possible that there are processes acting on the MMR1 vaccination process, acting at a local level, whose explanatory power cannot be accessed by the statistical model in its current form. This hypothesis is consistent with the factors previously identified (Chapter 1) as “parent and community interaction” (which could not be included in this model) and further supported as a potentially fruitful area for investigation when viewed in conjunction with the relative variation in MMR1 uptake at the different levels of spatial granularity. Furthermore, a global systematic review of multivariable quantitative child vaccination hesitancy studies [220] concluded that ‘Determinants of vaccine hesitancy are complex and context-specific - varying across time, place and vaccine’ (abstract, Larson et al [220]). Mindful that the concept of “vaccine hesitancy” is not fully interchangeable with vaccine non uptake (e.g. a vaccine-hesitant parent may be concerned about a specific vaccine yet still present their child for this vaccination), this indicates that spatial and temporal differences observed between the results of this study and the previous literature are not wholly unexpected and a consideration of locally-acting dynamic processes may offer more insight.

3. Initial Modelling

3.1. Motivation

Statistical analysis of ward-level data (Chapter 2) revealed geographically small scale variation in uptake of routine MMR 1 vaccination, and that demographic factors could only partially explain the observations and we proposed that locally acting dynamics may provide further explanation. In this thesis we hypothesise that that information shared between parents plays an important part in actively deciding whether or not to seek vaccination for their child, and may contribute to this small scale variation. A child's vaccination status is the result of a causal chain, so any non-random spatial clustering will depend firstly on the pattern of vaccination opinion in parents and then the conversion of intent to vaccination. In this chapter we consider the first of these stages, so we develop a mathematical model of information-sharing between individuals and the resultant pattern of opinions. The model is framed in the context of parents making a decision on whether to present their child for a specific vaccination in accordance with the routine recommendations (MMR1 in the UK schedule of vaccinations).

First we note that, although measles is endemic in the UK, recent national incidence per annum (6-31 cases per million UK population in 2010-14 [221-224]) implies that the majority of parents are making this decision outside of outbreak conditions. Reports of adverse reactions to MMR are of the same order of magnitude or lower (Chapter 1). Hence in our modelling we assume the evidence available to parents on the disease-risk to their child is not local incidence (of infection or adverse reactions, as used in several vaccination-information models [143, 146, 147, 149, 162-165, 225, 226]) but more general information (indirect reports, statistics and opinion). Hence we do not consider a coupled concurrent decision-incidence model (for incidence of infection nor adverse vaccine reactions). Our over-arching hypothesis proposes that information shared across social networks plays an important role in routine childhood vaccination decisions, so within our modelling we consider a mathematical representation of these connections.

Previously published models [166-170] potentially relevant to this scenario and its proposed treatment were identified in Chapter 1. These have demonstrated either increased outbreak probability [168-170] or increased outbreak size (at intermediate relative vaccination costs [166] or subject to initial vaccination levels [167]) when individuals place greater emphasis on their neighbours within their decision process. These results are attributed to the presence of clusters of unvaccinated individuals, which have been generated by the decision

process, either measured explicitly [166, 169, 170] or inferred Eames [168] (although cluster prevalence may be masked by the non-linear infection-transmission process and transmission-network characteristics [116]).

However, both the distribution and the proportion of unvaccinated individuals affect potential for outbreaks. Some of these models [168-170] artificially hold the vaccine-supporters' proportion constant during the decision process, whereas those where is it unconstrained [166, 167] show the decision-process increases in vaccination cover (with intermediate relative vaccination costs and baseline majority support for vaccination). Hence, it is unclear if the reported local clustering effects are artefacts associated with the mechanics of the constraint on global vaccination support. Furthermore the unconstrained models may contain an artefact in their use of steady state vaccine levels (determined after repeated "decisions" [166, 167]). Some game theory vaccination models are known to reach equilibrium after order 10^2 cycles [164, 165] ; whilst plausible for passively-observable health behaviours (e.g. obesity, smoking) this interrogation of neighbours is less appropriate for the active enquires needed to ascertain their vaccination opinion or status. Additionally, given the paucity of quantitative data on the characteristics of the network defined by those from whom parents seek advice to inform their vaccination-opinion (explored in more detail in (Chapter 4), it is concerning to note that all these models have each explored a narrow subset of potential network-structures (either in terms of structure-type or the average number of neighbours for each vertex) and, similarly a limited selection of the decision representations available in the literature.

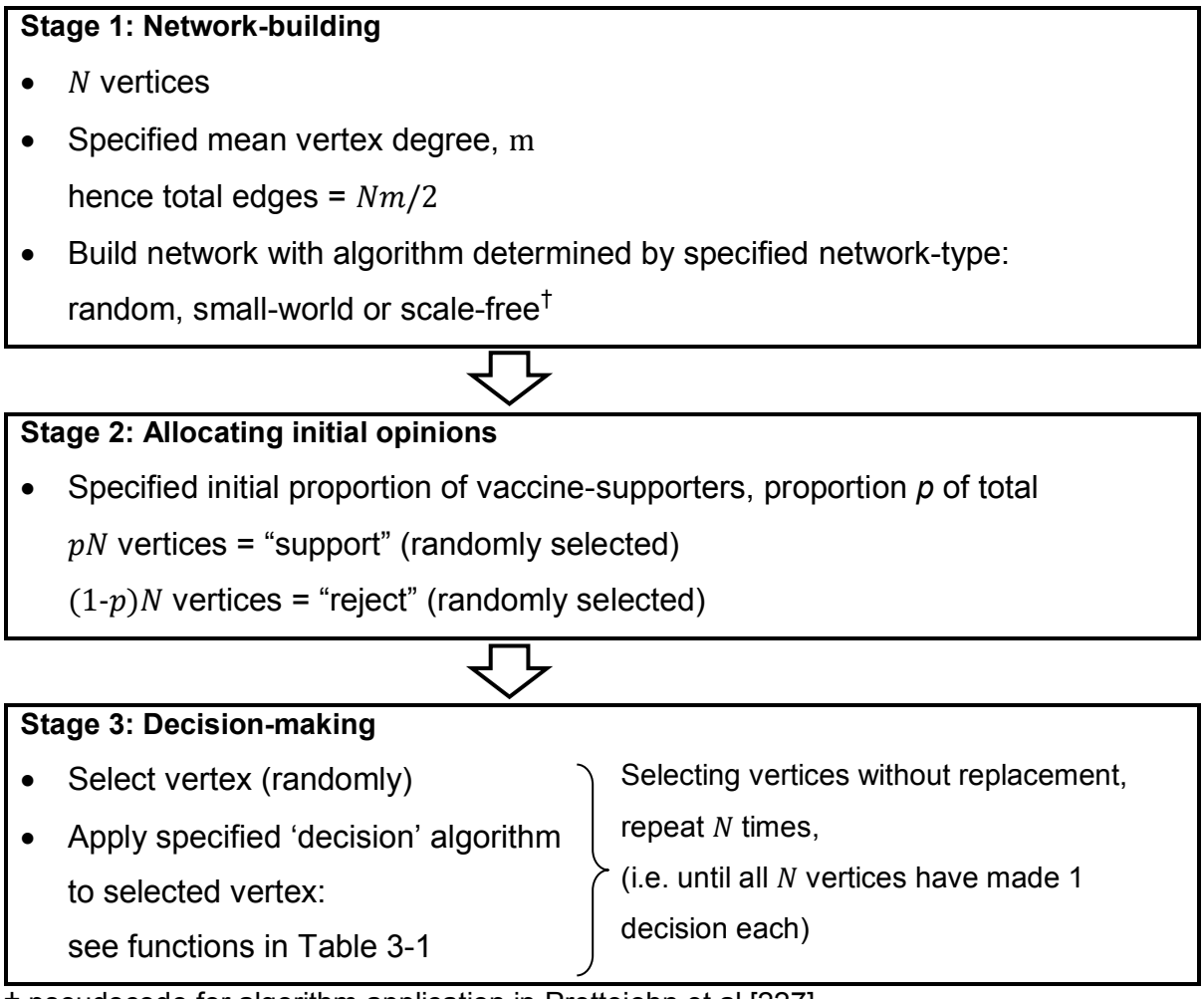
Hence, we develop a mathematical model to simulate the formation of parents' vaccination-opinion via an active process influenced by the opinions of their social contacts (with whom they discuss this subject), and use it to directly examine this process's ability to affect the both numbers and distribution of vaccine-rejecters in a cohort of parents considering routine childhood vaccination. Furthermore, given the limited knowledge on the networks over which parents seek advice to inform their vaccine-decisions, we explore if different outcomes are obtained according to the assumptions made about network type, mean number of contacts and the algorithmic representation used to model the decision-process.

3.2. Methods

3.2.1. Building the mathematical model

The model builds on the methods employed by Salathé & Bonhoeffer [170] and Eames [168]. There are three sequential stages in the model: (i) network-building, (ii) allocating initial opinions and (iii) decision-making, (schematic Figure 3-1). The variables of interest – vaccine-supporters proportion and clustering of like-minded individuals – are measured before and after the decision-making stage.

Figure 3-1 Model stages for one simulation within a specified scenario



† pseudocode for algorithm application in Prettejohn et al [227]

Stage 1: Network-building

We build a network with N vertices to represent N individuals within an information-sharing contact network. We have used N =4000, which corresponds to the annual vaccination

cohort for a typical Primary Care Trust. Each individual has a state variable representing their current binary opinion regarding vaccination (supporting or rejecting).

Three types of networks, each with different structural characteristics, are constructed using the following algorithms:

- random networks with approximately Poisson distribution of vertex degree (Erdős-Rényi [228]),
- small-world networks (Watts-Strogatz [229], using a rewiring probability of 0.02)
- scale-free networks (Barabási-Albert [230]).

In each case, the network is constructed to have a specified mean vertex degree (MVD) - the average number of contacts for an individual. The USA General Social Survey has collected data on the number of contacts with whom "important matters" are discussed [231, 242] reporting means of 3 and 2 respectively. To explore model sensitivity to MVD, we have chosen a range that includes a theoretical value used in other relevant models (ten [168-170]) and, motivated by the General Social Survey results, is skewed to lower values. We consider 15 network-structures (defined by the combination of a network-build algorithm with a specified MVD):

- random network with MVD values 4, 6, 8, 10 and 12
- small-world network with MVD values 4, 6, 8, 10 and 12
- scale-free network with MVD values 4, 6, 8, 10 and 12

The Watts-Strogatz algorithm [229] specifies MVD>$ln\,(N)$, and $ln\,(4000)\approx 8.29$. This is specified to generate a single-component network, but as connectivity uncertainty is a weakness of network data collection by sampling [232] (the only feasible option for networks like our parental cohort), we do not know if this is an appropriate property for the model. Therefore, we have relaxed the MVD>$ln\,(N)$ constraint, and (for small-world networks with MVD≤8) checked results with those obtained if consideration is restricted to single-component networks (negligible differences found in the average results). We have also checked all scenarios using N =400 ($ln\,(400)\approx 5.99$), and all results were found to be robust to this change in population (see Appendix).

Step 2: Allocating initial opinions

Initial vaccination-opinion-states are allocated randomly to each vertex, but constrained to give desired proportion of vaccine-supporters across the whole network (set at 90%, same order as recent England MMR-uptake average [42]). Opinion-clustering is measured using an intra-dyad agreement (IDA) value, i.e. the proportion of edges in the network that are

between vertices with concordant opinion-states. In a population with proportion p of individuals supporting vaccination, without opinion clustering, we expect to observe an intra-dyad agreement value of $p^2 + (1-p)^2$. Populations with opinion-clustering have a higher value for intra-dyad agreement.

Stage 3: Decision-making

All individuals make a 'decision', whether to maintain or change their initial opinion, using information obtained from their network contacts. All individuals reach their final opinion via one modelled 'decision'. We mimic the process of a cohort of parents reaching final vaccination-intent for their child (assumed to be finalised - using the latest available information - when each child reaches, in turn, the age when vaccination is due) by performing these 'decisions' sequentially across the cohort, and the sequence order is determined by random selection of individuals without replacement. We summarise the information received from the individual's contacts by one of two values: the count c of their contacts whose current opinion is opposite to the individual's own, or the proportion f of the individual's contacts formed by these opposing-thinkers. These two measures correspond to two types of complex contagion described by Centola & Macy [119]: count is appropriate for uncontested complex contagion and proportion for contested contagion, In all cases we assume that there is a positive association between the quantity of opposing information at the likelihood that the individual will change their opinion. The 'decision' is modelled using a representative algorithm, which applies a simple function to one of these values to determine if the selected individual changes their opinion-status. We consider four decision algorithms (Table 3-1):

- 'majority rule'
- 'threshold ' (Campbell & Salathé [169] use a threshold formulation, but acting on a different measure)
- 'fraction' (the algorithm used in Salathé & Bonhoeffer [170] and Eames [168].
- 'count' (a similar function has been used in infectious disease transmission models in discrete time [233])

Table 3-1 Decision algorithm formulations

p(change) is the probability that the canvassing individual changes their opinion

f 'opposing fraction' (contacts with opposite opinion, as proportion of all contacts)

c 'opposing count' (number of contacts with opposite opinion)

Algorithm	p(change)	
'majority rule'	$\begin{cases} 1 & if\ f > 0.5 \\ 0 & if\ f \leq 0.5 \end{cases}$	A deterministic decision. The canvasser changes opinion if and only if the strict majority of their contacts hold an opposing opinion.
'threshold'	$\begin{cases} 1 & if\ c \geq \alpha \\ 0 & if\ c < \alpha \end{cases}$	A deterministic decision, with parameter $\alpha \in \mathbb{N}$. The canvasser changes opinion if and only if at least a threshold number of their contacts hold an opposing opinion. We set $\alpha = 4$ †
'fraction'	f	This is a stochastic decision. The probability of opinion-changing is proportional to the 'opposing fraction' value.
'count'	$1 - e^{-\beta c}$	This is a stochastic decision, with parameter β. The probability of opinion-changing approximately proportional to the 'opposing count' when βc is small. We set $\beta = 0.05$ ‡

† sensitivity explored for $\alpha \in [3, 6]$

‡ sensitivity explored for $\beta \in [0.0125,\ 0.1]$

For the two formulations with a parameter, threshold (α) and count (β), sensitivity to parameter-value was explored (see Appendix). Results for 'count' were found to be qualitatively robust for $\beta \geq 0.025$; for 'threshold' the ratio between the network MVD and 2α was found to be of interest (§3.3).

Numbers of simulations

We explore the effect of each of the four decision-algorithms being applied on each of the 15 network-structures; so there are 60 scenarios. For each scenario, 100 networks are built, and on each of these networks 100 simulations of the decision-making process are run. Hence 10,000 simulations are run for each of 60 scenarios.

For each simulation we compare the final value of intra-dyad agreement (the clustering measure) to the expected value in a randomly-mixed population. The expected value is calculated for both the initial vaccine-support proportion and for the vaccine-support proportion observed in the network after decision-making.

3.2.2. Applying the decision algorithms in a randomly-mixed population without underlying contact network structure

We calculate the expected outcome of each decision-making algorithm applied to an individual in a randomly-mixed population in the absence of an underlying network structure.

As with the mathematical model, we assume initial opinion-states are determined by independent identical Bernoulli trials, with a 0.9 probability of "vaccine-support". To provide the 'opposing fraction' and 'opposing count' values for use in the algorithm, a specified number, k, of individuals (acting as 'contacts') are drawn at random from the population; hence the number of supporters 'contacted' is binomially distributed ~ B (k, 0.9) using same 90% vaccine-support initialisation as the mathematical model.

For each decision-making algorithm and specified value of k, we calculate the probability that the opinion state will change, for both possible initial states ("support" and "reject"). These two values are then aggregated, weighted by the probability that a randomly selected individual is initially a supporter or rejecter, to produce an expected change in supporter numbers in the population as a result of this 'decision' by a single randomly-selected population member. We also calculated the resulting change in the number of the canvassed set of individuals who have the same opinion as the deciding-individual (not an IDA as there are no dyads, but presumed to be informative when analysing the dyad characteristics).

3.3. Results

3.3.1. Numbers of vaccine-supporters

The mean percentage of vaccine-supporters observed after the decision-making process (Table 3-2) when compared with the initial value (90%) is qualitatively independent of network-type, but varies by decision algorithm. (The only exception is the 'threshold' decisions' on the highest MVD scale-free network). The 'majority rule' algorithm demonstrates a "rational herding" effect, with final vaccine-support approaching 100%.

Table 3-2 Post decision process vaccine-support (%) - mean

Proportion of all network vertices which have a final opinion-status "support".

Mean taken across 10,000 simulations for each combination of network structure (network-type and MVD) and decision-making algorithm.

	random network-type				
			MVD		
decision-making algorithm	4	6	8	10	12
'majority rule'	98.8%	99.6%	99.9%	100.0%	100.0%
'threshold' (α=4)	95.0%	98.0%	99.2%	99.4%	99.1%
'fraction'	90.0%	90.0%	90.0%	90.0%	90.0%
'count' (β=0.05)	89.9%	89.7%	89.4%	89.1%	88.6%
	small-world network-type				
			MVD		
decision-making algorithm	4	6	8	10	12
'majority rule'	99.2%	99.8%	100.0%	100.0%	100.0%
'threshold' (α=4)	95.7%	99.7%	99.8%	99.5%	98.8%
'fraction'	90.0%	90.0%	90.0%	90.0%	90.0%
'count' (β=0.05)	89.9%	89.7%	89.5%	89.2%	88.8%
	scale-free network-type				
			MVD		
decision-making algorithm	4	6	8	10	12
'majority rule'	99.0%	99.5%	99.9%	99.9%	100.0%
'threshold' (α=4)	91.7%	93.2%	94.7%	92.6%	86.2%
'fraction'	90.0%	89.9%	90.0%	90.1%	90.0%
'count' (β=0.05)	89.6%	89.1%	88.6%	87.9%	87.1%

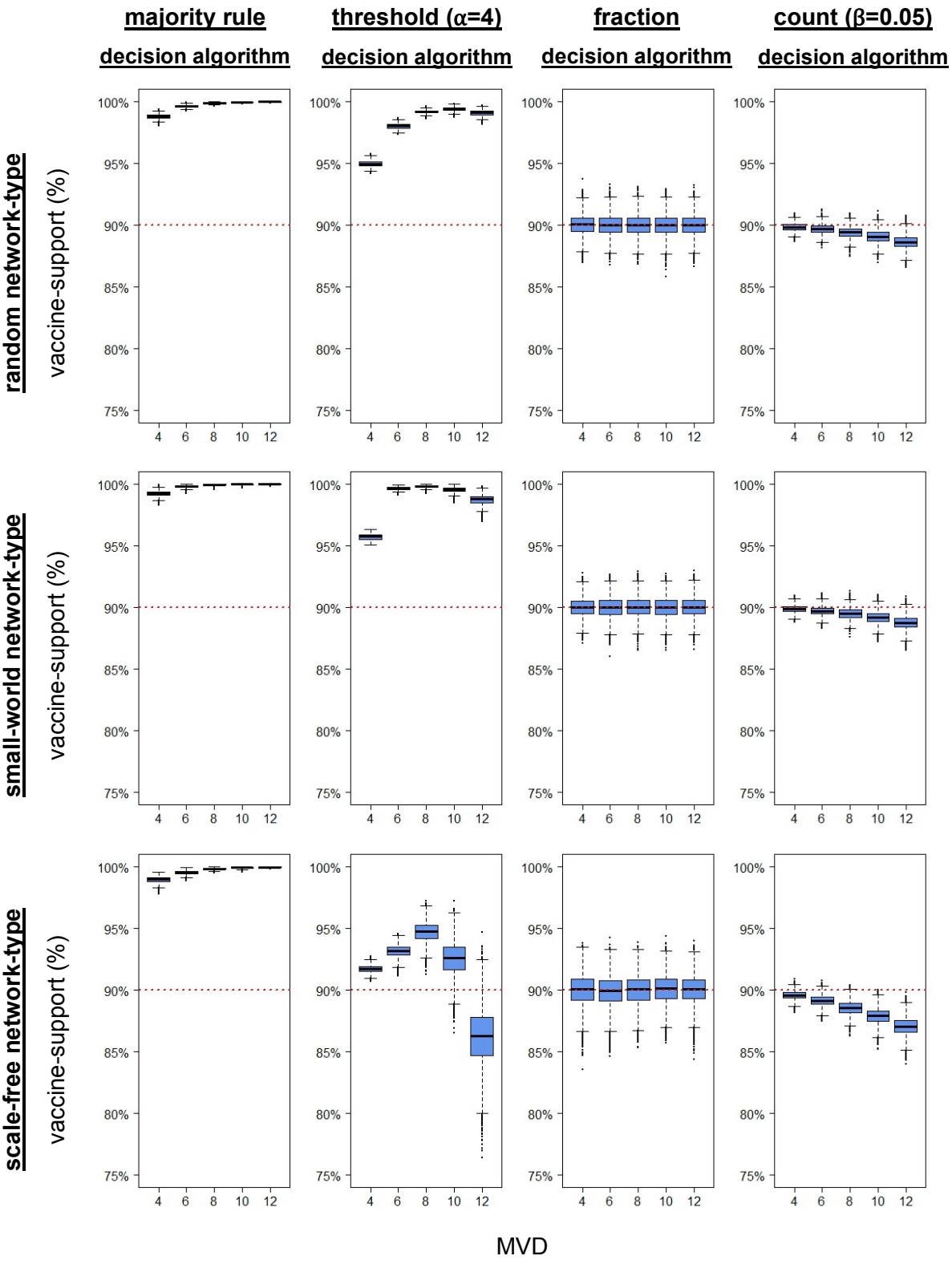
The vaccine-support proportion by simulation (Figure 3-2) also shows qualitative results independent of network type (with the same exception). The majority of simulations using 'majority rule' and 'count' result in increased and decreased vaccine-support, respectively, vs. the initial 90%, with greater differences associated with higher MVD. The 'fraction' algorithm does not substantially move the median value, but does show a wider range of results on the same network-structure.

Considering the increase in support as a function of the network's MVD, there is a turning point for the 'threshold' algorithm within the range examined. The sensitivity analysis to the parameter α within this formulation (see Appendix) demonstrates this maximum occurs at MVD=2α. This non-monotonic result is independent of network-type.

Figure 3-2 Distribution of vaccine-support (%) post decision process

Proportion of all network vertices which have of final opinion-status “support”.

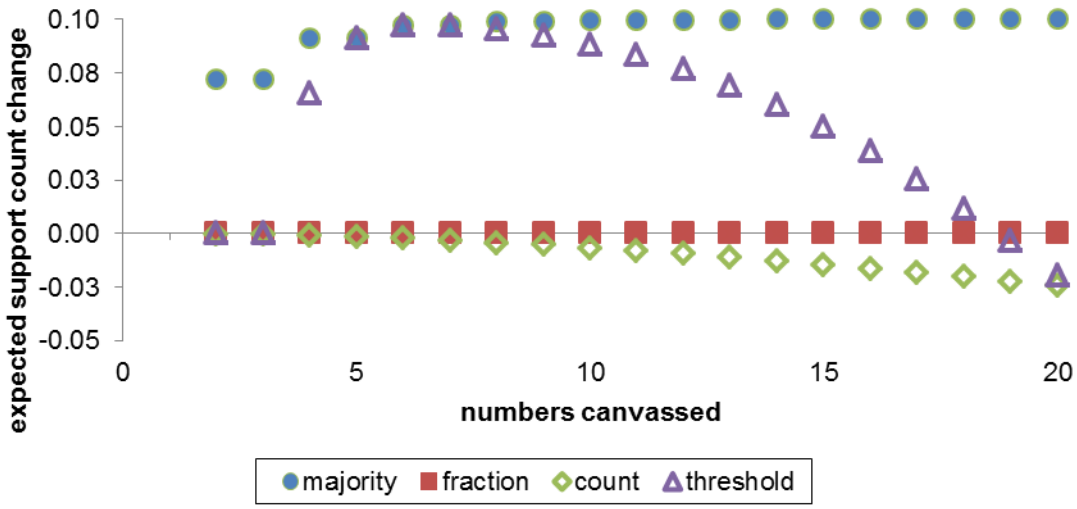
Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.



These results are directionally consistent with the expected outcome of each decision-making algorithm applied to an individual in a randomly-mixed population in the absence of an underlying network structure (Figure 3-3).

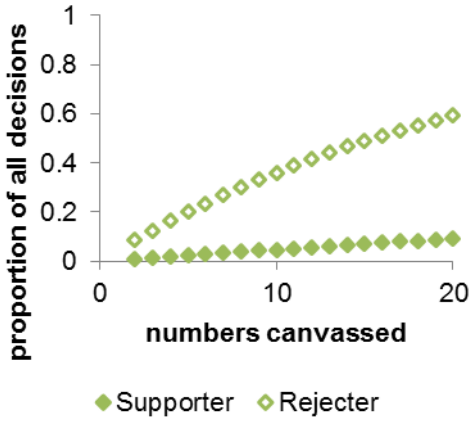
Figure 3-3 A single decision in a population without network structure – expected outcomes

a] expected change (support=positive) resulting from a randomly selected individual's decision (supporter and rejecter outcomes weighted by population prevalence)

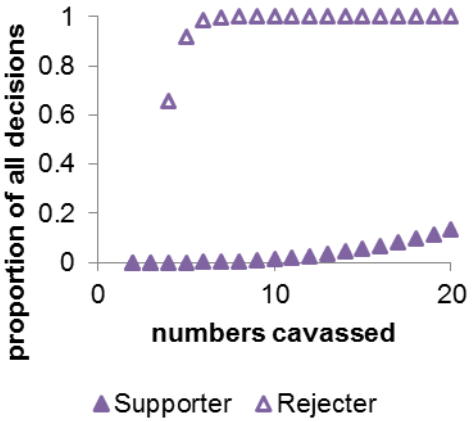


expected proportion of decisions which result in an opinion change

b] count algorithm



c] threshold algorithm

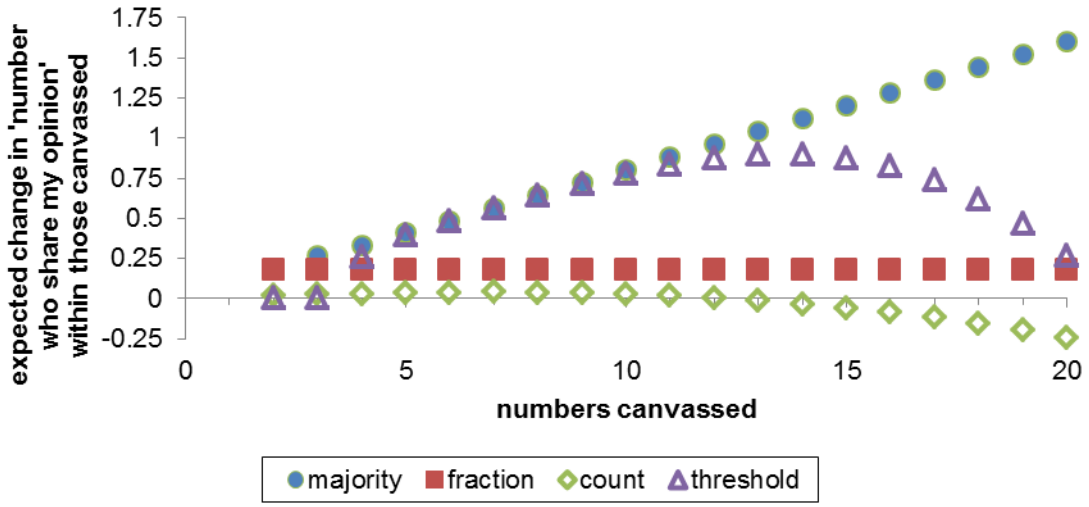


3.3.2. Clustering of vaccination opinion

After the decision-making process, the observed intra-dyad agreement (IDA), for a specified scenario (Figure 3-5) was increased for 'majority rule' and 'fraction' algorithms (across all network-structures). This IDA measure indicates increased opinion-clustering was present in these networks after the decision process. However the IDA quantitatively varied with MVD: higher MVD associated with smaller increases when using 'fraction', but associated with larger increases when using 'majority rule'.

The qualitative results of both 'opposing count'-based algorithms ('count' and 'threshold') varied by network structure, with scale-free networks differing from random and small world structures. For the 'count' algorithm, on random and small-world networks the opinion-clustering increased on lower MVD networks, but decreased on higher MVD networks. For the 'threshold' algorithm the opinion-clustering increased on the random and small world structures, but decreased (and by the largest observed quantitative amounts) when applied on a scale-free network. The single non-network decision scenario (comparing the individual's expected post-decision opinion with their canvassed individual's opinions, Figure 3-4), leads us to conclude that the heavy tail of the scale-free network's degree distribution contributes to this phenomenon because (for sets of canvassed individuals with set size in the top quartile examined) both these algorithms begin to be associated with opinions being more likely to change against the local majority as the canvass 'sample size' increases.

Figure 3-4 A single decision in a population without network structure - expected outcome compared with opinions of those canvassed



However, the expected clustering varies with the proportion of vaccine-supporters, which itself varied considerably between individual simulations (Figure 3-2). Hence, to distinguish clustering that is not solely due to changes in population-level vaccine-support, we compare the observed value with the expected value in a randomly-mixed population with the observed proportion of vaccine-support (calculated on a simulation-by-simulation basis).

These results (Figure 3-6) are not presented for the 'majority rule' algorithm due to the ceiling-effect at 100% vaccine-support, where no clustering is possible; we also note that the threshold algorithm also has this ceiling effect for non small-world networks with mvd≈2α. Restricting consideration to the remaining scenarios, all displayed median clustering changes. For the 'fraction' algorithm, the majority of simulations on all network-structures have increased clustering, with greater clustering (at equal MVD) produced on small-world networks and greater clustering was associated with lower MVD. For the 'count' algorithm, whereas increased clustering was again observed on random and small-world networks, by contrast reduced clustering was produced on scale-free networks and greater effect was associated with higher MVD. A third pattern was observed for the 'threshold' algorithm, with lower than expected levels of clustering on random and scale free networks, with the greatest deviations from expected clustering of the examined scenarios. As with the vaccine-support measure, there is some evidence for a turning point in relationship between this measure and network MVD (a minimum but at a higher MVD than 2α).

Figure 3-5 Intra-dyad agreement (%) post decision process

Proportion of all network edges which connect vertices of same final opinion-status.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.

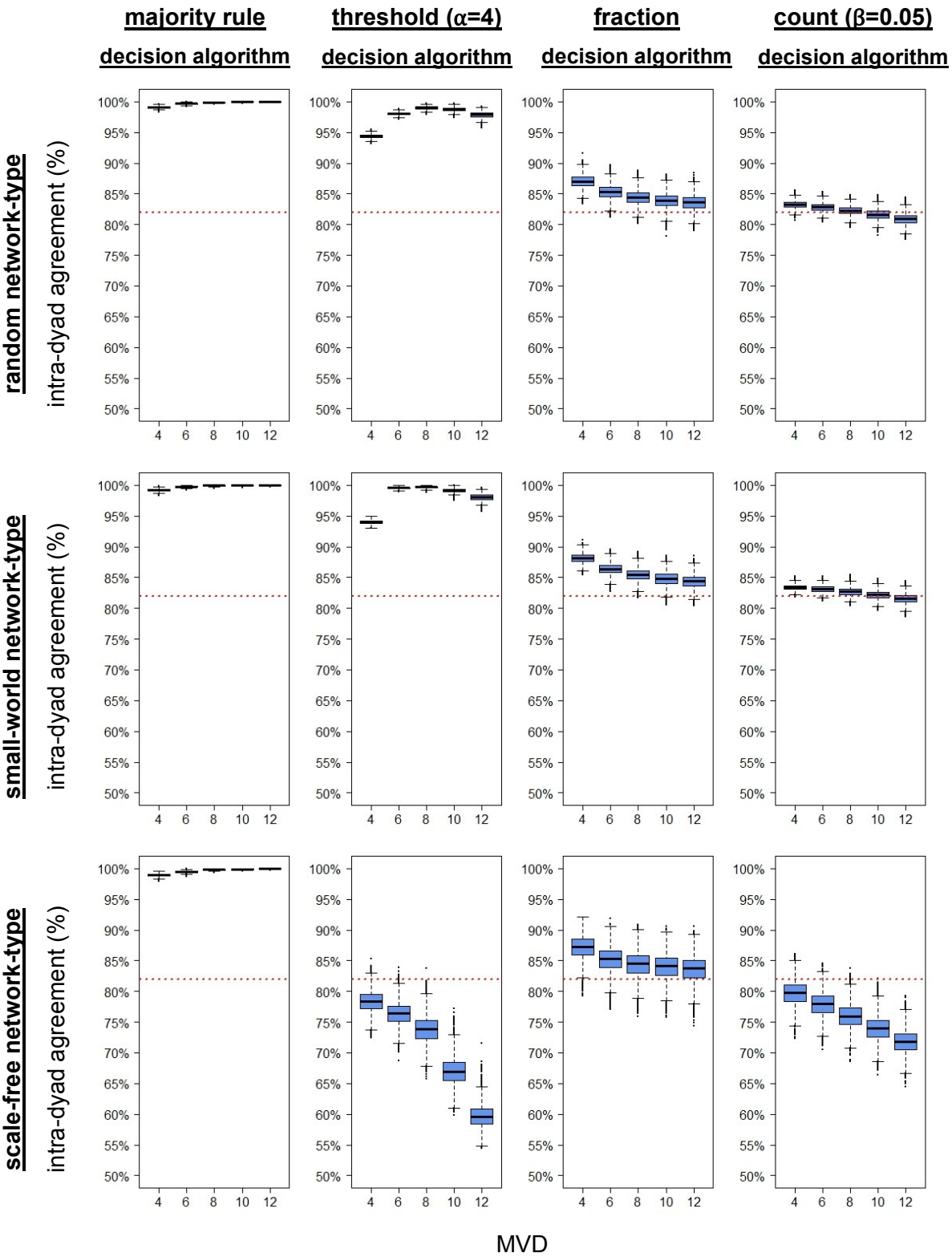
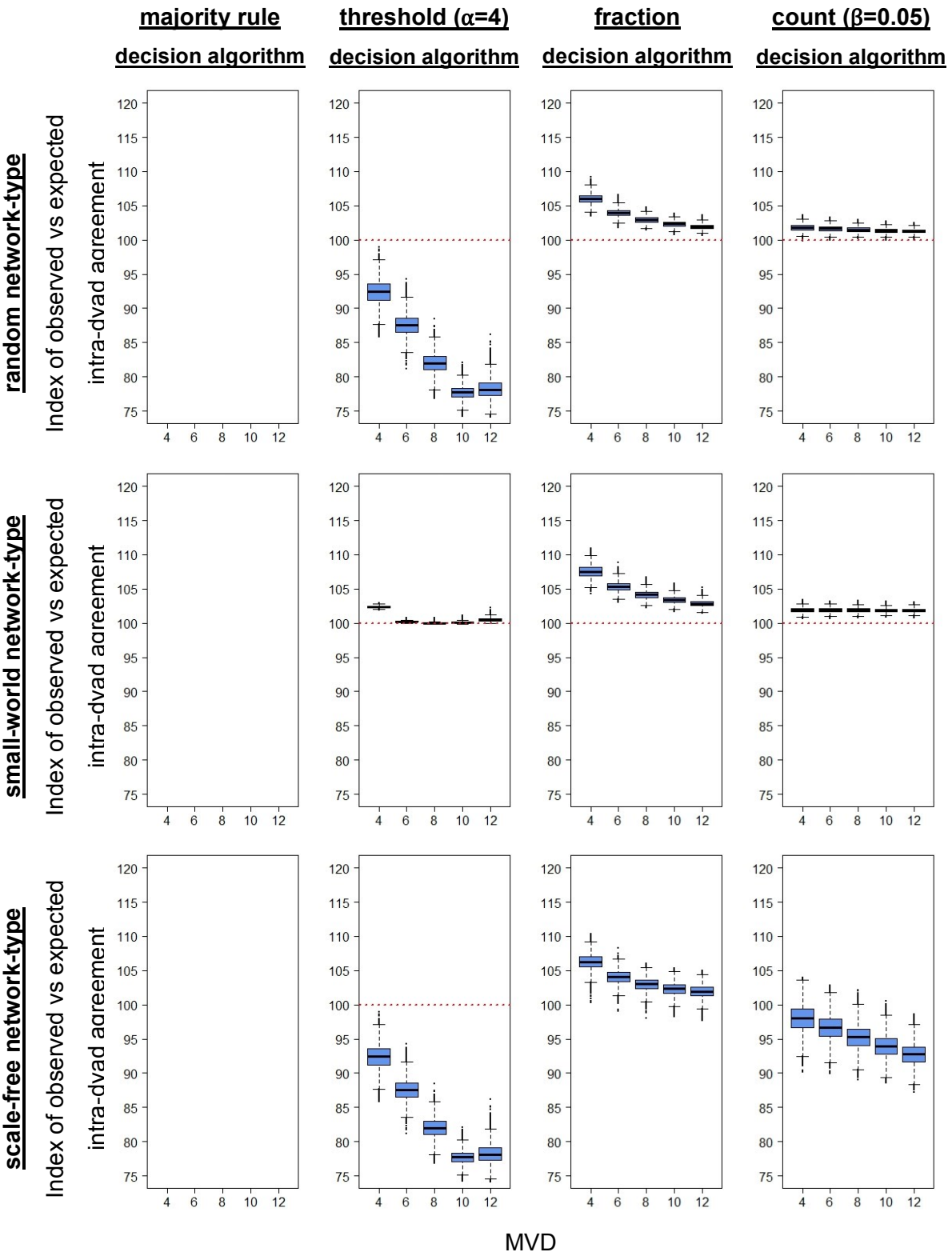


Figure 3-6 Intra-dyad agreement post decision process, observed vs expected value

Index: expected value = 100, calculated by simulation.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.



3.4. Discussion

The model indicates application of a decision-algorithm (which uses information sourced from network-contacts) is able to alter vaccine-support levels within a cohort of parents considering their child's routine vaccination. However, this effect is qualitatively-dependent on the representative decision-algorithm, and quantitatively-dependent on the network-structure. These results indicate that, even with a constant 'default' vaccine-support, this individual-level process may contribute to dynamics in vaccine-support observed at a population-level [42]. Specifically, under certain assumptions, a non-normative anti-vaccination sentiment spreads within a highly pro-vaccine population, without a pre-requisite of significant initial opinion-clustering.

This study also explicitly demonstrates that such active decisions can create vaccine-rejecter clusters. This effect was present in five of the seven combinations of decision algorithms and network-structures which are not compromised by ceiling effects (within the range of MVD explored). This confirms the inferences of previous work [166, 169, 170] and extends this result across a broader range of plausible contexts than previously examined.

However, for a given decision algorithm, the results are sensitive to network-structure: quantitative differences are observed (mainly by MVD) for all algorithms, the 'count' algorithm displays qualitative differences across the range of networks explored here and the 'threshold' algorithm's results are non-monotonic with respect to the underlying network's MVD, for both vaccine-support and clustering . Hence understanding this mechanism's potential to contribute to increased outbreak probabilities in highly vaccinated populations (via weakened (local) herd immunity, as previously proposed [166, 169, 170], requires knowledge of the information sharing network-structure (and examination of later stages in the proposed causal pathway, e.g. parent and child networks overlap [168]). As noted previously, there is little suitable data on information-sharing networks in the context of routine childhood vaccination. Cross sectional studies investigating patterns of social contacts (as appropriate to pathogen transmission) have been conducted in the UK [234, 235]. These studies' questionnaire methodologies may be usefully adapted to generate data that can be used to validate network-structure choice, and hence determine which patterns of vaccine-rejecters might realistically be produced by an active decision process. Results are also sensitive to the algorithm representing decision process, both in the assumption of contested or uncontested complex contagion and its parameterisation (absolute and interaction with the network MVD). Quantitative surveys touch on factors related to the former, but do not include explicit measurement of the latter.

Model limitations include the simplifying assumptions that all parents make an active decision, that this decision is based solely on the information shared, that all parents' decisions can be mathematically represented with the same function, that this function is 'symmetric' and that initial opinions are randomly-mixed. Before this process can be confidently incorporated into a fuller model of the action of information shared on social networks on subsequent dynamics of vaccine preventable-diseases, it is necessary to consider the suitability of these assumptions, and adapt the model if appropriate.

Addressing the first assumption by including some parents who make no active decision could exploit existing data [21]. However, those researchers warn about the validity of their measure of "automatically" getting the child vaccinated [236], concerns which transfer to other measurements appropriate to address model assumptions, which may also be reliant on subjective measures. It may be more appropriate to use decision-pattern data to validate model output where assumptions are relaxed by including 'spontaneous' individual opinion-change (changes attributable to other individual-specific reasons [139]), including information external to the network (e.g. media [169]) and introducing heterogeneity in decision representation by specifying a distribution of parameter values (e.g. α or β in the 'count' and 'threshold' algorithms respectively).

Our example algorithms representing the decision-response to the information-shared are plausible, but simplistic. Based on psychological theory and evidence regarding MMR decision-making, a full model would benefit from more asymmetrical formulations in two aspects. Firstly, we have used the same function irrespective of opinion-status; whereas psychological studies suggest decision-making is not independent of one's current opinion and that reference heuristics [127] and normative influences [134] affect decisions. Specific to childhood vaccination, some parents fear that rejecting the societal norm (vaccination) makes them a 'bad parent' [87] and omission bias may be present [237, 238]. This suggests state dependent decision-parameterisation is more appropriate. Secondly, all contacts are equally weighted (in calculating f and c). Psychologists propose primacy and recency effects lead to unequal treatment of received information [122] and studies of MMR-decisions suggest the advisor's relationship with the parent affects the weight their opinion carries [21, 87, 100, 101]. These sources however do not provide a quantitative metric to apply weights to network edges, representing the relative importance accorded to the information it carries, Previous work [167] has used frequency of physical contacts as a proxy, although data to confirm the spatial range of vaccination advisors is lacking. We have also assumed a positive relationship between the information categorisation and its effect the opinion of all recipients' opinion. However, experiments have shown a "backfire effect" may be present [239], also it may be possible that parents use the information

received as, in effect, a survey of (future) local uptake [240] (whereby pro-vaccine information may prompt (anticipated) ‘free-riding’ vaccination-rejection).

Finally, this study considers a closed cohort (without initial clusters) in order to more clearly identify the potential effect of an active decision process on the levels of vaccine-support and opinion-clustering. When incorporating this process into a full information-infection model it may be more appropriate to relax these constraints. Specifically, homophily within social networks has been observed [241]; hence a non-random opinion mix may be a more appropriate initial condition. Also social contagion theory proposes network ties inherently encourage the development of shared beliefs and behaviours; (including some health behaviours [137, 138] . It is plausible that this process may have acted on the parents, prior to the active decision-making modelled here, although the limited published data on the contact-network consulted during active decision-making implies there is also a lack of data on its relationship to the social network over which social contagion is proposed to act.

Whilst this study has specifically shown this process of making active vaccination-opinion decisions (using information from one’s social contacts) is able to produce changes in population-level vaccine-support (where vaccination was the ‘default’ opinion for the vast majority) and to create clusters of like-minded individuals, it has also demonstrated that these effects are both qualitatively and quantitatively dependent on the underlying information-sharing network properties and the assumptions used to model the decision process. Hence in order to better ascertain the credibility of the presence of this mechanism to produce clusters of vaccine –rejecters within the context of routine childhood vaccination in the UK, requires more empirical data to determine the characteristics of the networks involved and to enable validation of decision-representation.

4. Survey

4.1. Motivation

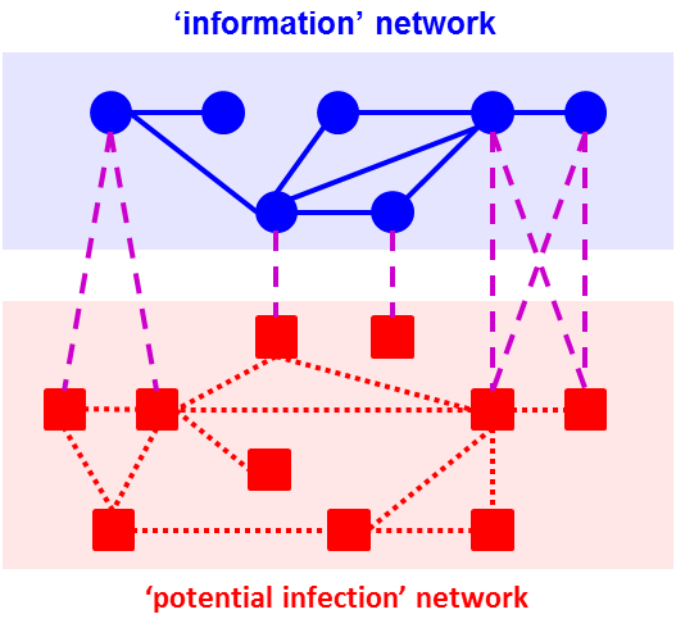
The mathematical modelling (Chapter 3) finds results that are sensitive to the network structure and decision formulation. We therefore seek evidence for the empirical values of these parameters, as they relate to the research question. Also, previous mathematical modelling (Chapter 1) has shown that when the outcome of a vaccination decision process is enacted in a population which is then challenged by an imported infection, the resultant infection dynamics are additionally sensitive to the contact network structure across which the infection can be transmitted [166-170] and, in the case where the decision and infection processes act on different populations, the links between them [168].

We therefore wish to describe the structure of 'information' network, the 'potential infection' network and for the connections between these two networks (Figure 4-1). These networks are specific to the decision/infection under consideration; we consider decisions as to whether to vaccinate with MMR(1) in adherence with the UK schedule of childhood immunisations [5], and social contacts between pre-school children that are appropriate to the transmission of measles.

An aside on the terminology adopted here:
To distinguish between empirical data and mathematical models: social network theory terminology is used when describing the empirical networks (i.e. 'nodes' connected by 'ties') whereas graph theory terminology is retained for mathematical models (i.e. 'vertices' connected by 'edges'). When viewing the network from the perspective of a specific node we refer to it as an 'ego' and its 'alter(s)' are the node(s) with whom it has a network tie. A set of three connected nodes [206] (a 'triad') is 'transitive' if ties exist between all three pairings ("my friends are also friends with each other"), and 'intransitive' otherwise ("we are not friends but do share a mutual friend").

Figure 4-1 Networks to be discovered

Networks to be discovered by survey



Legend

●	'information' network node	adult
—	'information' network tie	vaccination advice shared
■	'potential infection' network node ...	child
···	'potential infection' network tie	social contact relevant for measles transmission
—	tie between the two networks	adult (node in 'information' network) is parent of child (node in 'potential infection' network)

Empirical examples of the 'potential infection' network are obtained from contact tracing which is part of standard measles case management [243] but this information is not publically available. There have been a number of published studies which collect detail on networks of social contacts, if the participants and type of contact measured are relevant for measles transmission within the UK pre-school population, then they could considered for use here. Measles transmission requires physical proximity; data on such contacts may be collected from reports provided by study participants, from proximity-sensors carried by the participants or by observation of the contacts by researchers [244].

Both sensor-based and report-based studies have been used to collect data on social contact patterns of children. Sensor-based contact studies are only able to provide data on contacts within a closed, pre-defined set of participants. They have been used to study networks within schools in USA [245, 246], and France [247, 248]. Report based studies have also collected school-based datasets [246, 249, 250] . Whilst some of these studies

have included primary school age-children in UK [249], USA [251] and France [248], they have not captured the contacts of younger, pre-school children.

In the UK, data on young children's contact patterns has been captured within larger all-age report-based studies [234, 235]. In the POLYMOD study [234], the youngest reported child age-group was 0-4 years, which was oversampled as compared to its proportion in the census (n=95, 5.7% of GB respondents). In the British Social Contact Survey (BSCS) [235], the youngest reported child age-group was 0-10 years, which was "not well represented" (Danon et al [235] SI p1) within the respondents (n=18, 0.3% of respondents who gave an age). Both of these studies used prospective anonymous diaries for a single day to capture the social contacts (responses for young children being completed by their parents); frequency of contact was collected and used to calculate and estimate the contacts across longer periods of time. Both studies request contacts that were either face-to-face conversations or skin-to-skin physical contact. BSCS included an estimate of transitivity via third party reporting of contact to contact meetings [235].

The details of the social networks across which parents' vaccination decisions may be influenced are not captured in these studies as proximity contact studies, such as POLYMOD and BSCS, do not capture information transfer opportunities that are not face-to-face conversations. Online or phone records offer rich datasets to recreate networks over which information may be transferred [235, 252, 253]. However, there are channel-specific concerns as regards the fit with the definitions (tie and node) required here e.g. phone records are unable to filter contacts by communication content and content-specific search engine use [254] does not represent human-human contacts. However "social media" datasets are available with details of ties and message content that, with machine learning, become logistically categorisable by sentiment. Analysis of social media on the introduction of a new vaccine found assortativity by message [255]. However for established routine child vaccinations the timespan over which parents are receptive to information is beyond that which is practical for these datasets, as active consideration of child-vaccination opinions starts during pregnancy [256] and the scheduled age for MMR1 (12 months) further lengthens this period as compared with other routine vaccinations [5]. Additionally, there are limitations on the spatial proximity measures of online network members, and hence how online ties map into potential infection transmission contacts (especially for children who have a further separation from the online network).

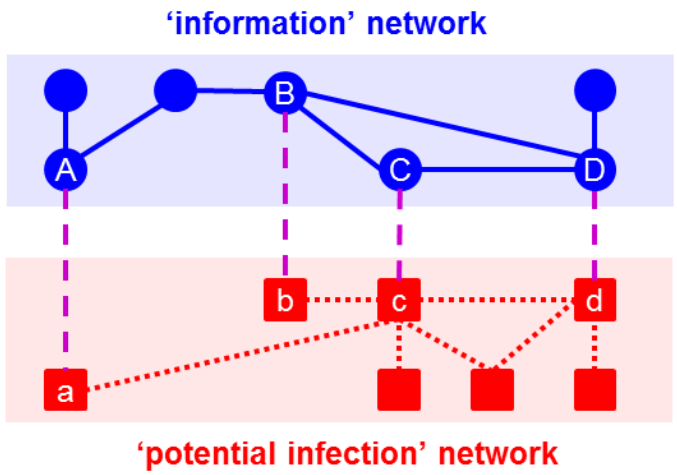
A number of studies in the UK have sought information on the types of people from whom (MMR) vaccination information is sought [21, 80, 82, 87, 94, 95, 97, 100-102, 257], but they have not collected data on the numbers of these contacts that would provide node-degree

information for the associated ‘information’ network. Surveys in the USA have captured the contacts with which “important matters” are discussed [231, 242] and have reported the average number of such contacts. These studies have not attempted to capture network structure and the contacts recorded may not reflect discussion of routine childhood vaccinations. Also, the different health service provision and vaccination legislation in the USA and the UK (in the USA the vaccinations are not necessarily free of at point of access and some states require proof of vaccination prior to school enrolment) indicates that the generalizability of data on the vaccination decision process across these two locations cannot be assumed.

This lack of data on ‘information’ networks means that the relationship of adult contacts within this network to children within the corresponding ‘potential infection’ network is inherently unknown. This uncertainty also extends to the ‘overlap’ of nodes and the ‘overlap’ of ties (Figure 4-2) (this comparison may be facilitated via alternative conceptualisation of the nodes in the separate networks as parent-child family-unit nodes in a duplex). For example, the influence of another generation, such as the parents’ own parents, has been observed in new parents’ breast-feeding decisions [258] and is specified in one study of vaccine information sources [21]. The inclusion of these individuals in the ‘information’ network would mean that not all ‘information’ network members are parents with dependent pre-school children (who might be in within the ‘potential infection’ network).

Figure 4-2 Networks' overlap

a] 'Information' and 'potential infection' networks' overlap toy example



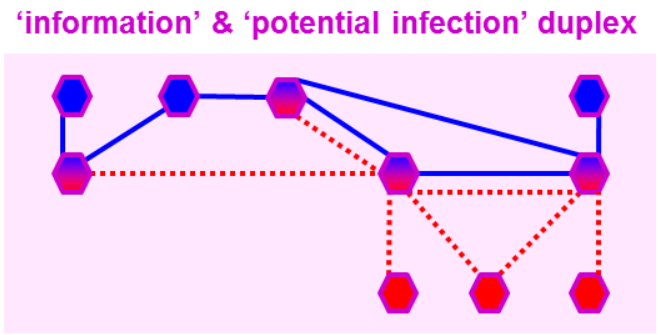
Pairs of nodes which 'overlap':
A & a, B & b, C & c, D & d

Pairs of ties which 'overlap':
B-C & b-c, C-D & c-d

Legend

- 'information' network node
- 'information' network tie
- tie between the two networks
- 'potential infection' network node
- 'potential infection' network tie

b] Same toy example reconfigured as a duplex



Duplex nodes with both 'information' and 'potential infection' ties correspond to 'overlapping' nodes in the original paired networks. Where both types of ties join the same pair of duplex nodes, these ties correspond 'overlapping' ties in the original paired networks.

Legend

- duplex node ..... family unit (internal colours inherited from original networks)
- 'information' network tie ..... vaccination advice shared
- 'potential infection' network tie ..... social contact relevant for measles transmission

In summary, few empirical studies of UK pre-school children's social contacts have been undertaken, none of which directly measure the cumulative contacts across the 6-8 day period appropriate to measles infectiousness [28] and the only study identified as attempting to capture transitivity of this contact network [235] has a very small number of young children

in its sample. Whilst several studies have found evidence of peer-to-peer vaccination information sharing, empirical quantitative data including the total numbers of information contacts by UK parents and the structure of the network they form were not found in the literature.

Given this paucity of required data on these networks, a survey was undertaken to discover the network structure parameters for use in mathematical modelling. This also provides the opportunity to collect data on the qualitative properties of the nodes and ties within each network (social relationship between ego-alter, sentiment of the vaccine-related information shared, and MMR status of the children).

The data on sentiment of information shared between 'information' network nodes provide direct empirical evidence to inform a choice of mathematical formulation to represent the decision-process, addressing another sensitive assumption in the mathematical model.

Also, these additional data open the possibility of obtaining empirical data on the amount of clustering of vaccination opinion in the 'information' network (as predicted by theoretical work but, to our knowledge, not previously quantitatively measured) and on clustering of vaccination status in the 'potential infection' network.

'Information' network data were collected from parents of pre-school children, who were also asked to report 'potential infection' network data for their pre-school children. Assumed open communities directed the choice of social network sampling methodology [232] (ego-centric networks with alter-connections, with snowball sampling to increase network penetration) and data collection (participant reports). In brief, a self-completion questionnaire was used, collecting retrospective data including non-anonymous network tie data (to enable consideration of network reconstruction).

4.2. Initial design and pilot

Questions, questionnaire format and survey logistics were tested via a pilot survey to minimise measurement error, and pilot respondents were invited to debrief interviews to aid further refinement prior to the main fieldwork. The pilot was completed without the snowball element so the survey content could be prepared and tested within a more controlled environment. Those involved in the pilot survey were also invited to volunteer to participate in face to face debrief interviews.

4.2.1. Methods

4.2.1.1. Variables measured

Collection of the data identified above was prioritised (network contacts, social relationship between ego-alter, sentiment of the vaccine-related information shared, and MMR status of children), but limited additional questions were included to guide the respondent through the questionnaire and to enable the comparison of this survey's participation and results with data from other sources, including surveillance data. In order to keep the respondent burden low, the anticipated completion time was set for a maximum of ten minutes.

Network structure variables are the ties in the 'information' network for the respondent, the ties in the 'potential infection' network for each pre-school child of the respondent and the links between these two groups (i.e. an adult alter who is the parent of a child alter). The specific definition provided for the determination of ties in the 'information' and 'potential infection' networks was MMR-related information and measles transmission respectively (see §4.2.1.4). Ties were non-directional.

Other variables on the decision and infection processes are the social relationship between ego-alter, sentiment of the vaccine-related information shared, MMR status of children and also personal knowledge of both measles cases and adverse events attributed to MMR.

The following demographic details were collected for the respondents and the sample children: location, sex, age, ethnic group and education for respondents, the number of children in the respondent's family and age for respondent's child. This includes adult demographics identified as associated with MMR hesitancy in a synthesis of previous studies (Chapter 1).

4.2.1.2. Survey format

A self-completion questionnaire was used, to enable respondents to complete the survey at their convenience and, given the perceived judgemental attitudes of peers [87, 97], to minimise social desirability bias. Data were non-anonymous (to identify reciprocal ties or mutual nodes), but confidential. The questionnaire was prepared in both paper and online formats; both formats contained the same content. All respondents were initially invited to complete and submit the questionnaire online, but paper copies were available within the

recruitment centre for the respondents use if they preferred (paper questionnaires were submitted by posting to the researchers).

4.2.1.3. Participants (recruitment)

The study population was parents with children aged 1-4 years. Childcare settings were identified as suitable recruitment centres. For the pilot, one recruitment centre was used, a child-care setting situated within a Primary Care Trust (PCT) on the shortlist for the main survey (Table 4-3).

All parents of children aged 1-4 years enrolled at the selected childcare setting were invited to participate (subject to ethical requirement on vulnerable individuals) via email from a senior member of setting staff. The email included the survey website address, and survey materials supplied by the researchers (Invitation to Participate and Participant Information Sheet).

4.2.1.4. Instrument development

Question-wording and answer-options (for closed questions) were informed by existing, validated instruments and a synthesis of previous MMR studies (Chapter 1), see Table 4-1. The Questionnaire Appraisal System [259] was used to check questionnaire content during development. A single instrument was used by all respondents – the pilot itself was not used to test alternative question content or presentation, but alternative format were presented to debrief participants for their feedback.

POLYMOD data have evidence of an artificial capping of contact numbers by the spaces provided [234]. Hence, we include a ‘group’ contact option as used in BSCS [235] for children; also for both networks, the online version allowed unlimited entries and the paper version had spaces in excess of twice the expected numbers based on the closest relevant studies (Table 4-1).

Table 4-1 Questionnaire specification

Survey element	Detail / Literature basis
Demographics	Wording from 2011 Census [185]
'Information' network	
Tie definition: proximity	Information shared regarding measles protecting vaccines.
Tie definition: timespan	No time limit (assume ability to recall is associated with ability to influence decision)
Spaces for ties (paper)	Estimates from Marsden, McPherson et al [231, 242] and inferred estimates from DH/COI CITS [21]
Alter's social relationship	Categories from DH/COI CITS [21]
Sentiment of information	Categories informed by responses in qualitative studies [91, 93, 94, 96, 97, 100, 102, 103].
'Potential infection' network	
Sample children spaces	Twice mean dependent children [260]
Tie definition; proximity	"in the same room for 15+ minutes or face-to-face contact" (Appendix 4 [243])
Tie definition; timespan	Weekly (approx. infectious period for measles [28])
Spaces for ties (paper)	mean number of all-age contacts made by 0-4 children [234].
MMR status	Parent recall (good agreement with medical sources [218, 261])

Restrictions were applied to the 'potential infection' contacts request. Inclusion was restricted to other pre-school children: providing clear respondent guidelines to improve data quality, matching the sample inclusion definition and given evidence [234] that this age group mix assortatively or with a generation unlikely to be susceptible to measles [243]. Also contacts were requested for "term-time", given the differences in term time vs holiday contact patterns observed for school age children [262] might also be found in this group.

In addition to the target limit on time-burden on respondents, compact presentation units (single folded paper sheet and lack of scrolling online) were used to reduce barriers to non-completion).

4.2.1.5. Ethical considerations

The study was reviewed and approved by the Imperial College Research Ethics Committee (reference ICREC_12_2_2). Procedures and safeguards relating to informed consent, data protection, collection of non-anonymous data, response confidentiality, protection of

vulnerable individuals and the respect of patient confidentiality are given as an appendix (see Appendix)

4.2.1.6. Data analysis

The first stage in the pilot survey analysis was network reconstruction, matching-up the connecting names for each interacting pair, using the non-anonymous data. After the two networks ('information' and 'potential infection') and the parent child links between them were constructed, the data was anonymised before further analysis. The data analysis is then completed, including calculation of network characteristics, vaccination patterns and respondents' demographic profiles.

4.2.1.7. Participant debrief

Both respondents and management staff at the recruiting centre were invited to participate in face-to-face debriefing on the pilot survey, the respondent debrief used a structured interview for consistency across the sample.

The pilot survey debrief staff interview was planned to include discussion of the survey distribution logistics, survey return logistics, appropriateness of generic instructions to their specific circumstances, if the centre had received any comments regarding the placement of the survey (within any confidentiality constraints) and any other subjects which the interviewee wished to raise.

The pilot survey debrief respondent interview included specific investigation of the validity of questions relating to the network structure, both in the question wording and the answer-collection formatting. This investigation included general, open questions and comments solicited via the presentation of pre-prepared questionnaire alternatives to the interviewee. The debrief interview also investigated the perceived burden on the respondent (including time taken to complete the questionnaire and the practicalities of returning a completed questionnaire), any issues of questionnaire comprehension, any difficulties with providing their answers, the clarity and comprehensiveness of the instructions and support information and their willingness to snowball (and preferred snowball mechanic from a list of alternatives). The respondent debrief interview structure also included more open questioning to enable respondents to volunteer feedback on any other elements of the survey.

4.2.2. Results

Fieldwork for the survey pilot was conducted at a nursery within the Hammersmith & Fulham PCT in July 2012. It was known that the nursery is used by a large number of healthcare professionals (although not exclusively used by those in this occupation sector). Hence predominance of a particular occupational sector was expected to produce a bias in respondent characteristics, such as educational qualifications, which reduces the generalizability of the results of the pilot. However, this clientele is suitable for the pilot study, as they were thought to be able to give constructive criticism during the feedback and debrief process.

Response rate and Sample characteristics

The response rate was 20%, based on the size of the nursery's roll. 80% of respondents responded online, the remaining 20% used paper questionnaires. The demographic profile of responding adults is shown in Table 4-2. The pilot sample is highly educated, as expected given the known bias in the childcare facility's clientele.

Network Structure

The mean number of reported 'information' contacts reported was 1.7 (n=15). Reported contacts included family, friends and healthcare professionals. No transitivity was observed within the reported contacts. The mean number of reported contacts within the 'potential infection' network was 13.1 (n=15). One case of transitivity was observed.

There was 1 'overlap' tie observed, i.e. an 'information' network tie between an adult ego and one of their adult alters and a 'potential infection' network tie between that ego's child and that alter's child.

MMR status and decision context

The recalled uptake of MMR was 85% of children (n=13; censoring data for 2 children aged 13 months or less, so ineligible for routine MMR vaccination, and for 1 respondent who was unsure of their child's vaccination status). 83% of parents had vaccinated their eligible pre-school children (n=14; censoring for the unsure respondent). Eligible siblings shared the same vaccination status.

The sample includes one respondent who had received mixed advice as to whether to vaccinate their child against measles. There was also one respondent who knew of a recent

measles case and another who knew of an adverse reaction attributed to MMR. All the children of these respondents, who were eligible for MMR vaccination, had been vaccinated.

Table 4-2 Pilot survey respondent characteristics

Pilot survey adult respondents		
Sex	Male	13%
	Female	87%
Age-group	35-34 years	47%
	35-44 years	47%
	45+ years	7%
Ethnicity	White / White British	73%
	Black / Black British	7%
	Asian / Asian British	13%
	Other including Mixed	7%
Education	Postgraduate	60%
	Graduate	27%
	A-level	7%
	Other	7%
Children aged under 5 years	1 child	93%
	2 children	7%
		Base: all adult respondents (n=15)
Pilot survey child sample		
Age	Under 1 year (0 – 11 months)	7%
	1 year (12 – 23 months)	33%
	2 years (24 – 35 months)	20%
	3 year (36 – 47 months)	13%
	4 years (48 – 59 months)	27%
		Base: all sample children (n=16)

Debrief participation

A third of respondents offered feedback on the pilot material, and three respondents volunteered to participate in face to face structured debrief interviews. The childcare facility management also agreed to be interviewed, post-survey.

4.2.3. Discussion

The pilot survey has demonstrated that the survey could be successfully administered and is capable of producing data that can be analysed to produce characteristics for the parameterisation of mathematical models exploring the influence of social networks on vaccination decisions and on the resulting potential for vaccine-preventable disease outbreaks.

The nursery management reported no disruption to their operation, which is desirable for co-operation from these settings as recruitment centres. Both formats of the survey (paper and online) were used, the technical success of the latter is essential for the implementation of the snowball element of the full survey. The 20% response rate is acceptable and can be used in the power calculations for the main survey to estimate the numbers that must be approached to obtain the required sample size.

It was satisfying that the proportion of respondents with unvaccinated children (who were old enough to be eligible for MMR) was in line with the most recent MMR1 uptake level for the PCT (81% [38]) allaying fears that these individuals who are acting contrary to the social norm may be reluctant to participate.

The number of 'information' network contacts was lower than prior estimates (made using Marsden, McPherson et al and DH/COI CITS [21, 231, 242]). The number of 'potential infection' pre-school contacts exceeded the mean number of all-age contacts reported in POLYMOD [234]. This latter finding raised concern that the paper version may not have sufficient space for these answers to avoid a capping phenomenon.

However, the contact-listing questions were not both answered by all respondents; which was of concern given the primary objective of the survey is the discovery of these contacts. This subject was explored in detail during the structured debrief interviews and was the subject of the other feedback comments. These clarified barriers to eliciting a response to both these questions. Using four pre prepared alternative questionnaire layouts (including the pilot version), the debrief interviews also unanimously identified a revised questionnaire wording which is more likely to prompt fuller disclosure of all recollectable contacts (who meet the survey definitions) (see §4.3.1.5)

4.3. Full survey

The full survey included the snowball mechanic. Following the pilot some elements of the questionnaire were amended (wording and layout) and administration logistics were refined, these differences are described below.

4.3.1. Methods

4.3.1.1. Variables measured

The questionnaire entered fieldwork with the same variables being measured as in the pilot.

During the survey fieldwork period, a large outbreak of measles in Wales was reported in the national media. Furthermore, the Department of Health launched a national MMR catch-up campaign in late April 2013, primarily targeting children aged 10-16 years old (Chapter 1).

This context may increase urgency with which parents seek scheduled vaccinations, prompt previously vaccine-hesitant parents to reconsider and present unvaccinated children out of schedule, and increase the vigour with which HCP encourage adherence with the recommended vaccinations (on schedule or belatedly). The absolute timing of any MMR vaccination therefore became a new variable of interest and, from mid-May 2013 a revised questionnaire was used. The question regarding the MMR status of the respondent's children was amended to distinguish vaccination decisions that were completed prior to these events and those that may have been influenced by them. Questionnaires already in field with recruitment centres were not withdrawn and replaced.

4.3.1.2. Survey format

As with the pilot, a self-completion questionnaire was used, prepared in paper and online formats, with both formats containing identical questions. Further details on the specific format(s) presented to individual potential participants are given below (§4.3.1.3). A snowball mechanism was included; respondents were asked to forward the online access details to other adults whom they had included as answers within the questionnaire. No incentive was offered for snowball participation, as the numbers of contacts was a measured variable of the survey itself.

4.3.1.3. Participants (recruitment)

The study population was parents with children aged 0-4 years, widened from 1-4 years in the pilot (based on advice regarding practicalities from the management debrief). Participants were recruited via two channels – 'direct' recruitment (via recruitment centres) and 'snowball' sampling (recruited online by the existing participants, being one of their listed 'information' network contacts, so not necessarily the parent of a child aged 0-4).

Direct recruitment was undertaken in regions of epidemiological interest. These were defined as those PCTs which reported MMR1 uptake figures in the lowest decile for England in any of the following COVER reports (the most recent published prior to Ethical Approval submission): annual reports for 2008-9, 2009-2010 and quarterly reports for 2010-2011 [35, 36, 181-184]). 33 PCTs fall into this shortlist definition (Table 4-3). (PCT names and boundaries used are as were in operation in March 2011). The number of recruitment centres approached was largely capacity-driven and focussed into a limited number of PCT areas, selected from this shortlist. (No geographical restriction was placed on snowball recruited respondents.)

Candidate settings were identified via the National Association of Family Information Services and management approval was required before they were used as a recruitment centre. Participating recruitment centres describe themselves as various types of childcare (nursery, playgroup, kindergarden etc), but no childminders were approached (as local authorities advised that significant additional approvals would be required to do so). Given this funnel of approval logistics, the sample of recruitment centres was a sample of convenience and coverage bias may therefore be present, although unlikely for children age 3-4 years old [263].

Table 4-3 Shortlisted PCT

In alphabetical order with mean MMR1 uptake 2008-2011 (weighted by annual eligible population)

PCT	MMR1		PCT	MMR1	
Barking & Dagenham PCT	79.5%	♣♥♠	Hartlepool PCT	84.7%	♦
Barnet PCT	86.0%	†	Havering PCT	80.3%	‡
Bexley Care Trust	75.8%	‡♣♥	Herefordshire PCT	84.9%	♠
Brent Teaching PCT	74.8%	‡♣	Hounslow PCT	76.6%	‡♣♦♠
Bristol PCT	83.2%	†	Islington PCT	80.6%	‡♦
Bromley PCT	82.4%	‡	Kingston PCT	84.6%	♠
Camden PCT	74.2%	‡♣♦♥♠	Kensington & Chelsea PCT	85.5%	♣♦♥♠
City & Hackney Teaching PCT	74.8%	♣♦♥♠	Lambeth PCT	79.1%	†‡♦
Coventry Teaching PCT	89.6%	†	Lewisham PCT	77.9%	†‡♣♦♥♠
Croydon PCT	81.7%	†♣♦♥♠	Newham PCT	82.3%	♥
Dorset PCT	85.5%	♥♠	Nottingham City PCT	81.0%	†
Ealing PCT	82.7%	†	Richmond & Twickenham PCT	83.9%	‡♣♦
Enfield PCT	77.2%	†‡♣♦♥♠	Southwark PCT	78.5%	†‡♣♦♥
Great Yarmouth & Waveney PCT	84.0%	†	Surrey PCT	80.4%	‡♣♦♥♠
Greenwich Teaching PCT	74.9%	†‡♣♥	Sutton & Merton PCT	82.7%	♦
Hammersmith & Fulham PCT	75.8%	†‡♦♠	Wandsworth PCT	84.6%	♥♠
Haringey Teaching PCT	75.8%	♣♥♠			

Key

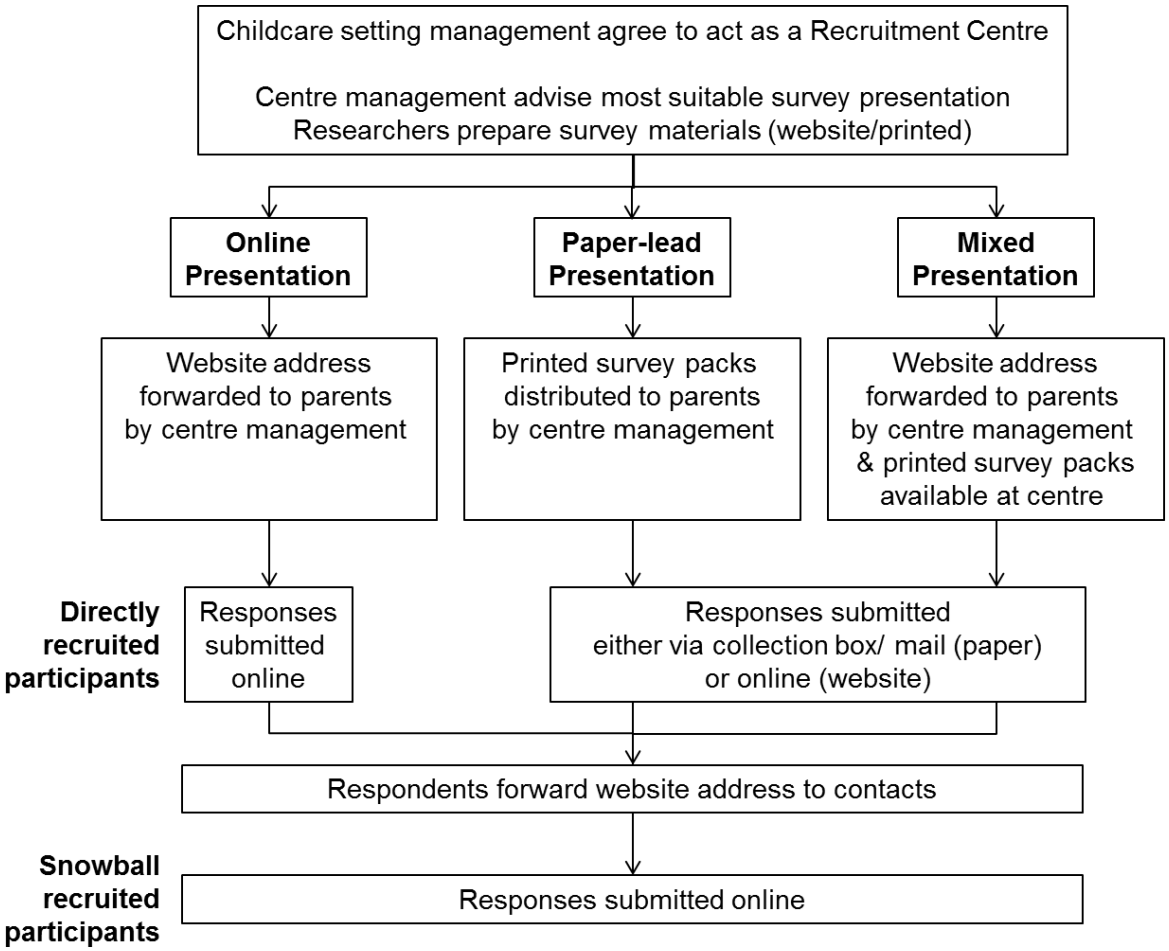
† lowest decile in 2008-09 [35]
‡ lowest decile in 2009-10 [36]

♣ lowest decile in 2010-11 quarter 1 [181]
♦ lowest decile in 2010-11 quarter 2 [182]
♥ lowest decile in 2010-11 quarter 3 [183]
♠ lowest decile in 2010-11 quarter 4 [184]

Three survey presentation routes were used: online, paper-lead and mixed (Figure 4-3). The presentation used at a centre was selected by the setting management, to fit-in with how they normally communicate with parents and to incorporate the management's experience of getting parents to respond to information requests.

Centres choosing the online presentation were provided with a clickable link to the survey website (hosting all the survey documentation) and a pro forma invitation email that they could forward or include as part of e-newsletters etc. Centres choosing the paper format were provided with sufficient printed survey packs for all enrolled children. Centres choosing to use a mixed format were supplied with the same information as online centres, to use in the initial communication with parents, and a smaller number of printed survey packs that parents could collect from the setting.

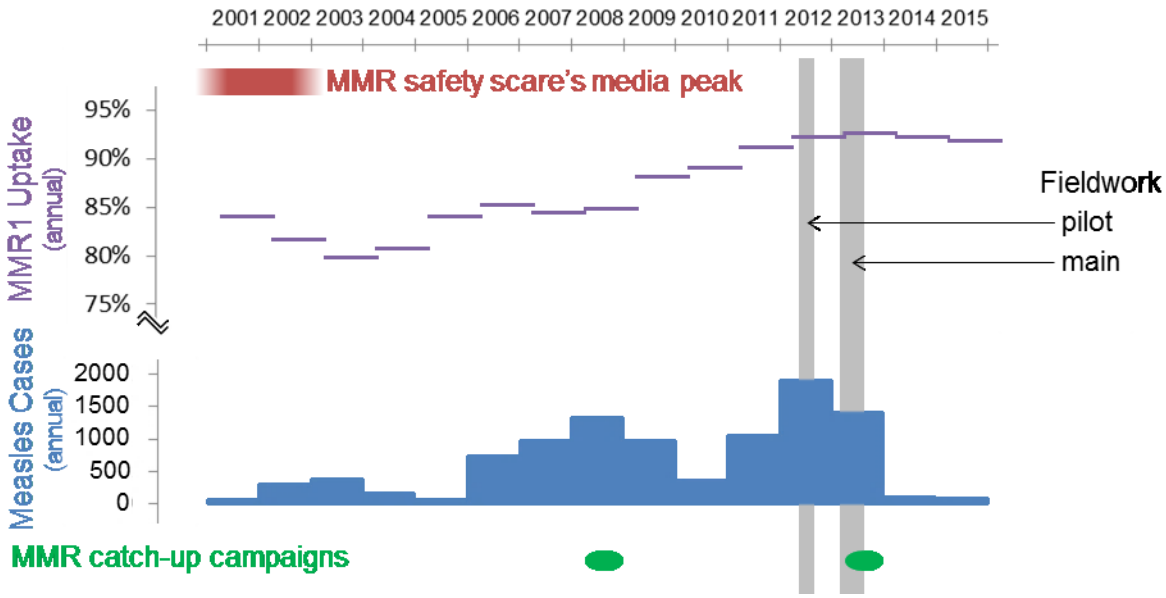
Figure 4-3 Survey presentation routes



The fieldwork period is shown in Figure 4-4, together with selected context for MMR (annual MMR1 uptake and catch-up campaigns) and measles (annual confirmed cases) in England,

for the time since the peak of the MMR safety-scare media coverage (in 2001-2), details of which were given in Chapter 1.

Figure 4-4 Fieldwork period in context



MMR1 uptake *NHS Digital* [42], Confirmed measles cases *PHE HPA* [48, 49]

4.3.1.4. Power calculation

We consider the pattern of vaccination opinions on the 'information' network. The sample size calculated is that required to detect a 10% increase in the observed intra-dyad agreement value (as used in Eames [168]) vs that expected under random opinion allocation, at 5% significance and 80% power (using sample size calculator [264]).

We assume the proportion of pro-MMR nodes is the same as the MMR1 uptake, calculating the required sample size for the inter-quartile uptake values from the shortlisted PCT (Table 4-3). To convert from dyad sample size to node sample size requires the mean ties per note (uses handshaking lemma); we use the value from the pilot, 1.70, which provides a more cautious sample size than the (higher) prior estimates [21, 231, 242]. Similarly, to estimate the number of potential participants to approach to deliver the calculated sample size, a more conservative response rate than the pilot survey (20%) has been applied, also any contribution from the snowball is excluded:

Node sample size	310	251
Estimated response rate for survey	15%	15%
Potential participants to be approached	2067	1674

The number of recruitment centres approached (i.e. the sampling unit for measurements using network units) will be largely capacity-driven and it is acknowledged that the likely sample size of networks (<20 networks) is insufficient to analyse a dataset of network-level measurements.

4.3.1.5. Instrument development

The questionnaire was amended, based on the debrief interviews and the analysis of results vs prior estimates.

The alternative contact-collection layout preferred by debrief participants was adopted, and filled examples for the child data answers were provided alongside the answer grid. Additional guidance on question completion was also provided, using information provided by pilot respondents. Informed by pilot responses, spaces for fewer children were included but more spaces for 'potential infection' network contacts were provided per child. Contacts that children made at the childcare setting were collected separately from those made elsewhere, using these context prompts to reduce recall bias. For adults and children, an additional answer option "no contacts" was added to differentiate zero contacts from an unanswered question.

The list of questions (Box 4-1) is common to both formats; examples of the finalised paper are in the appendix (see Appendix).

The survey materials were prepared in English and this language was used on the website. It is acknowledged that language choice may introduce both coverage and sample bias. The potential scale of sampling bias was assessed via the analysis data regarding use English as Addition Language (EAL) in the survey population (see Appendix) and found to be a concern. Hence, centres were asked which language(s) they used to communicate with parents and if use of English would restrict survey access (and the most common first-languages for any parents affected, so appropriate translation services could be engaged). Printed materials were distributed in the language used by the centre and the English-language "Invitation to Participate" included instructions for those with limited English-language literacy skills to request translated materials.

Box 4-1 Survey questions

Full survey questionnaire: list of questions

Please tell us about yourself

What is your name?

What is the postcode of your home?

What is your sex?

How old are you?

What is your ethnic group?

Which of these academic qualifications do you have?

Please tell us about your family

How many children (born in 1997 or later) do you have? †

Do you have any children aged under 5? *If 'no' go to* ✸

[For each under 5]

What is their name?

How old are they?

Have they ever received an MMR jab? ‡

Please tell us the names of the pre-school children that your child mixes with in a typical week (during school term-time), include weekdays and weekends.

- Children who your child mixes with at childcare
- Children your child meets in other places

[Thinking about your MMR jab decision, for a nominated child]

Please tell us the names of all the people with whom you can remember discussing vaccinations to protect children against measles (e.g. MMR) including giving/receiving advice, information or opinions on this subject.

How do you know this person?

Where do they live?

Which of these descriptions is the closest match to this person's advice or opinion?

Is this person the parent of a pre-school child? *If no parents go to* ✸

[For every person marked as having pre-school child(ren) in the previous question]

Do they have a child (or children) who was included in your answer to [pre-school children that your child mixes with]?

✸ Have you, or anyone you know personally, had measles recently?

Have you, or anyone you know personally, had a serious adverse reaction attributed to MMR jab?

† children born in 1997 or later would turn 16 in the fieldwork year

‡ options provided were revised mid-May 2013

4.3.1.6. Ethical considerations

This survey was reviewed and approved by the Imperial College Research Ethics Committee (reference ICREC_12_2_2). The considerations outlined for the pilot were maintained, regarding informed consent, collection of non-anonymous data, data protection, response confidentiality (including additional procedures for the confidential return of paper questionnaires), protection of vulnerable individuals and the respect of patient confidentiality (see Appendix).

4.3.1.7. Data analysis

4.3.1.7.1. Processing and cleaning

Paper questionnaire responses were transferred into the same electronic database as the online responses. To facilitate analysis, missing answers were imputed where these were unambiguously determined by other answers. Independent double-entry was used for a random 10% of the returned paper questionnaires, and the resultant electronic records were cross-checked and compared for accuracy of data-transfer (proportions of matching answers / Cohen's Kappa values). We accepted only 100% agreement for questions with pre-determined answer options. Subject to an acceptable level of accuracy in transfer and cleaning, single-entry was used for the remaining questionnaires (else further random samples would undergo double-entry and checking until acceptable accuracy was achieved).

4.3.1.7.2. Network reconstruction

Subject to sufficient responses per recruitment centre (including snowball) networks are constructed from the data prior to anonymisation (by matching-up names for each interacting pair). All individual names were assumed to be unique, unless there was evidence to the contrary from location and one other data element. As with the data transfer a sample is processed using double-entry, checked for agreement before single-entry was used. After any network reconstruction is completed the data are anonymised before further analysis.

4.3.1.7.3. Ego-centric data analysis including inference of unreported vaccine-opinion and MMR1 status

The data analysis is then completed, including calculation of network characteristics, and the patterns of vaccination status and of vaccination information transfer. The comparison of the properties of the survey data with those of networks produced by standard network algorithms is reported later (Chapter 5).

To enable calculation of intra-dyad agreement from ego-centric data, the vaccine-opinion of members of the 'information' network and the vaccine status for contacts within the 'potential infection' networks are inferred, as far as possible.

We infer the measles-vaccination opinion of the respondents' advisors, based on the recalled sentiment of the communication shared with the respondent.

We can make an inference of the MMR status for the some of the contacts of the sample children who were not themselves included in the sample, namely those named contacts who are both individually named by the respondent and were identified as being the children of an advisor to that respondent. We infer their MMR status based on the advice their parent gave the respondent, i.e. the inferred opinion for the advisor (Figure 4-5). Additionally we have reported data on the MMR status of contacts who are siblings to that child.

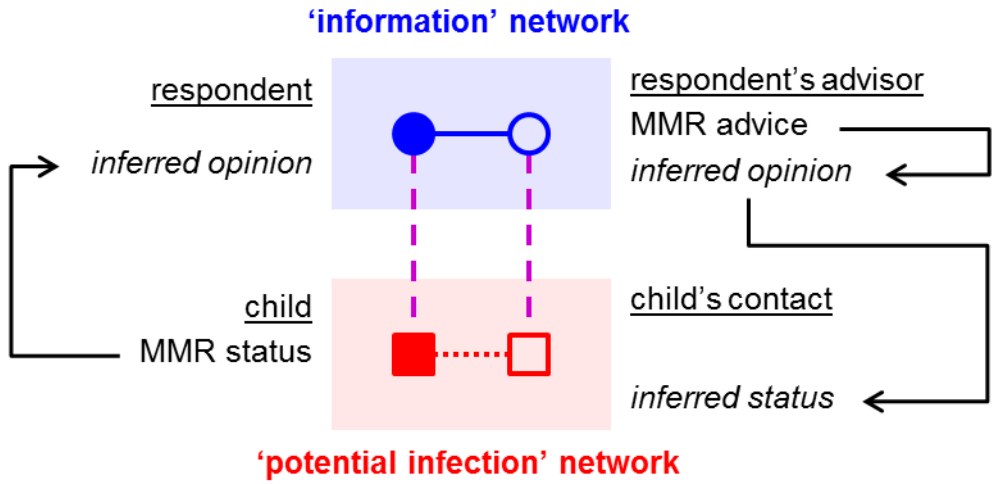
We infer the measles-vaccination opinion of the respondents based on their reported behaviour in vaccinating their child (where a respondent has children with differently categorised vaccinations we assume the respondent has the opinion corresponding to their vaccination behaviour with the younger child).

Assuming that vaccination-supporters seek vaccination at the earliest possible opportunity within the routine schedule, the MMR status of children aged 14 months can be used to infer vaccine-support for the respondents. Vaccine-hesitancy may be expressed as a delay in presentation of the child for vaccination or non-presentation; hence (mindful of the COVER methodology) we infer that if a child is still unvaccinated by 24 months then the parent intends to "never" have them vaccinated with MMR, and vaccinations occurring between 14-24 months have been purposefully deferred.

Figure 4-5 Inference of MMR opinion and MMR1 status

Nodes within networks discovered by survey

with associated MMR-related data



Notes:

* if mulitple offspring render a respondent's opinion inferrence abiguous, the youngest child with an uncensored MMR status is used as the source

* for contacts which are siblings of the child their reported status is used

Legend

'information' network tie

'potential infection' network tie

parent-child family tie

inference

4.3.2. Results

4.3.2.1. Implementation

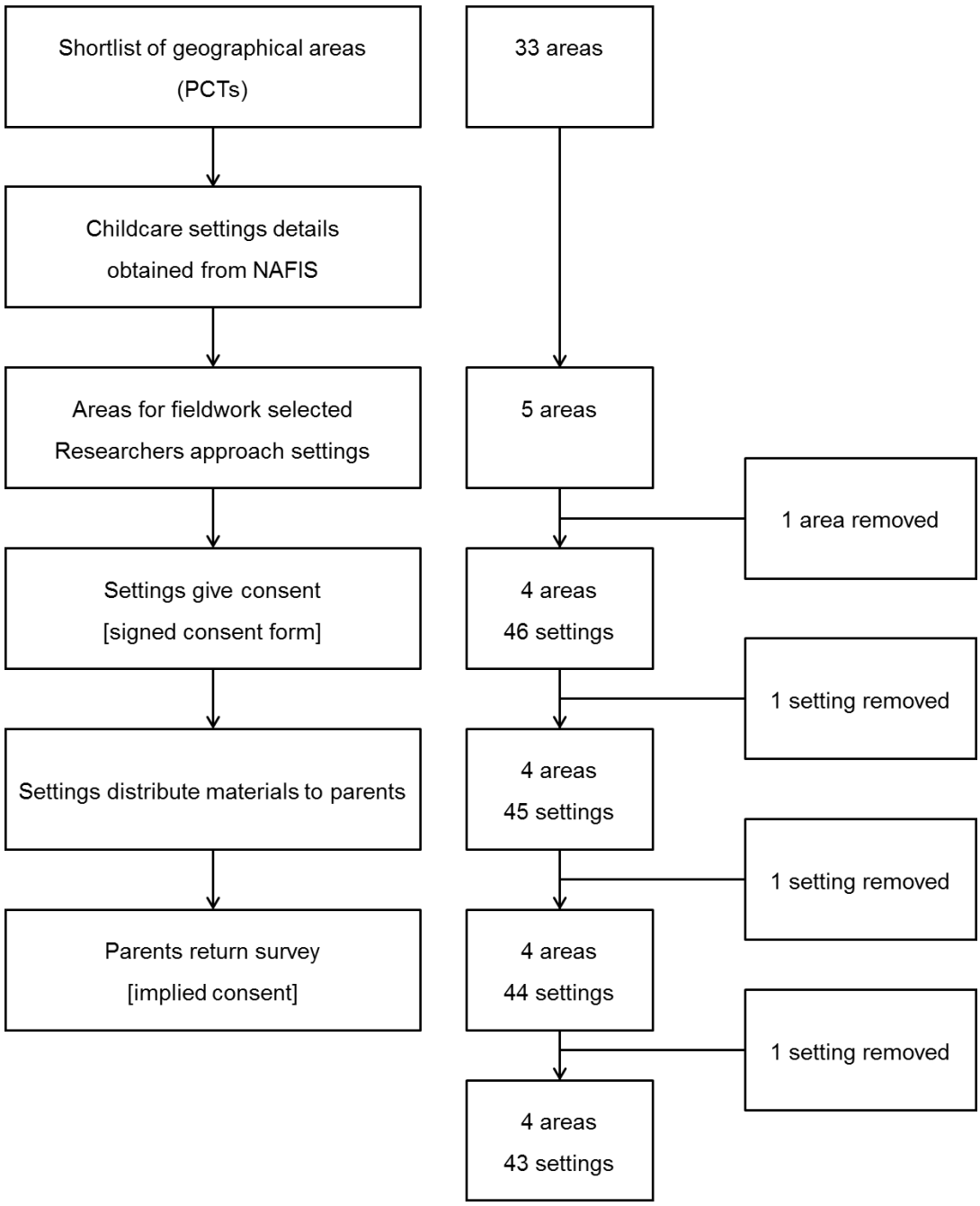
4.3.2.1.1. Setting recruitment

The survey fieldwork was conducted in the following PCTs: Camden, Ealing, Enfield, Great Yarmouth & Waveney (GY&W) and Wandsworth. The final fieldwork areas were purposively chosen to reflect the London bias of lower MMR1 uptake (with both inner and outer London areas included) but to also include a non-London region.

Ealing PCT questionnaires were all placed prior to the questionnaire change to incorporate absolute timing of MMR uptake (Q11), other areas used the amended version of the questionnaire. However the analysis does not include any data from questionnaires placed in Ealing PCT, because there was a small local measles outbreak [265] and several recruitment centres withdrew co-operation (including where individual setting management had agreed to participate but were subsequently over-ruled at group level, and so the settings withdrew their consent). Three settings in other areas were removed from the survey after having given consent to act as recruitment centres and receiving questionnaires for distribution (one centre withdrew consent following a change of management, one centre distributed materials incorrectly, one centre accidentally destroyed returned questionnaires before they could be collected for analysis).

The potential parent participants are a thus a sample of convenience. A summary of the recruitment funnel for recruitment centre funnel is shown in Figure 4-6 together with the geographical distribution of the participating recruitment (the total enrolled children at the participating recruitment centres represent 3.3% of the pre-school-age population of the four areas [211]).

Figure 4-6 Recruitment funnel

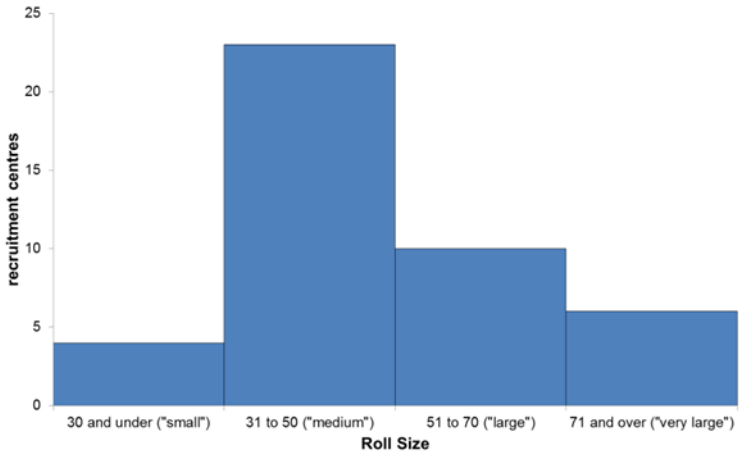


Final sample recruiting: 4 areas, 43 settings (recruitment centres)

	No. of settings	Enrolled children	Population age 0-4 years
Camden PCT	8	317	13168
Enfield PCT	6	320	24513
GY&W PCT	11	731	11758
Wandsworth PCT	18	996	21670
Total	**43**	**2364**	**71109**

Settings with a range of enrollment sizes were included in the recruitment centre sample (Figure 4-7). The median roll size of recruitment centres (48 children, n=43, minimum 19, maximum 184) is between the mean weekly attendance for all full day care settings and sessional day care settings, 59 and 44 respectively, in England [266].

Figure 4-7 Roll sizes of recruitment centres



Base: all recruitment centres (n=43)

Recruitment centres utilised all available format offers (Table 4-4). The majority (corresponding to 56% of gross enrolment) of potential respondents were primarily offered the paper questionnaire, with online (parents only given web access details) and mixed-format (parents notified electronically, but paper questionnaires available within the setting) accounting for 26% and 18% respectively.

Table 4-4 Survey formats offered by recruitment centres to parents

	Paper lead	Online	Mixed	Total
Camden PCT	4	4	0	8
Enfield PCT	5	0	1	6
GY&W PCT	7	1	3	11
Wandsworth PCT	7	9	2	18
Total	**23**	**14**	**6**	**43**

Base: all recruitment centres (n=43)

All settings in the final sample communicated with their parents using English, so all printed materials were produced in English. Although several settings indicated that they had a minority of parents who did not have English as their first language (with a range of European and Asian languages used), they did not anticipate that these parents would be

prevented from participation through use of English materials. The invitation to request materials in other languages was not taken up by any potential participants.

4.3.2.1.2. Participant response

A total of 170 questionnaires were returned to researchers. 169 questionnaires came from participants recruited directly by the centres, 1 questionnaire via the snowball. Hence the conversion rate from directly recruited participants to a completed snowball questionnaire was 0.6%. 2 returned questionnaires have been removed from the analysis (the only questions answered were those regarding the respondent demographics), so the final analysed sample size is 168 adult participants – all directly recruited.

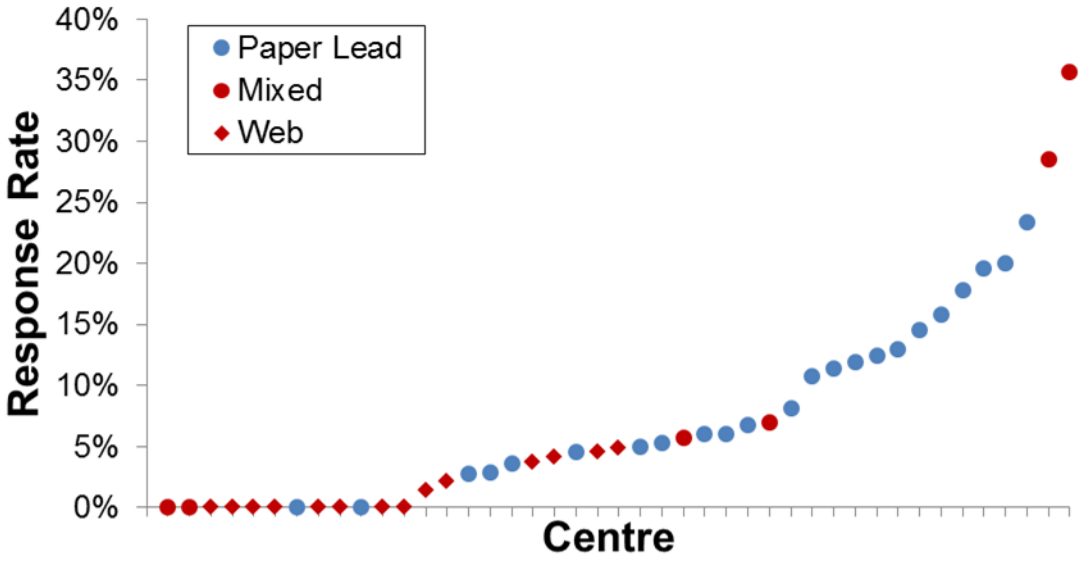
45 questionnaires were submitted online (27%, n=168), of the remaining 123 paper questionnaires, 119 (71%, n=168) were submitted via the recruitment centre collection boxes and 4 (2%, n=168) were submitted directly to the researchers.

There are 212 pre-school children in the sample (reported offspring of the adult respondents, not restricted to those attending the recruitment centre).

The response rate was at least 7.1% (due to parent-list confidentiality, enrolled children is used as the denominator not parents, hence this rate is a lower bound, due to sibling co-enrollment - 11% of parents in the sample have co-enrolled offspring). Subject to the same denominator caveat, parental response was lowest for those with child(ren) are enrolled in a setting with roll count of 30 or below (1.9%) and for those offered only the online survey (1.6%).

Considering the corresponding response rate measured within recruitment centre, the median response by centre was 4.9% (range = 0.0% to 35.7%, n=43). 12 centres did not produce any analysable response (28% of recruitment centres, with a combined child enrolment of 719, 30% of the total). The poorly-responding centres are more strongly associated with the (online) survey presentation (Kruskal-Wallis H=15.10, 2 df, $p<0.001$) than centre size (Kruskal-Wallis H=3.20, 3 df, p=0.36) or location (Krusal-Wallis H=3.43, 3 df, p=0.33). However the best-responding centres used the mixed presentation (Figure 4-8), which was web-led (and 94% of responses received from centres with mixed presentation were submitted using the online questionnaire), which allays concerns that the website itself might be depressing response levels.

Figure 4-8 Response rate by centre by presentation



4.3.2.1.3. Sample characteristics: respondents and their children

The demographic composition of the sample – adult respondents and their children in the sample – are given in Table 4-5 and Table 4-6.

As with the pilot survey, the majority of adult respondents are female. The range of ages represented is greater than the pilot, but skewed older. As expected the educational achievement is less skewed towards postgraduate level qualifications than the pilot, but still contains more graduates than the population (adults with dependent children in survey areas [267]). Black/Black British parents are under-represented in the sample [190].

Children aged under 1 year old are under-represented in the sample (as a proportion of all under 5s [191]), although well distributed between ages 1 to 4 years (inclusive). The majority of children in the sample are first-born for the respondent. The geographical location of the children differs little from that of the adult sample. 90% of the children attended childcare (n=206); fewer under 1 year olds attend childcare (17%, n=18) compared to the older age groups (97%, n=188). There is strong evidence that formal childcare attendance by sample 0-2 year olds differs from the UK average (35% in 2013 [263], $z=11.03$, $p<0.01$), but not for the older children.

There is strong evidence that sample has a higher uptake of MMR(1) than reported by the COVER surveillance. More details are given in §4.3.2.8.

Table 4-5 Sample characteristics – adults

Adult respondents

Sex	Male	10.8%	
	Female	89.2%	
Age-group	18-24 years	4.2%	
	25-34 years	32.7%	
	35-44 years	60.1%	
	45+ years	3.0%	
Ethnicity	White / White British	88.1%	White (British) 72.0%, White (Other) 16.1%
	Black / Black British	3.6%	
	Asian / Asian British	6.0%	
	Other including Mixed	2.4%	
Education	Postgraduate	28.1%	
	Graduate	44.9%	
	A-Level	16.8%	
	5+ GCSE	3.6%	
	1-4 GCSE	4.2%	
	None	2.4%	
Children	1	52.4%	
	2	38.0%	
	3	6.6%	
	4	3.0%	
Children under 5	1	74.4%	
	2	25.0%	
	3	0.6%	

Base: all adult respondents censoring for missing answers; n=166-168

Table 4-6 Respondent characteristics – children

Child sample

Age	Under 1 year (0-11 months)	9.0%	0-5 months 3.3%, 6-11 months 5.7%
	1 year (12-23 months)	25.5%	12-17 months 10.4%, 18-24 months 15.1%
	2 years (24-35 months)	23.1%	
	3 years (36-47 months)	21.2%	
	4 years (48-59 months)	21.2%	
Ordinal (known twins have shared ordinal)	1st	61.0%	
	2nd	31.0%	
	3rd	5.7%	
	4th	2.4%	
Centre PCT	Camden PCT	14.6%	
	Enfield PCT	11.8%	
	GY&W PCT	41.0%	
	Wandsworth PCT	32.5%	
Childcare attendance	Yes	87.3%	
	No	9.9%	

Base: all sample children censoring for missing answers; n=206-212

MMR Vaccinated	Yes	97.8%	

Base: all sample children aged 24 months or older, censoring for missing answers; n=136

4.3.2.2. Inferred opinions and vaccination status

4.3.2.2.1. Vaccination status inference

Vaccination status is inferred for all contacts for whom it is possible (not age-censored as contacts' age is not a measured variable); we note this methodology has a bias toward "vaccinated" status. On this basis, we have a vaccination status for 202 sample children (95%, n=212) and an inferred vaccination status for 213 contacts (17% of the total contacts who are named or siblings, n=1255).

We note that non-availability of inferred vaccination status for 83% of the contacts may introduce additional sample bias, and there is strong evidence that the available ego and alter samples are taken from different populations (ego 92% vs alter 84%, χ^2=6.968, 1 df, p=0.01), which is unexpected as they are drawn from the same demograph. If we are more cautious with assumptions of statistical independence, and so censor sibling contacts, the evidence for sample heterogeneity remains (alter 80%, χ^2= 9.707, 1 df, p=0.02). Furthermore this heterogeneity is not driven by the inclusion of children below routine vaccination age (with non-assortative mixing) as excluding this group does not weaken the evidence (alter 86%, 16.420, 1 df, p<0.001).

4.3.2.2.2. Vaccination opinion inference

'Information' network alters

We infer the measles-vaccination opinion of the 'information' network contacts based on the sentiment of the advice they gave the respondent. The inferred measles-vaccination opinion of the 'advisors' (i.e. 'information' network alters) uses a binary measurement of either supporting adherence to the scheduled MMR1 vaccination or opposing adherence (via deferral or rejection) (see §4.3.2.7.1). Therefore it is possible to infer opinion for 421 advisors (92% of all advisors, n=456).

'Information' network egos

We infer the measles-vaccination opinion of the respondents based on the vaccination status of their (youngest uncensored) child with vaccination behaviour categorised as "timely", "late" or "never". We use the longitudinal data (for those children in the pre-catch-up cohort, as used in §4.3.2.8) censoring children still under 24 months at data collection (n=136) to distinguish "late" and "never" behaviour. (We note applying this method to this type of data has a bias towards "timely" vaccination.)

124 respondents (74% of all respondents, n=168) have children in this subset. 94% vaccinated in a "timely" manner, 4% were "late" vaccinators and 2% "never" intended to vaccinate the child with MMR. To mirror the binary measure of advisor opinion, we collate "late" and "never" vaccinators as non-adherents to the recommended schedule (6%) and "timely" vaccinators as adherents. There is little evidence for differences in inferred opinion by respondent's demographics: sex, age, education, ethnicity, location (PCT of centre) and number of offspring (Fisher Exact, significance measured at p<0.05).

‘Information’ network sample

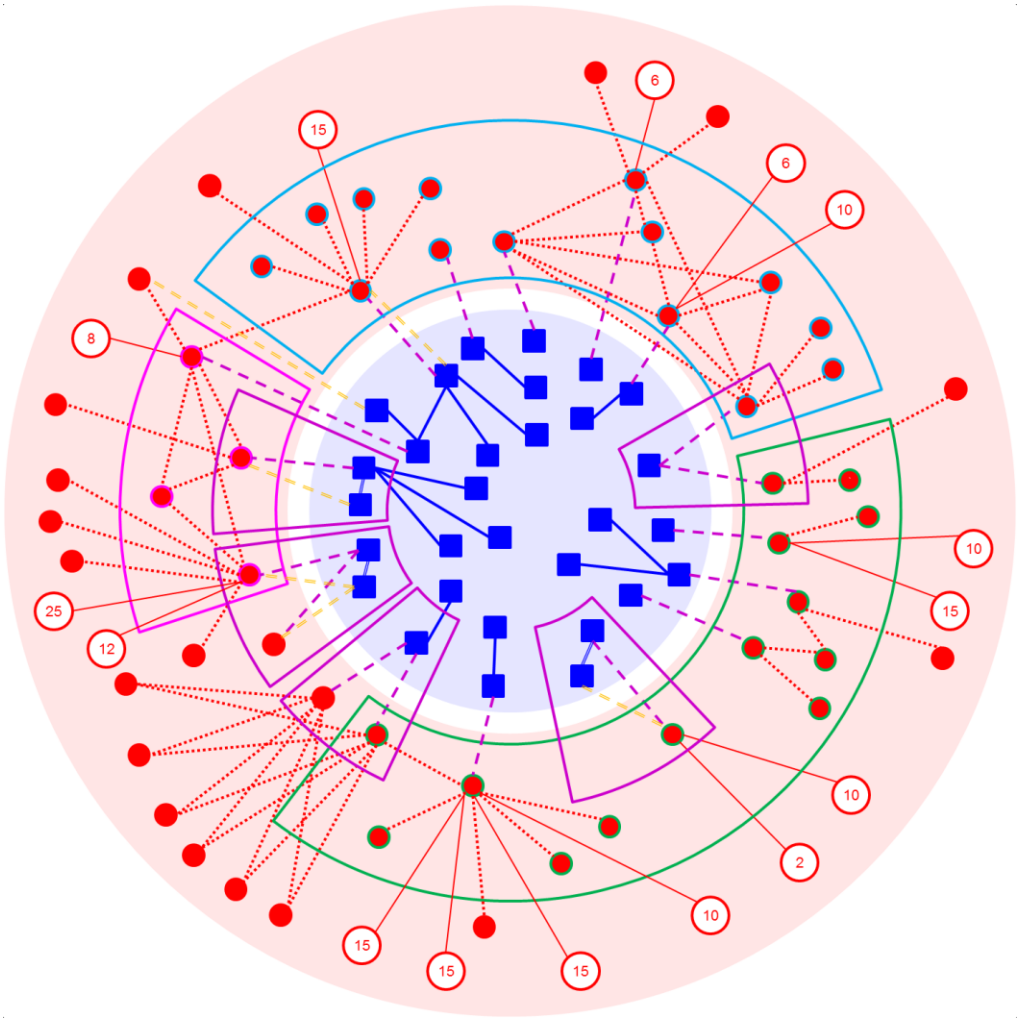
We have vaccination opinions for 92% of all advisors (n=456) and 74% of all respondents (n=168). There is very weak evidence to reject a null hypothesis that the samples of respondents and their advisors are drawn from populations with the same profile of binary vaccination opinion (χ^2=1.807, 1 df, p=0.21), this is despite the presence of HCP within the latter group and its cross-generational make-up.

4.3.2.3. Full networks structure

Given the response by centre (both absolute value and skew, §4.3.2.1.2) it was not thought reasonable to reconstruct networks for all centres. Similarly, all the network structure measurement is restricted to ego-centric data analysis.

An example of the linked networks from the best-responding centre is shown in Figure 4-9, unreported but inferrable ties (between siblings and within complete network subgraphs) are not shown. 14% of the childcare setting nodes (based on total enrolled children) remain cryptic under this sample, and are not shown in Figure 4-9. We note the proportion (92%) of child notes that are located within complete network subgraphs (all children within the same room at the childcare setting and each set of grouped contacts). We can estimate the sampling error for this measure from under-reporting the setting subgraphs (the cryptic nodes above), but the data does not allow de-duplication of unnamed nodes within the non-childcare groups.

Figure 4-9 Example reconstructed linked networks from one centre



Key

- Extent of 'information' network
- Extent of 'potential infection' network
- Extent of family unit
- Extent of setting room (coloured by room)
- Adult 'information' network node
- Single child 'potential infection' network node (edge colour-coded for setting room)
- n — Grouped child 'potential infection' network nodes (represents the complete subgraph size n, K_n)
- Adult-adult 'information' tie (double line indicates between partners)
- Child-child 'potential infection' tie (solid line to grouped nodes)
- Offspring-parent family link, adult is survey respondent
- Offspring-parent family link, adult is alter of survey respondent(s)

4.3.2.4. ‘Information’ networks

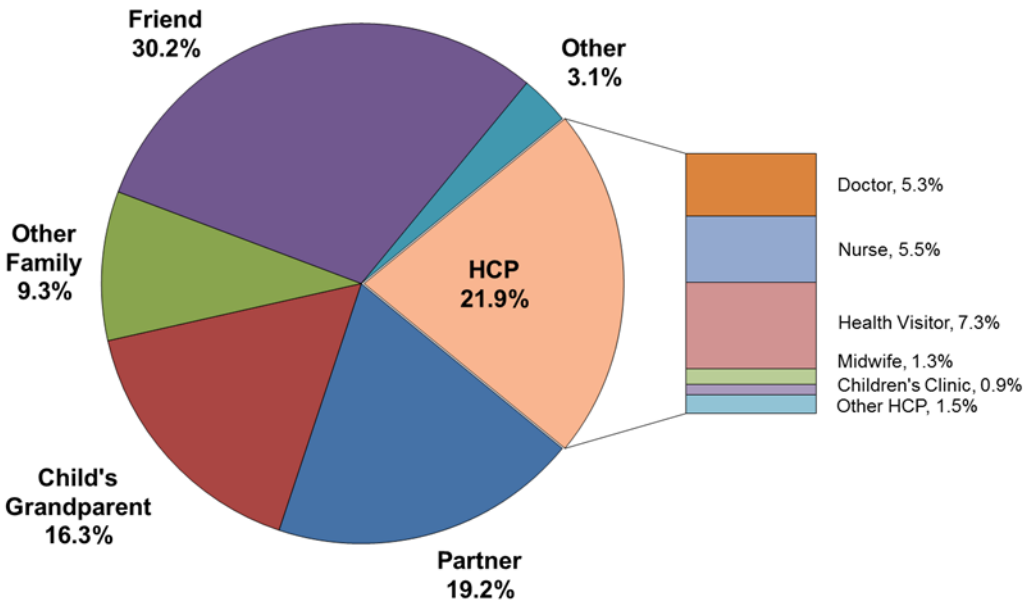
Ego-centric information network data are available for 161 respondents (96%, n=168), with a total of 456 alters reported.

4.3.2.4.1. ‘Information’ network nodes

Social relationship of the alter nodes to the ego

Relationship categorisation (Figure 4-10) is available for 453 alters (99%, n=456). The majority of advisors are not qualified healthcare professionals (HCP): friends represent the largest category of advisors (30%, n=453, 13 prompted categories), and a total 45% of advisors were family members.

Figure 4-10 Advisors sample by relationship with respondent



Base: all advisors censoring for missing answers (n=453)

Spatial relationship of alter nodes to the ego

We examine the relative location of the advisors as the information network ties are not necessarily dependent on physical proximity. The location information supplied for respondents and their advisors was used to infer the PCT for each [268]. However given the range of spatial magnitudes of the survey PCT (2180 - 54387 hectares [194], with the three

London PCT in the smallest decile nationally), we have also categorised estimates for geodesic ego-alter distance (<10km, >10km) to compensate for edge effects, and categorised estimated driving time (<15minutes, >15minutes) to compensate for different built environments. 14% of advisors (n=391) have vague "London" locations (and advised London residents), and they form the majority of insufficiently precise locations (under the PCT metric, 22% of advisors, n=391).

The majority of the information network alters are reside in the same PCT as their ego (71%, n=301) or the adjacent PCT (further 17%, n=301). The proportion of alters within the same PCT is significantly different stratified by centre PCT (χ^2=20.296, 3 df, p<0.001) with the values for the spatially-smaller London PCTs lower than GYW PCT. Including a geodesic categorisation, 73% of advisors are located either in the same PCT or within 10km of the ego's address (n=334). Under the driving-time metric, 74% of information network alters are within 15 minutes of the ego and this proportion is not significantly different across the survey areas (χ^2=0.237, 3 df, p=0.97).

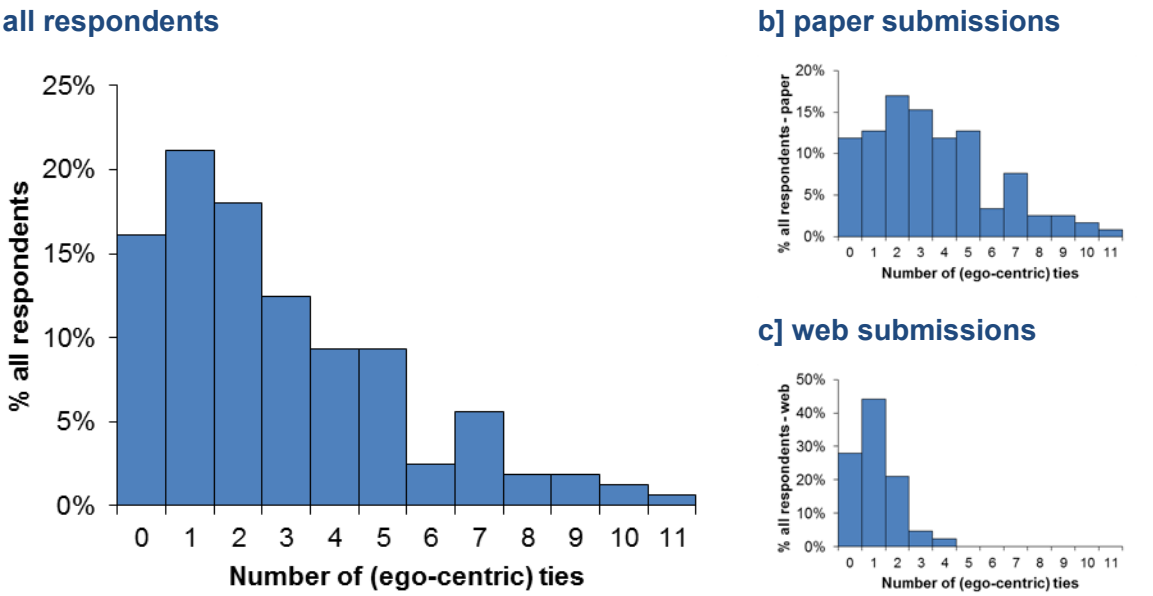
Thus the observation that majority of the information network ties are contained within the closer of binary categorisations of physical location is robust under all examined metrics (although the temporal stability of this observation over the vaccine-decision process is unknown). We also note the corollary that over 25% are located further away, which may limit face-to-face encounters, and suggest inter-PCT links were this decision-influence process to be examined at that granularity.

4.3.2.4.2. 'Information' network ties

The mean number of ego-alter ties reported in the 'information' network is 2.83 (95% CI 2.44 – 3.22, n=161). 16.1% of respondents specified that they had discussed measles-related vaccination with no-one (Figure 4-11). There is little evidence for difference in mean ties, when stratified by respondent demographics – sex, age, ethnicity, education, family size, and PCT location (p>0.05, using Mann-Whitney and Krusal-Wallis as appropriate).

Considering the inferred opinion of the respondents, non-adherents have fewer advisors (mean 1.14, (n=7), than adherents (mean 2.90, n=115), with moderate evidence that this difference is statistically significant (Mann-Whitney, U=228.5, p=0.05).

Figure 4-11 Advisor distribution



Base: all adult respondents censoring for missing answers (n=161)

Given the maximum number of reported contacts exceeds the number of spaces in the paper questionnaire, we check for a format bias (there was no maximum imposed online). Evidence for a format effect is found, but not in the anticipated direction (Figure 4-11b&c) as the mean number of reported ties for paper responses (3.47) is significantly greater than that for web responses (1.09) (Mann-Whitney U=1033.5, p<0.001, n=161). Age and education are the only demographics to have non-homogeneous use of format (χ^2 tests, p<0.05, web-use skewed towards older or better-educated respondents), and format use is not associated with inferred opinion (Fisher Exact, p=1.00). However the data do not permit a satisfactory application of two-way ANOVA analyses adjusting for format and demographics. Nevertheless, all the demographically-stratified means are higher for paper than for the corresponding web-submitted sample, so we conclude there is evidence that the number of advisors reported is subject to a format bias, with fuller disclosure by the respondents using paper questionnaires (mean 3.47, 95% CI 3.00 – 3.94).

Censoring web respondents also retains the moderate evidence for a statistically significant by inferred vaccination-opinion difference with mean numbers of advisors 1.60 and 3.60 (non-adherents and adherents respectively) (Mann-Whitney U=116.5, p=0.09). Non-adherents were also more likely to report no discussions (57% vs 18%, n=121, p=0.03 adjusted for format use).

Less than half of respondents had discussed measles-vaccination issues with a HCP: 46% (n=134) of respondents with categorised advisors (Table 4-7), an estimated 38% of total

when adjusting for respondents with no reported advisors). By contrast, an estimated 71% of all parents had discussed this subject with friends or family; the majority of respondents reported discussions with their partner (estimated 54% of total) with friends and the child's grandparents also each estimated to be consulted by more than a third of parents (estimated 38% and 34% respectively) . Again, there is little evidence (Fisher Exact, significance measured at $p<0.05$) for difference in the adherents and non-adherents accessing different types of people.

Table 4-7 Categories of advisors accessed by respondent

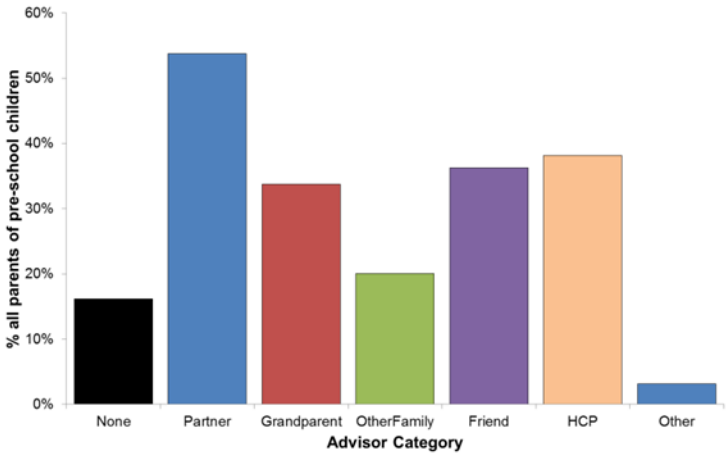
Respondents with at least one advisor in this relationship category

	Total
Partner	64.2%
Child's grandparent	40.3%
Other family	23.9%
Friend	43.3%
HCP	45.5%
Other	3.7%

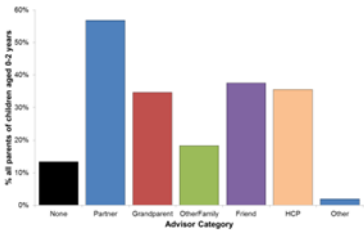
Base: all respondents with at least one advisor, censoring missing relationships (n=134)

Figure 4-12 Estimates of advisor categories accessed

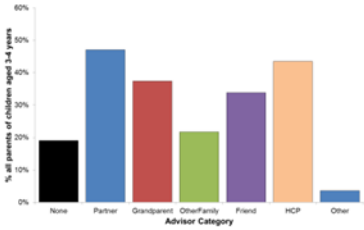
a] all respondents



b] respondents with children aged 0-2years



c] respondents with children aged 3-4 years



Base: all adult respondents censoring for missing answers (n=161, 105, 84)

Having previously noted format effect on the numbers of advisors listed (with web-based respondents under-reporting), we find that the types of advisors listed also differ. Fewer web-based than paper-based respondents included friends (χ^2=12.114, 1 df, p<0.001), the child's grandparents (χ^2=7.353, 1 df, p<0.001) or other family (χ^2=6.740, 1 df, p<0.001) in their listed advisors (n=31 web, n=103 paper), but comparable proportions included partners across both formats (χ^2=0.656, 1 df, p=0.42). Under the assumption that the format used is actually independent of the nature of the parent's advisors this may indicate the advisors "omitted" by web-based respondents were more like to be friends or family members (other than their partners).

4.3.2.5. 'Potential infection' networks

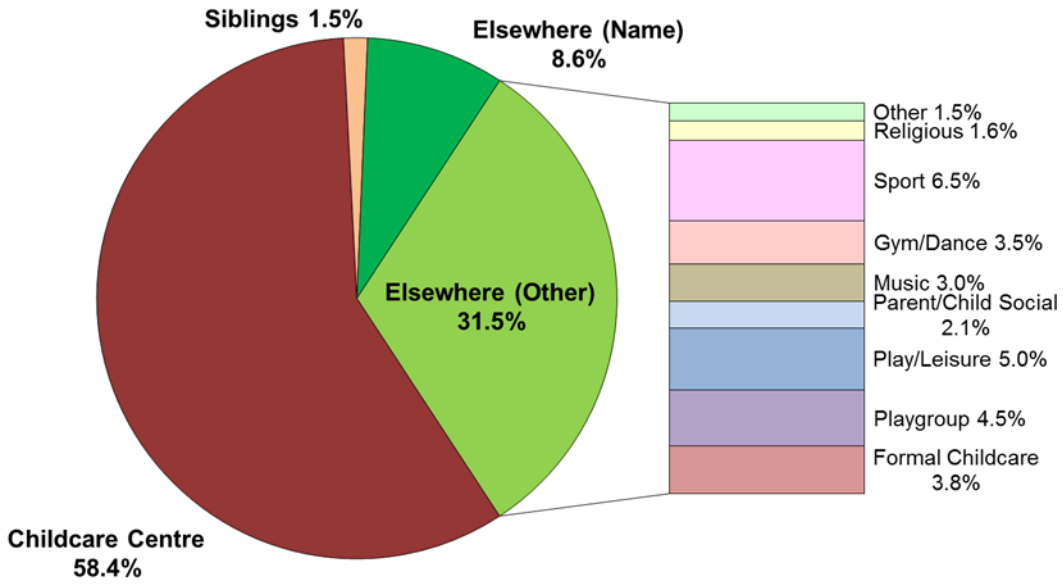
Data imputation was applied to counter a consistent oversight across all responses: i.e. siblings were not named in the contacts lists. Hence we imputed responses of all pre-school siblings as contacts for all survey children (74 ties affecting 73 children).

Ego-centric potential infection network data are available for 195 respondents (92%, n=212), with a total of 4971 alters reported.

4.3.2.5.1. 'Potential infection' network alters

Contacts were uniquely categorised by context in which they were primarily encountered: siblings, at the centre, elsewhere (named and un-named). The majority of contacts are made with those at childcare (58%, n=4971) (Figure 4-13), and this is the largest category for all demographic strata (ordinal and age of child, age, education and ethnicity of parent, PCT location, family size). Groups of contacts were met in both formal (e.g. sports sessions, playgroups, music classes) and informal contexts.

Figure 4-13 Contacts sample by context



Base: contact ties for all sample children (n=4971)

4.3.2.5.2. Potential infection network ties

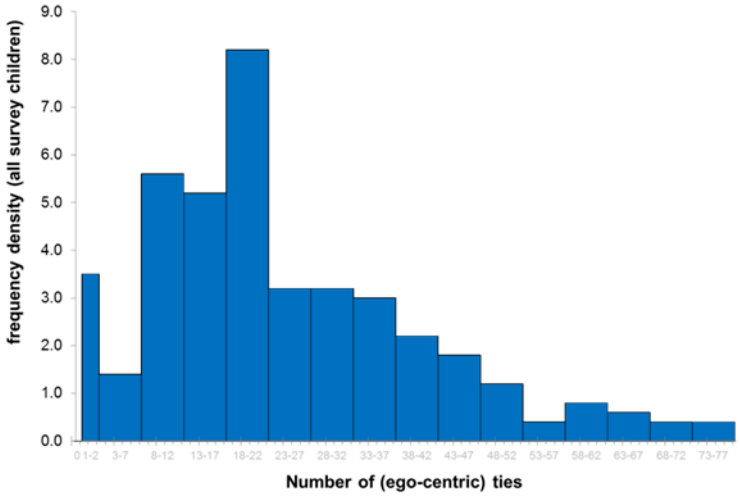
No children had zero contacts – although before the imputation of sibling ties 7 (4%) were declared as meeting no other pre-schoolers during a typical term-time week. The mean number of ego-alter ties in the 'potential information' network is 25.49 (95% CI 23.26 – 27.72). The distribution of the number of ties is given in Figure 4-14. Unlike the listing of adults advisors, we find no strong evidence for difference in the mean entries by format (number of line entries for contacts made at places other than at the centre ANOVA F=2.490, 1 & 210 df, p=0.12).

1163 contacts were reported by name (23%), the mean number of named ties is 5.96 (95% CI 5.27 – 6.66). There is very little evidence that the number of additional (un-named) contacts reported is associated with the size of the network of named contacts only (Spearman's rho=0.019, p=0.80).

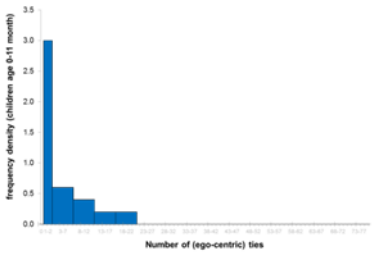
Figure 4-14 Contact distribution

All contacts

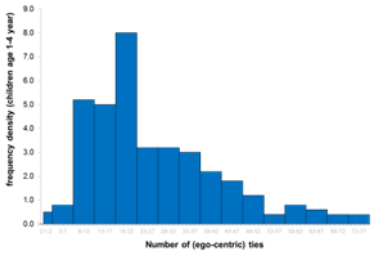
a] all survey children



b] ego age 0-11 months

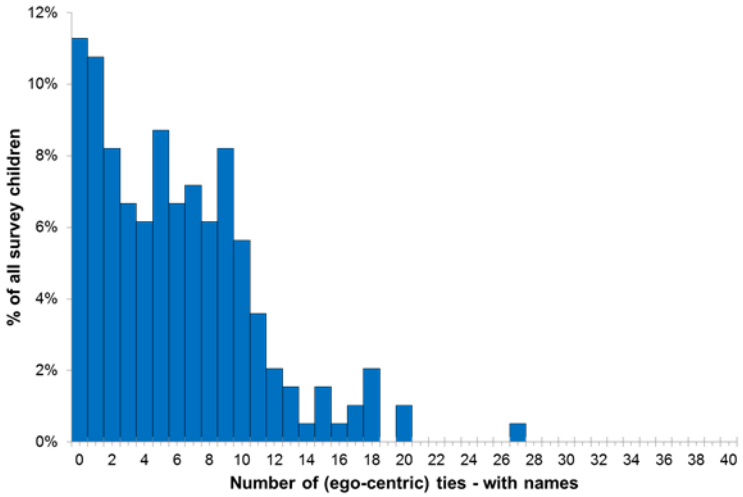


c] ego age 1-4 years

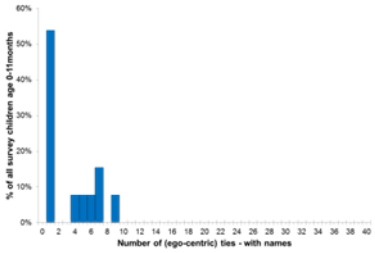


Named contacts only

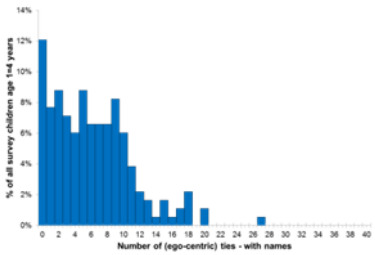
d] all survey children



e] ego age 0-11 months



f] ego age 1-4 years



Base: all sample children censoring for missing answers (n=195, 13, 182)

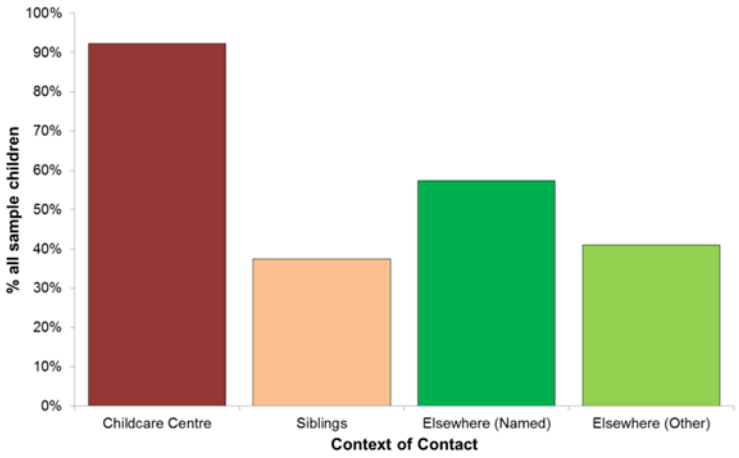
There is strong evidence for a difference in mean number of ties when stratified by the child's age and child's enrolment at the centres. Children attending the centre have a higher mean contacts than non-attendees (27.07 vs 6.60, Mann-Whitney U=2492, p<0.001) as do older children (Krusal-Wallis H=41.155, 4 df, p<0.001, with increasing trend by age

Jonckheere T_{JT}=9874.5, p<0.001). For other measured demographics, there is little evidence for difference in mean number of ties when stratified by PCT location, and parent's age and ethnicity; the evidence for different means when stratified by family size becomes insignificant when the child's age or centre attendance are also taken into consideration (p>0.05, using Mann-Whitney and Krusal-Wallis as appropriate).

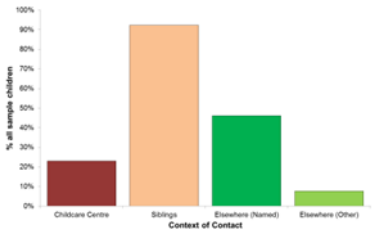
Also, more children meet contacts at the centre than in the other categories (92%, n=195), and this was also the case for all demographic strata examined (as above), except for those under one year old (Figure 4-15). Of the children who do not attend the centre, 47% (n=15) had no contacts other than siblings, whereas 26% of attendees (n=180) also had non-sibling contacts outside the childcare context.

Figure 4-15 Proportion of children with any contacts in specified context

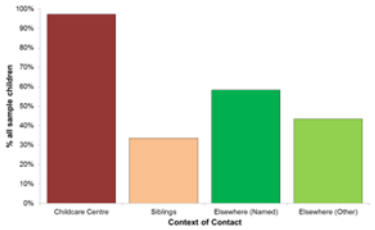
a] all survey children



b] ego age 0-12 months



c] ego age 1-4 years



Base: all sample children censoring for missing answers (n=195, 13, 182)

We recall that the contact definition is based on weekly contacts and pre-school contacts only. We use weekly contact data weighted by context (§4.3.2.5.1) to estimate mean daily contacts. We assume siblings meet every day, childcare contacts 5 days per week (mean attendance frequency at school-based nurseries in England [266]) and all others contacts are met once a week; we obtain a daily mean contacts of 12.48 (of whom 3.11 are named contacts). A more conservative childcare attendance of 3 days per week (mean attendance at full-daycare settings in England is 3-4 days per week [266]) gives daily mean of 8.22, (of whom 2.15 are named contacts).

Table 4-8 Mean weekly contacts by context

Context	Mean	(95% CI)
Centre	14.89	(13.81 – 15.97)
Elsewhere	10.22	(8.33 – 12.11)
Siblings	0.38	(0.31 – 0.45)

Base: all sample children censoring for missing answers (n=195)

4.3.2.6. Network overlap

104 advisors are reported to have offspring who are named contacts in a linked 'potential infection' network (28%, n=366, censoring partner advisors), and 140 'potential infection' ties overlap. (The named contacts restriction may create under-reporting bias, for parents of un-named contacts.) Assuming partners are parents to the sibling contacts (data not reported as a corollary of reporting oversight 4.3.2.5), a total of 191 advisors are parents of named contacts (42%, n=453) and 248 'potential infection' ties overlap.

More generally, respondents reported that 65% of their advisors were themselves parents of under 5s (n=306, censoring for advisors with whom respondents have a primarily professional relationship - as respondents were unable to provide an informative parent-status for the majority (58%) of these advisors). We note that 20% of all non-professionally-known advisors are grandparents of the respondent's child and so are unlikely to be parents of a pre-schooler, providing a ceiling for the possible overlap.

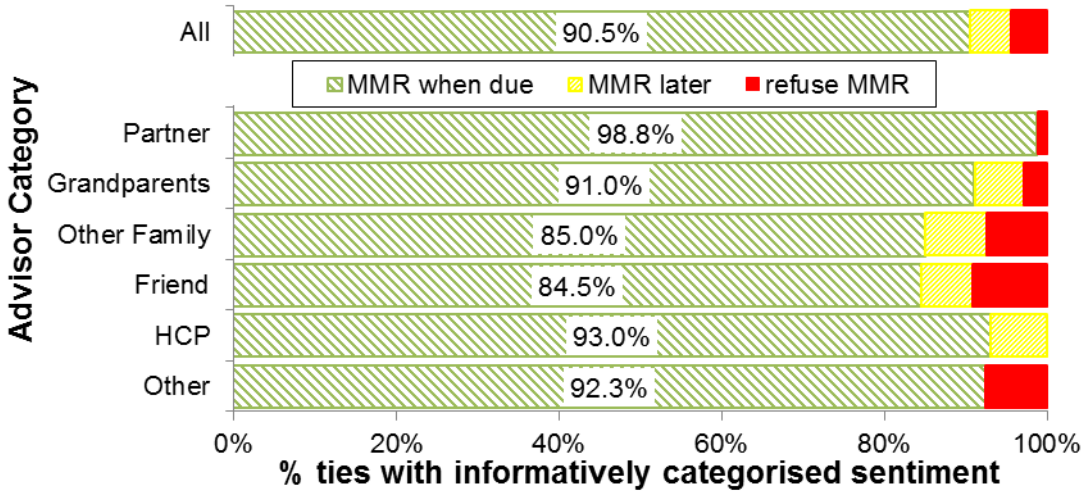
4.3.2.7. Vaccination information received

4.3.2.7.1. Inter-personal communication content

The sentiment of the communication shared across the information network ties was informatively-categorised for 93% of the ties (n=456, 6% had unclear sentiment or poorly recalled). The majority (90%, n=421) carried sentiment that supported the adherence to the recommended schedule ("should get the MMR jab done when it is due"). The 10% that opposed adherence were evenly split between delaying and refusing MMR. Support for receiving MMR as scheduled forms the majority of the communication shared across ties with every relationship category of advisors (Figure 4-16), but there is evidence the proportion of sentiment that counsels against the schedule varies by this relationship

(χ^2=14.07, 5 df, p=0.02) with the highest proportions of non-adherence advice carried by tie with friends and other family (not partner not child's grandparents).

Figure 4-16 Sentiment communicated



Base: 'information' network ties, censoring missing answers and unclear/poorly recalled content (n=418-421)

The relationship between the respondent's inferred vaccination opinion and the sentiment(s) of the information shared with their alters is explored further in Chapter 5.

4.3.2.7.2. Direct exposure

10 respondents (6%, n=166) claimed to personally know recent measles cases and 4 respondents (2%, n=166) claimed to personally know someone who experienced a "serious adverse reaction" attributed to the MMR immunisation. (To put these numbers in context, see Box 4-2.)

The events recalled are subject to perceptions of severity and attribution by the respondent; they do not necessarily correspond to events which were included (or excluded) within the surveillance systems for measles cases (NOIDS), or adverse drug reactions (MHRA Yellow Card scheme) - surveillance details are given in Chapter 1.

Box 4-2 Indications of scale

Enrolled children at recruitment centres	0.07%	(of under 5's)
10 unique infection events is equivalent to		
notified measles cases (all)	0.06%	(in previous 5 years)
notified measles cases (in under 5's)	0.09%	(in previous 5 years)
confirmed measles cases (all)	0.18%	(in previous 5 years)
confirmed measles cases (in under 5's)	0.58%	(in previous 5 years)
		Base: England

Enrolled children at recruitment centres	0.06%	(of under 5's)
4 unique vaccination events is equivalent to		
reported adverse events for the MMR vaccine	0.62%	(in previous 5 years)
		Base: UK

under 5 population [191], *measles cases* [48-51], *adverse events* [45].

They do reflect (perceived) evidence which respondents could chose to draw upon when judging the risk of infection or of adverse vaccine reactions, which are variables in several models applied to vaccination decisions (Chapter 1). However, as all these respondents' children were vaccinated, there is no evidence (Fisher Exact, p=1.00) of a relationship between knowing a case and seeking vaccination, nor knowing an adverse event and avoiding.

Given this somewhat unexpected total lack of association, we investigate if there is differential advice experience between parents making their vaccination decision with or without this direct knowledge.

Respondents declaring knowledge of "measles cases"

There is evidence for difference in the mean number of advisors consulted: 5.80 for those with this knowledge vs 2.64 for those without (Mann-Whitney U=1177, p<0.001, and Mann-Whitney U=788, p=0.01 for paper responses only - given lack of web responses for the exposed group and the suspected format effect on reported numbers of advisors). If this association were to be symptomatic of causality the direction is unclear: does having more advisors increases the likelihood of case discovery, or does knowing of a case drive increased advice-seeking?

Respondents declaring knowledge of "adverse reactions to MMR"

There is no little evidence for a difference in the mean number of advisors (Mann-Whitney U=290.5, p=0.83) nor of different proportions of advice in favour of schedule adherence (Mann-Whitney U=199.5, exact p=0.64) vs respondents with no such knowledge. Given that vaccination is a non-reversible action, we note that temporal order of the child's vaccination and becoming aware of an adverse event is unknown.

4.3.2.8. Vaccination status

Censoring children under 14 months (MMR1 routinely administered at 12-13 months [5]), MMR cover is 98% (n=181, Table 4-9) and 98% of parents have vaccinated at least one child (n=157, Table 4-10). Stratifying the child MMR status data by location, child's ordinal, and the age, ethnicity and educational status of the respondent reveals no significantly different levels of uptake by demography (Fisher Exact, significance measured at $p<0.05$).

Censoring children under 24 months (the age used in MMR1 COVER statistics [10]), the recalled uptake is 98% (n=136). There is very strong evidence that the figure is different ($z=3.84$, $p<0.001$) to the corresponding COVER uptake for MMR1 of 86.6% (annual COVER for April 2010-March 2013, weighted by sample size by year by PCT) [37-39]. We note that the survey measurement is not directly comparable with COVER data, with a bias towards higher coverage, as there may be children who were unvaccinated at 24 months (their status for COVER) but who have been subsequently vaccinated.

Table 4-9 Children's MMR status

		% Vaccinated	
Current Status	Age 14 months and over	97.8%	n=181
	Age 24 months and over	97.8%	n=136
Pre-catch-up Status	Age 14 months and over	92.1%	n=152
	Age 24 months and over	94.3%	n=105

Base: all sample children with informative MMR status, censored by age (at time indicated)

Table 4-10 Parent's MMR participation

	At least one child Vaccinated	At least one child Unvaccinated	
Current Status			
Children 14 months and over	98.1%	2.5%	n=157
Children 24 months and over	97.6%	2.4%	n=124
Pre-catch-up Status			
Children 14 months and over	92.7%	8.0%	n=137
Children 24 months and over	94.9%	6.1%	n=99

Base: all adults respondents with at least one uncensored child (children with informative MMR status and age at time indicated)

As noted in §4.3.1.1, a high-profile measles outbreak and national MMR catch-up campaign were concurrent with the fieldwork, which we thought could alter vaccination decision-making and behaviour, so a retrospective "pre-catch-up" vaccination status was added to measured variables.

Cross-sectional and longitudinal data all point to a general increase in pro-MMR behaviour since the "pre-catch-up" period. All four cross-sectional measurements of MMR uptake are lower pre-catch-up (Table 4-9 and Table 4-10) and current coverage is higher for all previously unsaturated demographic strata except black ethnicity (demographic status assumed invariant, so age strata excluded) (stratified data in Appendix). The higher current coverage has contributions from new vaccinations in the pre-catch-up cohort (longitudinal data shows 92% cover rising to 98% in those then aged 14months - McNemar, one-way, $p<0.001$) and higher uptake in the cohort that has since turned 14 months (97% cover, n=29). Comparison of the cohorts of children within the MMR1 schedule window is consistent with an increased urgency, albeit with very small samples (current vs December 2012 cohort; 12 month olds: 33% vs 25% vaccinated, n=7; 13 month olds: 100% vs 60% vaccinated, n=9).

We therefore conclude that the original vaccination decisions (uninfluenced by the atypical fieldwork context) are better represented by pre-catch-up data. Strong evidence for sample bias (against vaccine hesitancy) is still present, as the pre-catch-up uptake in children then 24 months or older is significantly different from the weighted COVER MMR1 uptake ($z=2.64$, $p<0.01$).

The definition previous adopted for inferred vaccination opinion (§4.3.1.7.3) incorporates vaccine acceptance evaluated using pre-catch-up data.

4.3.2.9. Intradyad agreement

We examine the distribution of the vaccination opinion and vaccination status across the networks via intra-dyad agreement (IDA), utilising the opinions and status inferred for this purpose (§4.3.1.7.3)

4.3.2.9.1. Intradyad agreement – vaccination opinion

It is possible to measure the proportion of intra-dyad agreement in 313 ties (69% of ties, 94 respondents are included). There is 89% IDA across these ties (Table 4-11.). There is strong evidence that this differs from that expected if the ties were allocated randomly (Fisher exact, p=0.04). This result (direction and strength of evidence) remains robust under alternative definitions to infer the respondent's opinion (child's current vaccination status censoring under 14month olds, child's pre-catch-up status censoring then under 24 month olds), which offer differing potential for recall bias or window for timely vaccination.

Given availability of only ego-centric data, we have also considered an alternative randomisation: fixing the ego-opinion marginal totals but randomly generating ties with no restriction on the alter-opinion marginal totals (using Bernouilli trials each with the probability forming a tie with a vaccine-supporting alter set at the proportion of all advisors who are categorised as supporting scheduled MMR adherence). In this scenario the evidence that the observed IDA differs from that of randomly-generated ties is less strong (exact, p=0.77). (Poisson Binomial calculated using poibin package (version 1.1, 2012) [269]) in R (version 2.5.2, 2012) [270].

We find no evidence for non-random IDA across professional relationship ties (HCP: Fisher exact =1.00, n=72), in contrast to the stronger evidence for socially-focused relationships (Fisher exact, Friends & Family p=0.07, n=269; Partner p=0.03 n=59). The non-random IDA with partners suggests a degree of homophily in child-rearing decisions.

Table 4-11 Ego-alter ties within 'information' network

a] ties with all types of advisors

Inferred Opinions of connected nodes

		Alter opinion	
		Support	Oppose
Ego opinion	Adhere	275	30
	Non-Adhere	5	3

Intra-dyad agreement = 88.8%

Base: all 'information' network ties with inferred opinion for both nodes (n=313)

Comparison with tie randomisation

Expected intra-dyad agreement = 87.4%

Fisher exact, p=0.04

Comparison with alternative random generation of ties

Expected intra-dyad agreement = 88.4%

Exact, using Poisson Binomial distribution, p=0.77

b] excluding ties between partners

Inferred Opinions of connected nodes

		Alter opinion	
		Support	Oppose
Ego opinion	Adhere	215	30
	Non-Adhere	4	2

Intra-dyad agreement = 86.5%

Base: all 'information' network ties with inferred opinion for both nodes, censoring for ties between partners (n=251)

Comparison with tie randomisation

Expected intra-dyad agreement = 85.5%

Fisher exact, p=0.17

Comparison with alternative random generation of ties

Expected intra-dyad agreement = 86.5%

Exact, using Poisson Binomial distribution, p=0.92

4.3.2.9.2. Intradyad agreement – vaccination status

There are 209 ties for which we have the vaccine status for both nodes (4% of ties, 110 sample children included as ego node) hence it is difficult to make robust observations on the level of IDA within the child population.

There is 77% IDA across potential infection network ties (Table 4-12), and there is little evidence that this differs from that expected if the ties were allocated randomly (Fisher exact, p=0.23), similar results (exact, p=0.35) are obtained under a similar random tie-generation process as used for the 'Information' network IDA analysis.

We might expect the sibling-sibling ties (n=40) to be non-randomly paired in terms of vaccine status, inflating the above proportions of IDA. However the IDA in the sample with these ties censored is 79% and there is insufficient statistical power to examine this situation further for non-random values (using $p<0.05$).

Table 4-12 Ego-alter ties within 'potential infection' network

a] ties with all types of contacts

Inferred Status of connected nodes

		Alter status	
		Vaccinated	Unvaccinated
Ego status	Vaccinated	160	35
	Unvaccinated	12	0

Intra-dyad agreement = 77.3%

Base: all 'potential infection' network ties with inferred status for both nodes (n=207)

Comparison with tie randomisation

Expected intra-dyad agreement = 79.3%

Fisher exact, p=0.23

Comparison with alternative random generation of ties

Expected intra-dyad agreement = 79.7%

Exact, using Poisson Binomial distribution, p=0.35

b] excluding ties between siblings

Inferred Status of connected nodes

		Alter status	
		Vaccinated	Unvaccinated
Ego status	Vaccinated	100	25
	Unvaccinated	2	0

Intra-dyad agreement = 78.7%

Base: all 'potential infection' network ties with inferred status for both nodes, censoring for ties between siblings (n=127)

4.3.3. Discussion

This survey addresses multiple gaps in the quantitative literature regarding parents' vaccination decisions in the UK. This new evidence includes measuring the clustering of vaccination opinions, quantifying parents exposure to reports of adverse vaccine reactions, and both quantifying the numbers of contacts with whom parents discuss the MMR vaccination decision and the sentiment of their advice. We are unaware of any previous studies reporting the numbers of such advisors, within a jurisdiction with voluntary vaccination - after this survey went to field a study from the USA [271] was published with data on the number of vaccine-advice contacts (purposive non-adherent oversampling, n=196) however the fieldwork location has compulsory MMR vaccination [272] which may limits its generalizability to decisions in the UK. Furthermore this information is linked to data on pre-school children's social contacts – the latter collected specifically for this age-group which addresses inherent weaknesses with respect to the data on this age-group from all-age contact studies (i.e. measurement bias, sample size).

We find that the majority of vaccination-information discussants are not HCPs. Hence health promotion campaigns, which have the objective of improving the quality of vaccination-advice given to parents, will necessarily only be able to affect a minority proportion of advisors if the campaign is focussed solely on the HCP community. Other sources of information (people as identified here, and also recognising the use of online/printed media [21]) should be included in such campaigns to maximise the advice-sources that can be thus affected.

This survey found that adherents to the recommended vaccination schedule had a significantly higher number of advisors than non-adherents (3.6 advisors vs 1.6). Possible interpretations include that a higher number of advisors strengthens a normative effect or that non-adherents are more reluctant to seek advice, but inferences on the direction of causality (if any) of the relative egocentric network size is beyond the scope of this study. Comparison with the data from the USA study [271] – in which non-conformers have 6.7 advisors vs 4.8 for non-conformers – is uninformative on this point as in a mandatory-vaccination context advice may be sought not just on the decision itself but also on logistics of implementing a non-conformity decision. However, we do note a potential artefact in that social desirability bias may have led to non-normative opinion-holders reporting fewer contacts given the measles/MMR context which developed during the fieldwork (discussed below); a second wave at a less contentious time would prove useful.

As noted above, to our knowledge, this is the first UK-based quantitative study on the numbers of vaccine-decision advisors and age-specific measurement of pre-schooler's social contacts. Nonetheless we are able to compare results with previously identified USA studies for the former element and for UK all-age studies for the latter. The mean number of advisors is similar to the "important matters" measures from the USA General Social Survey [231, 241] and lower than the Brunson study [271]. The latter difference could plausibly be explained by the different geographies (regarding mandatory vaccination, as noted above); however the use of different data collection instruments means we cannot discount a methodological influence (specifically the prompts employed to encourage name generation by recognition, to improve disclosure vs unaided recall [273]) are used differently).

The preschool sample is larger than all-age contact surveys found in the epidemiological literature (e.g. twice POLYMOD's [234]), and this survey was able to include a more preschool-appropriate contact definition that all-age surveys (have a "two-way conversation" – contact definition in Mossong et al [234] - is difficult for parents to interpret for the younger age-groups).

For the associative social contacts of children aged 0-4 years, a direct comparison with POLYMOD [234] is not possible, however the most conservative estimate of mean daily contacts is higher in this study (8.2 vs 1.9 contacts) and remains so if we restrict consideration to named contacts only in this study (2.2 contacts) or to physical contacts only in POLYMOD (1.5 contacts). Although mindful of the rise in childcare attendance between the fieldwork periods [263], addressing measurement bias through question-wording appropriate for younger children (see below) and a more robust sample size (e.g. twice POLYMOD [234]), mean these data do suggest a note of caution when using all-age contact studies where preschool children are of specific interest. Hence adaptations of this element of the survey and including non assortative mixing may prove a useful exercise to improve data available for such uses.

There was no evidence of a relationship between knowing a measles case and seeking vaccination, nor knowing an adverse event and avoiding vaccination. Additionally the numbers reported such knowledge was an order of magnitude higher than the prevalence of these cases/events indicated by surveillance data. ADR surveillance data methodology requires careful interpretation and precludes specific statistical analysis, and the expected spread of information across network will vary by network path structure and source distribution. However, with these caveats, this disparity for both types of knowledge raises the possibility that unsubstantiated information is transmitted along with verifiable information

across social networks, and is therefore available as evidence to influence parents' vaccination decisions.

It is of direct interest to this project's hypothesis that there is evidence for the clustering of opinion, and to our knowledge this has not been measured previously. However data cannot inform us if vaccination opinion is the only homophilic measure across the tie (homophily is generally observed in social networks [241]) nor the temporal order of tie- and opinion-formation. We also discuss below difficulties regarding opinion inference from the data that can be collected.

Data collection tool performance

The recruitment methodology and data collection formats are key strengths of the study, specifically recruitment via childcare settings and use of a self-completion questionnaire.

Given the levels of vaccine coverage in the UK and mindful of the perceived judgements of other parents regarding one's decision [87, 97], social desirability bias was expected to be a concern. It is known that social desirability can depress participation or engagement with surveys (skipping questions, withholding information, or the providing inaccurate but 'desirable' answers [274]). These highlighted design factors were incorporated to reduce exposure to social desirability bias: self-completion reduces the bias as compared with interviewer led surveying [275], and placement via GP practices was specifically rejected as parents could think GPs would be biased regarding expected vaccination behaviour (and due to the evidence associating perception of GP's motivations with reduced uptake [83, 89, 91-94, 97, 100, 101]). The proportion of non-vaccinators in the pilot responses demonstrated that this survey's methodology and implementation could deliver unbiased samples.

Moreover, the value of these decisions increased with the unforeseeable events that affected the context of the survey during the fieldwork period, namely a high profile measles outbreak (albeit not in the surveyed area) and the reactive NHS MMR catch-up campaign (see Chapter 1) which are thought to have amplified social desirability bias. Paulhus' model of socially desirable responding [276] distinguishes between egoistical and moralistic bias, the concurrent disease and vaccination context could be interpreted as driving an element of both. Avoiding recruitment via GPs also avoided the potential for the survey invitation to be specifically associated with GP's attempts to contact parents in order to persuade them to 'catch-up' with missed MMR vaccinations.

The survey outperformed the response rate of the most nearly- comparable all-age contact survey BSCS [235], met the pre-fieldwork target for statistical power, and generated a larger database of social contact data for pre-school children in the UK than recent all-age surveys [234, 235]. However the response rate was lower than for earlier surveys relating to MMR [80, 82-84, 89, 93]. This may be from a combination of design factors (non-anonymous data on self and alters, lack of incentives, inclusion of web-responses) and external factors (due to the timing of the fieldwork). Given our objective to collect data to enable the reconstruction of networks the collecting anonymous data was an unavailable option. The decision to not offer an incentive for response *per se* was a pragmatic one, given the combination of logistical and ethical considerations (centres could not handle incentive logistics and maintain response confidentiality of response, respondent contact details would be for centrally-administrated post-fulfilment). Thirdly, there is some evidence web-based surveys depresses response vs paper [277], and our analysis did lead us to conclude there was a format effect on this survey (with reduced depth of response from web-based participants). Evidence from an unusually strong social desirability context, the historical timings of previous surveys (closer to the peak of the vaccine scare which may have been motivational) and the higher response rate from the pilot (overall and compared with similar settings included in the main sample) would support an absolute temporal effect on response rate. Additionally, anecdotal evidence exists for survey fatigue in the Wandsworth sample (several surveys were in field from a variety of sources).

The failure of the snowball may also be similarly affected by outside events, and given the variable of interest was size of the contact network of offering an incentive for snowballing would have severely compromised data integrity. Combined with the skew in within-centre responses, the ability to measure transitivity data for the networks was severely compromised, and no results are reported.

Given this survey is intended to contribute to a model of decision-making the issue of heuristics and cognitive biases in the respondents is important [122, 134], for quantitative as well as the qualitative data. The recall bias and availability heuristics [122] inherent in the retrospective design is a survey strength, as the responses more accurately represent the perceived, subjective evidence used in parents' decision-making, which would not be reflected in more objective measurements. Other cognitive biases remain a weakness, for example, choice-supportive bias [278] would lead to underreporting of advisors and advice contrary to the final (or current) opinion and so amplify the evidence supporting one's final decision. Further social desirability effects will be present too (skewing the bias effect by vaccination opinion).

We infer values for the vaccination opinion and behaviour (for ego and alters) from responses which are subject to recall bias and other specific cognitive concerns, namely misattribution and consistency bias (which may lead to public advice inconsistent with private opinion through trying present oneself in a consistent manner). Additionally, recalled information and behaviour may not reflect current thinking, more so for those now less disposed to vaccinate, as vaccination - unlike opinion – is irreversible. It is difficult to avoid this bias with cross-sectional studies such as this one (although an unplanned retrospective longitudinal element was added during fieldwork)

This survey includes a more preschool-appropriate contact definition that all-age surveys to e.g. have a “two-way conversation” – contact definition in Mossong et al (p382 [234]) - is difficult for parents to interpret for the younger age-groups. This reduces bias in child-child contacts measurement, but it remains an indirect and recalled measure. We have not been able to conceive a study design where this is not inherent, as the children are too young to respond themselves and the open sample (and ethical sensitivities for non-sample minors) limits use observational methods of network data collection.

It was intended to collect data in a “non-outbreak” context; although steps were taken to avoid respondents in proximity to an outbreak, respondents are likely to have been aware of the measles outbreak in Wales. Whilst not affecting vaccination decisions completed before then, it provided a different framing of our questions (external to the questionnaire itself), hence a limit may be placed on generalizability of some results to a truly “non-outbreak” context.

Future work

There are also opportunities to design extensions to facilitate incorporation with data from other age-groups, or to complement the demographic characterisations and vaccination opinion (by ego or homophily across information networks) with either relevant psychological characteristics or media exposure (as source of information informing the decision). Relevant psychological characteristics include altruism, categorisation from models such as theory of culture [130], or personality traits associated with decision-making, susceptibility to conspiracy theories [279] etc. Some individual responses also point to a rich vein of within family-unit dynamics – notably the making of joint decisions (using, say, paired questionnaire with both parents) – and within-family ordinal vaccination patterns.

In the case that further data collection is desired, survey methodology can be adjusted to directly address some of the survey weaknesses (inherent or unplanned) identified above is possible. The following are all relatively simple adaptations: a different (less abnormal)

context, reduced use of web (or web redesign vs paper questionnaires), purposive sampling of the vaccine hesitant, interviewer-led questioning for stronger name-generator prompts; though they may compromise other strengths (e.g. interviewers increase social desirability bias vs self-completion questionnaires [275]).

To address the lack of transitive data may require a semi-closed sample or indirect measurement (like in BSCS [235]). The unplanned introduction of limited longitudinal data could be expanded to gain clarity on some temporal uncertainties through a prospective study, in which case expectant mums are an interesting initial sample [255].

Specific investigation of the relationship between vaccination behaviour and direct knowledge (of infections or adverse reactions) would be valuable given its tacit inclusion as assumption in some decision model frameworks (Chapter 1) and the lack of association found in this dataset.

Beyond data collection, the survey was initiated in order to address the paucity of data to inform a mathematical model, so this is the most immediate future work arising.

5. Revisiting the MMR1 decision model

5.1. Introduction

From previous modelling (Chapter 3) we concluded that information-sharing which influences vaccination decisions is capable of both changing the overall proportion of vaccination-supporters within a population, and also producing opinion-clustering within the population. These effects were found across several different assumptions of network structure and of mathematical representation of the decision process. However, the change in the total vaccine-support level is qualitatively-dependent on the decision-representation and quantitatively-dependent on the network-structure. Furthermore, the opinion-clustering effect is qualitatively- and quantitatively-dependent on the decision-representation and network-structure and their combination.

The collection of empirical data (Chapter 4), to address the paucity of sources available to inform the selection of assumptions used in the model, enables the revisiting of the mathematical modelling. We use the survey data to determine empirically-informed parameters for network-structure and decision-representation. We consider a mathematical model with assumptions that incorporate these parameters (which may or may not be in the parameter space previously considered) to explore the patterns of vaccination-opinion generated via information-sharing. We continue with the specific example of the routine schedule MMR1 vaccination in the UK.

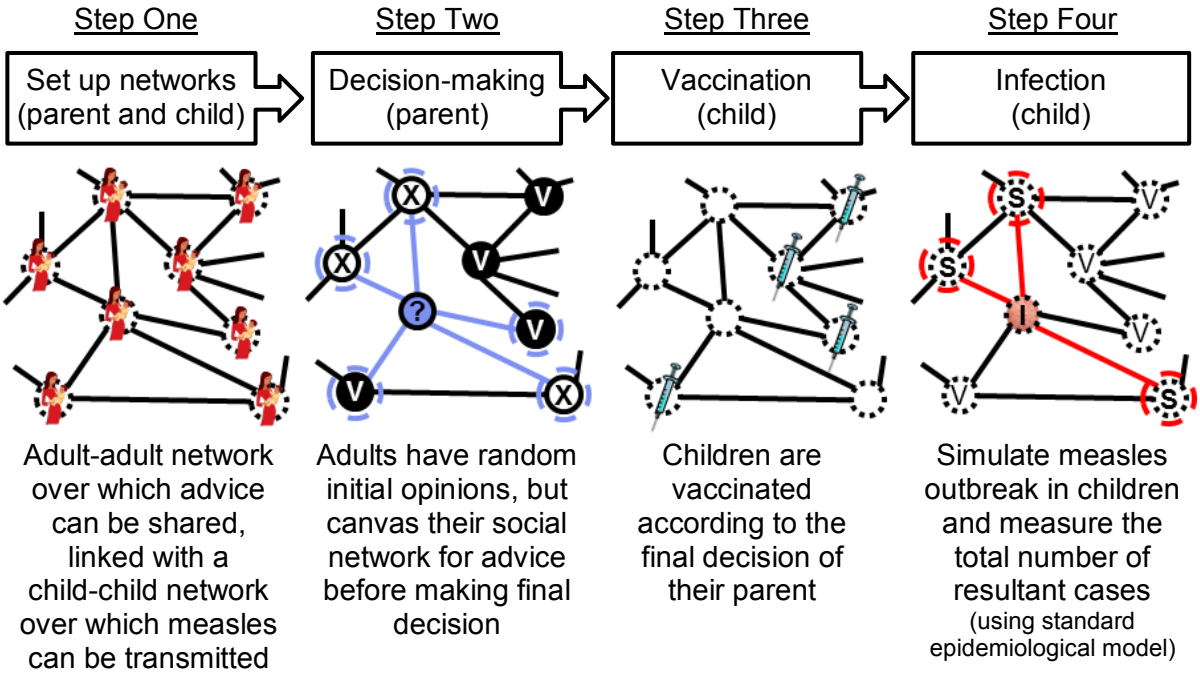
The empirical data also permits parameterisation of an extension to the model framework: to consider the pattern of vaccination uptake within the pre-school population, and resultant outbreak probability when challenged by an infection introduced from a vaccine-refusing community or another geography (the origin of generalised measles outbreaks in Sussex [280], France [109], Netherlands [281] and in Wales [53] respectively). Both the proportion of unvaccinated individuals and their distribution within the population affect the potential for outbreaks of vaccine-preventable diseases [47, 112]. Clusters of unvaccinated individuals enable outbreaks to establish in otherwise highly-vaccinated populations.

Similar to the approach used in the investigation of opinion clustering in the adult population (regarding MMR1), the pattern of (MMR1) vaccine uptake is considered using a network of social contacts (which offer measles transmission opportunities). The potential for the pattern of vaccine opinions to affect outbreak potential within the child population (via the resultant pattern of vaccination uptake) will be moderated by the overlap of edges in both

networks ('information' and 'potential infection'), with greater outbreak probability associated with larger overlaps [168]. Previous work [168] has assumed the adult and child modelled networks have the same mean vertex degree (MVD) and use similar construction algorithms. Previous empirical quantitative studies (albeit with imperfect matches to the assumptions for MMR1-measles in the UK) have focussed on either the information contacts [271] or the child transmission network [234, 235]; no single-source studies have been identified. This survey (Chapter 4) provides single-source empirical data on both networks, including some information on overlap). Additionally, the pre-school physical proximity contact data is both appropriate to measles-transmission opportunities and has a larger sample of children aged 0-4 years, than in previous all-age social contact studies in the UK [234, 235] .

We therefore extend the model, and its parameterisation, to include the information-sharing process (on the 'information' network), the translation of resultant vaccination decisions into the vaccination status of the offspring (age 0-4 years) of parents within that network, and the infection dynamics within that child population following the introduction of index case infection. We seek to use this combined decision-infection model to understand the impact of a peer-to-peer vaccine-information-sharing process on the pattern of opinion, the pattern of cover and outbreak probability.

Figure 5-1 Extended model overview



5.2. Parameter fitting

5.2.1. Methods

5.2.1.1. Network structure parameters

We determine parameters for the network structures of the information network (for the decision model) and of the potential-information network (for the infection model). In the previous modelling (Chapter 3) the structure of the decision model network was determined by the combination of network-build algorithm and mean vertex degree (MVD). Using the ego-centric survey data, we obtain values for the mean vertex degree (from those of the (ego) nodes), and another network structure characteristic (the degree distribution) which we use to determine an appropriate network structure for the model.

The degree distribution is the frequency distribution of the numbers of immediate network neighbours, across all vertices. We compare the degree distributions from the survey data with those obtained using the network-generation algorithms considered in the previous modelling and – given visual inspection of the plots – with some standard long-tail probability distributions: exponential (parameter λ), negative binomial (parameters w, p) and lognormal (parameters μ, σ). The algorithms considered are Erdős-Rényi for random networks [228], Watts-Strogatz for small-world networks [229]), and Barabási-Albert for scale-free networks [230]. Erdős-Rényi generates a degree distribution which the follows Poisson distribution (parameter λ) (an approximation for the Binomial distribution for large networks). The degree distribution of a network generated using the Watts-Strogatz algorithm (parameters m, β) is given by Equation 5-1 [282]. The term within the summation is equivalent to the product of p(n) under Bin(m, 1-β) and p(x-m-n) under Poi($m\beta$).

Equation 5-1 Watts-Strogatz algorithm degree distribution

$$\boldsymbol{P(x) = \sum_{n=0}^{min(x-m,m)} \frac{m!}{(m-n)!\,n!}(1-\beta)^n \beta^{m-n} \frac{e^{-m\beta}(m\beta)^{x-m-n}}{(x-m-n)!} \quad for\ x \geq m}$$

mean vertex degree is $2m$, rewiring parameter $\beta \in (0,1)$).

The degree distribution of a network generated by Barabási-Albert tends towards a power law distribution, $P(x) \sim x^{-3}$. We therefore compare with discrete power law probability distribution (parameters γ, x_{min}) and also consider the degree distribution for the non-limit situation given in Equation 5-2 [283].

Equation 5-2 Barabási Albert algorithm degree distribution (non-limit)

$$P(x) = \frac{2m(m+1)}{x(x+1)(x+2)} \quad for\ x \geq m$$

m is number of outgoing edges from each additional vertex in the algorithm).

We use Maximum Likelihood Estimation to optimise parameters (R package bbmle [284]), and, for the power law distribution, R package poweRlaw [285] based on the work by Clauset [286], comparing the fit between the candidate formulations using Akaike information criterion (AIC). We note that the not all the candidate distributions are defined at zero and, similarly, the small world and scale-free algorithms generate a single giant component; therefore we initially censor the elements with zero contacts when performing the fitting.

5.2.1.2. Decision model: Information network data

Both the model and the survey data focus on the parent's decision whether to present their child for the scheduled MMR1 vaccination, with egocentric data collected for one respondent (parent) per child. For simplicity, the model has a single (parent) vertex in the 'information network' with an opinion status which is transferred to inform the vaccination status of each single (child) vertex in the 'potential infection' network.

However, the vaccination decision for each child may be a joint decision made by more than one of their parents, if they are co-parenting during the decision-making period (active consideration of vaccinations starts before birth [256] and scheduled MMR1 vaccination would occur within a couple of months after the child's first birthday [5]). Hence a 'parent' vertex in the model may represent the combination of two individuals and the single-respondent survey data may not represent the full set of 'information network' neighbours for a joint decision.

Data on lone parent or couple status were not collected in the survey, but there is evidence (from two national datasets: registered births and census population [213, 287]) that raises the possibility that the sample is biased on this measure. 75% of families with dependent children aged 0-4 years old (in surveyed regions, weighted by uncensored response by PCT) are not lone-parent families [287]. This is significantly different (z=2.796, p=0.01) from the proportion of surveyed parents who include a partner as an advisor, 53% (n=160) (i.e. 64%, n=134, of those who cited any advisors). Both these proportions are lower than the proportion of children who are born to parents who were legally partners or assumed to be

cohabiting (85%), based on national birth registration data [213]). Both population measurements are consistent with either a biased survey sample bias (towards lone parents) or with 15%-25% of couple-parents not citing their partners as an advisor (either a similar phenomenon to the observed non-reporting of siblings (Chapter 4) or their partner was genuinely not consulted). That said, the proportion consulting with partners is higher than reported in the DH/COI CITS data (48% of parents of 0-2s) [21].

We therefore categorise the decision-model vertices into two types: those corresponding to "couples" and "lone parents" (at a proportion estimated from the population datasets) and we estimate degree distributions for each vertex type separately. Given the lack of data to inform our assumptions, for the couples we take two example estimates of the (net) egocentric network used to inform the decision: "solo" where the shared network is one parent's network only (their partner adds no unique contacts), and "joint" where the shared network includes unique contacts from both parents.

We inform the parameterisation of the degree for each vertex type using the data stratified by the inclusion or absence of the partner in the reported ego-centric network: "lone parents" uses data from respondents who do not include partners as advisors, and "couples" uses data from respondents who do include partners as advisors; assuming that both types of decision-making units are equally likely to have no advisors. For "solo" egocentric network assumption for a "couples" vertex we use that raw data. Under the alternative assumption of "joint" egocentric network for a "couples" vertex, weights are applied to the "solo" data (using the reported ego-alter relationship data) to calculate a plausible size for a combined set of contacts as follows. We assume both parents have the same pattern of advisors (i.e. mirrored) and de-duplicate shared contacts (contacts which are neither friends nor family are taken to be shared contacts: 94% of those assumed duplicates are healthcare professionals or childcare staff), and partners themselves are removed as no longer external to the "joint" decision. We test parameter and model sensitivity to these assumptions through the degree distributions obtained, and qualitative model outcome (in §5.3).

As noted previously (Chapter 4) there is strong evidence for a survey format effect for the number of advisors reported, with the web format depressing full disclosure. Additionally, the paper sample is closer to the population profiles of parent age and education. We therefore consider data which censors web responses to remove this artefact and reduce response bias.

From these analyses we determine a suitable range of MVD to explore, and compare the degree distributions for each to inform the structure for the final 'information network' model

5.2.1.3. Infection model: Potential infection network data

The previous analysis (Chapter 4) revealed the importance of the contacts made with roommates at the childcare setting within the egocentric networks, with such contacts forming 58% of all contacts. Also the measles transmission opportunity definition [243] determines that the 'potential transmission' subgraph for n roommates is the complete graph K_n. We wish to retain this structure within the model, and so regard the overall degree distribution as the sum of the distribution of setting-roommate contacts and the distribution of non-setting-based contacts.

However the recruitment methodology for the survey introduced a bias towards childcare attendees for the 0-2 year olds. Population data [263] shows formal childcare enrolment of 93%-98% in 3 and 4 year olds (UK, 2012) and 35% for 0-2 year olds (UK, 2013). Combined with data on daycare showing attendance of 2 year olds at 80% of that of 3 year olds [266], we therefore simplify to an assumption of universal childcare attendance by 2-4 year olds, and non-attendance by 0-1 year olds. (These population datasets [263, 266] do include robust figures for the size of different types of formal childcare settings, however the data is for the setting as a whole - not for the room, which the proximity required under the transmission opportunity definition – so these datasets cannot be directly used for parameterisation.)

We therefore perform the MVD and degree distribution fitting, stratified by the age-groups identified above, using data excluding childcare contacts and (given the only child model assumption) sibling contacts. Similarly to the adult data, we fit candidate distributions to the (non-childcare) observed stratified distributions, using the same set of candidate distributions and fitting methodology (parameterise using maximum likelihood estimation and compare fit using AIC). We also determine an appropriate distribution to model the number of room contacts (for age 2-4 years only), with a different set of candidate parameterisations based on visual inspection of the data.

5.2.1.4. Decision mathematical representation

In the decision model, binary vaccination opinions (support adherence or not) are initially randomly allocated to each vertex, before being acted upon by an information-sharing process. We assume this information (the opinion status of each adjacent vertex) is collated and the opinion status updated according to a decision algorithm based upon this evidence. As in the previous modelling (Chapter 3) the mathematical representation of this algorithm is

the probability that the opinion status of the vertex changes and is assumed to be a function of either the proportion of all adjacent vertices which have the opposing opinion status to the vertex ('fraction' f) or the count of adjacent vertices which have the opposing opinion status ('count' c). We represent the two states by A and H (adhere and non-adhere, i.e. "hesitate", respectively). Hence, f_A is the fraction of adjacent vertices with state A; and similarly for f_H, c_A and c_H.

We first visually investigate the observed state of nodes which have been categorised by 'quantity' of alters with a specified state. We consider node categorisation by the 'fraction' and by the 'count' measures of its alters. In addition to separate analysis of A and H states, we combine both analyses – defining each node within each 'quantity' category as having the "Same" (or opposing) state as that which resulted in that node having been allocated to that category (each node providing two values, one from its categorisation under each state).

Then using the same categorisations by 'quantity' of alters, we fit response function parameters for the observed marginal proportions of the opinion state. As an unknown variable we also allow a_{init} to vary in finding the best fit.

In order to estimate the format of 'response to evidence' function – here denoted as $r(\,)$ – and any associated parameters, we assume the data is the collation of results from identical, independent decisions for each (ego) node, with the ego's initial opinion having been randomly allocated via a Bernoulli trial with a probability a_{init} of an "Adhere" opinion. The distributions of alters and their opinion states are taken from the data. We apply the candidate function to each (ego) vertex to obtain the distribution of final opinion state.

An "Adhere" state ego vertex in the observed data arises from either an initial state A experiencing 'no change' on applying the response function or an initial state H experiencing 'change' on applying the response function (and conversely for a Hesitate state ego vertex). The candidate responses functions used are those in the previous modelling with summary recap in Table 5-1) and an identical response function is used for both decision directions, as data are insufficient to permit robust separate analysis by direction. Hence, the proportion of vertices expected to have an adhere opinion after a decision is given by Equation 5-3a(i) for a fraction-based algorithm (acting on $\{f_A, f_H\}$ for each vertex), and Equation 5-3a(ii) for count-based algorithms (acting on $\{c_A, c_H\}$). Expressions for the proportion of vertices we expect to observe with a Hesitate opinion state are similarly derived.

For the marginal distributions, we prefer an expression with a single alters-based variable (as in the response function itself). For 'fraction'-based algorithm we can simplify to obtain

the observed Adhere proportion as a function of f_A (fraction of adjacent vertices with state A) (Equation 5-3b(i)). We note that, unlike the pairings $\{f_A, f_H\}$ which are uniquely determined as $f_H + f_A = 1$ for each vertex, a given value of c_A may be paired with multiple values of c_H (dependent of the vertex degree, k). Hence the degree of the vertex is present in the equivalent expression for 'count'-based algorithm (Equation 5-3b(ii)), so we sum across all vertex degrees to calculate the Adhere proportion.

Table 5-1 Summary recap of candidate responses functions

Algorithm	Response function representation
'majority rule'	$r(f) = \begin{cases} 0 & f < 0.5 \\ 1 & f \geq 0.5 \end{cases}$
'fraction'	$r(f) = f$
'threshold' (parameter α)	$r(c) = \begin{cases} 0 & c < \alpha \\ 1 & c \geq \alpha \end{cases}$
'count' (parameter β)	$r(c) = \exp^{-\beta c}$

Equation 5-3 Expected proportion of post-decision 'Adhere' vertices

a] Proportion of vertices with observed (post-decision) state A

(i) fraction-based algorithm for vertices with alters' states $\{f_A, f_H\}$

$$a_{init}\ [1 - r(f_H)] + (1 - a_{init})\, r(f_A)$$

(ii) count-based algorithm for vertices with alters' advice $\{c_A, c_H\}$

$$a_{init}\ [1 - r(c_H)] + (1 - a_{init})\, r(c_A)$$

b] Proportion of state A (post-decision) in vertices with a given level of 'adhere' alters

(i) fraction-based algorithm for vertices with alters' states $f_A, = f$

$$a_{init}\ [1 - r(1 - f)] + (1 - a_{init})\, r(f)$$

(ii) count-based algorithm for vertices with alters' advice $c_A = c$ and degree k

$$a_{init}\ [1 - r(k - c)] + (1 - a_{init})\, r(c)$$

5.2.2. Results

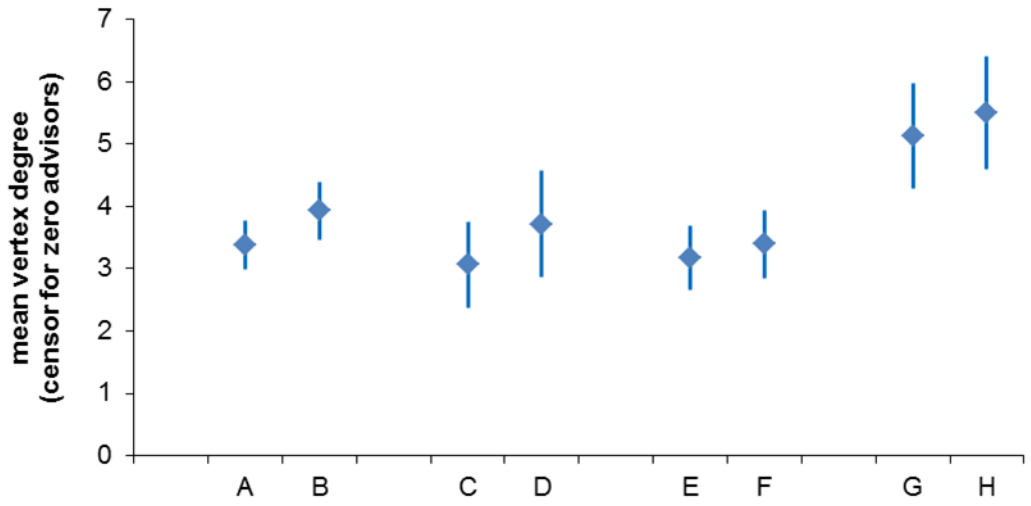
5.2.2.1. Decision model network structure

We initially explore the stratified data, and calculate estimates for the "solo" and "joint" options of treating "couples" vertices' egocentric network. From Figure 5-2, the range of

plausible MVD values (censoring for nodes with degree zero) is 2.4 - 4.6 for lone parents, 2.5-3.6 for couples (assuming solo decisions) and 4.2 - 6.4 for couples (assume joint decisions).

Figure 5-2 Information network – MVD estimation

Mean vertex degree (95% CI) – censored for zero advisors



Key		Format	No reported Advisors	n
	Raw data			
A	All responses	All	16%	161
B	All responses	Paper	12%	118
	Lone parent estimate			
C	Data: respondents with advisors,	All	-	48
D	and partner not reported as advisor	Paper	-	35
	Couples estimates - solo			
E	Data: respondents with advisors,	All	20% †	86
F	and partner reported as advisor	Paper	10% †	68
	Couples estimates – joint			
G	Estimated net advisors for both partner	All	22% †	86
H	("E" and "F" re-weighted)	Paper	13% †	68

(Base: all respondents with all advisors categorised by relationship

† zero degree nodes are formed as the partner is excluded from advisor count (as not external to the decision-making unit).

There little evidence to reject null hypotheses that the location of, and that the shape of, the distributions of advisor numbers is different for the lone parent and solo (unweighted) couples types (Mann Whitney, U=1357, p=0.24; Kolmogorov-Smirnov, D=0.582, p=0.88 respectively; paper responses, n=103).

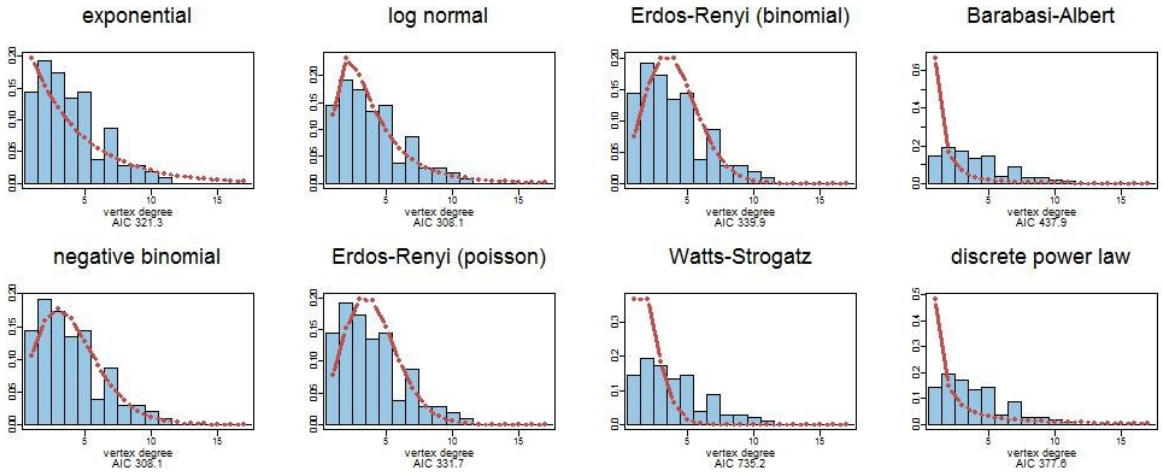
Hence, although there is the suggestion that sample bias maybe possible, under this definition for couples' decision-making such a bias will have little effect on the network structure parameterisation and so there would be little need to stratify the sample in degree distribution fitting below (and the model-building, §5.3). Conversely under the joint discussion assumption for couples, we retain separate distributions for lone parents and couples.

Fitted candidate degree distributions are shown in Figure 5-3, fitted to unstratified data (under the "solo" assumption, this is used for both lone parent and couple), and data for each of the lone parent and couple (joint) assumption. The small world and scale free distributions are the poorest degree distributions fits of those considered, as judged by the AIC. The data are best fitted by a lognormal or negative binomial distribution, with the lognormal marginally better; this is the case for all datasets. Parameter values for the best fit lognormal fits, for the datasets still under consideration are in Table 5-2

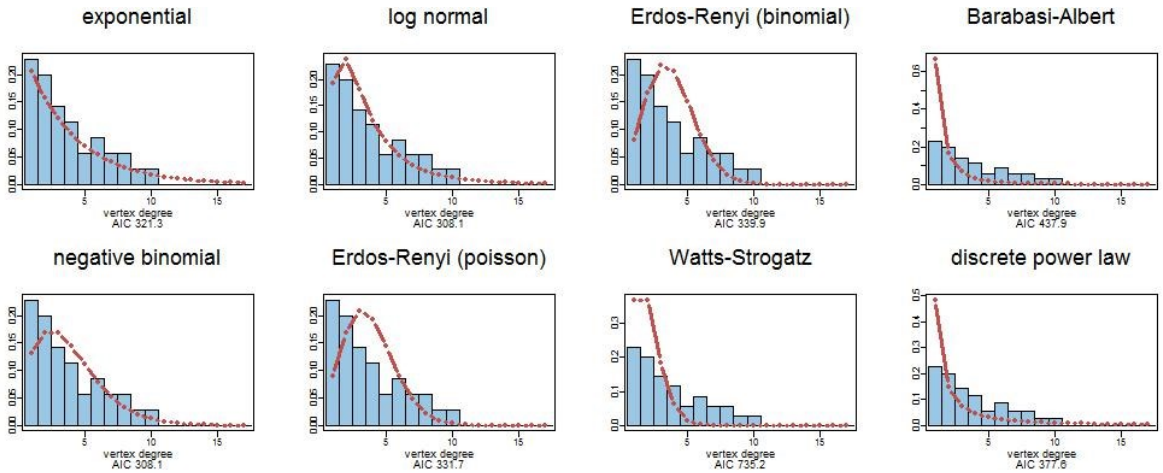
Figure 5-3 Best fits for information network candidate degree distributions

compared with normalised observations of information network survey data

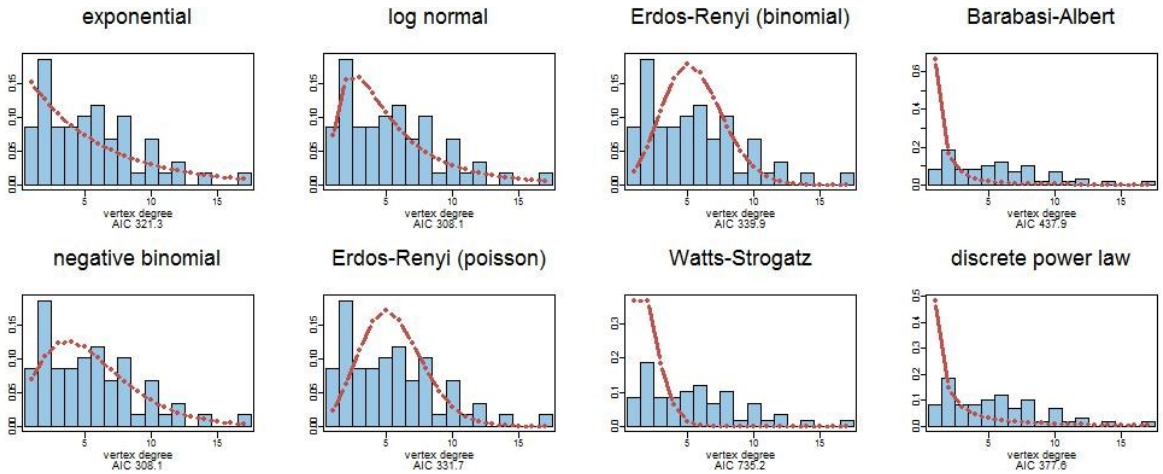
a] all respondents (unstratified data)



b] lone parent estimate



c] couples estimate : assuming joint decision, weighted data



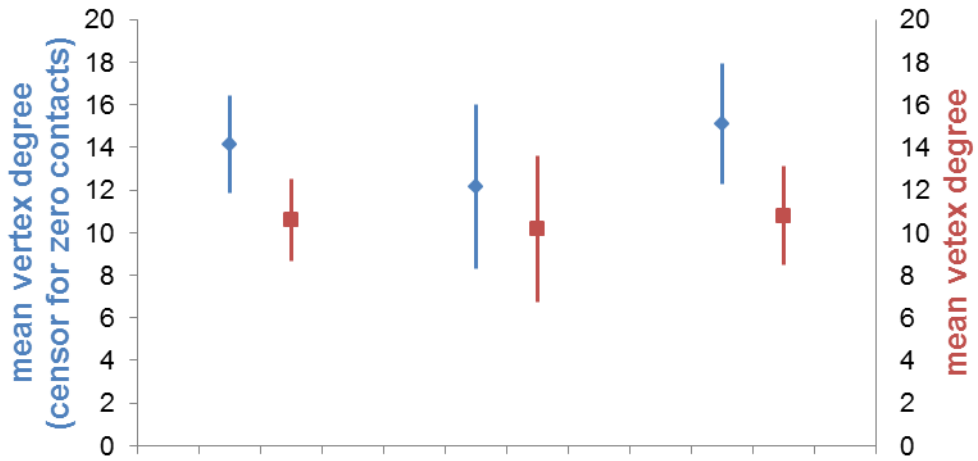
Legend for all sub-plots: fitted values observed values

5.2.2.2. Infection model network structure

Firstly we consider the non-childcare contacts, stratified by age (excluding sibling contacts). The MVD values, stratified by age and censored for zero non-childcare contacts, are 12.2 (95% CI 8.3-16.0) and 15.1 (95% CI 12.3-17.9) for 0-1 and 2-4 years old respectively (Figure 5-4). For model parameterisation, the former are assumed to not attend formal childcare whilst all of the older group do. However there is little evidence that the distributions of their contacts (under 5's) outside of the childcare setting are different, neither in terms of location measure (Mann Whitney, U= 3534, p=0.71) nor shape (Kolmogorov-Smirnov, D=0.761, p=0.61).

Figure 5-4 Potential infection network – MVD estimation (excluding childcare)

Mean vertex degree (95% CI) – contacts not met at childcare nor siblings



Age	0-4 years	0-1 years	2-4 years
No non-childcare contacts	25%	16%	29%
N	188	55	133

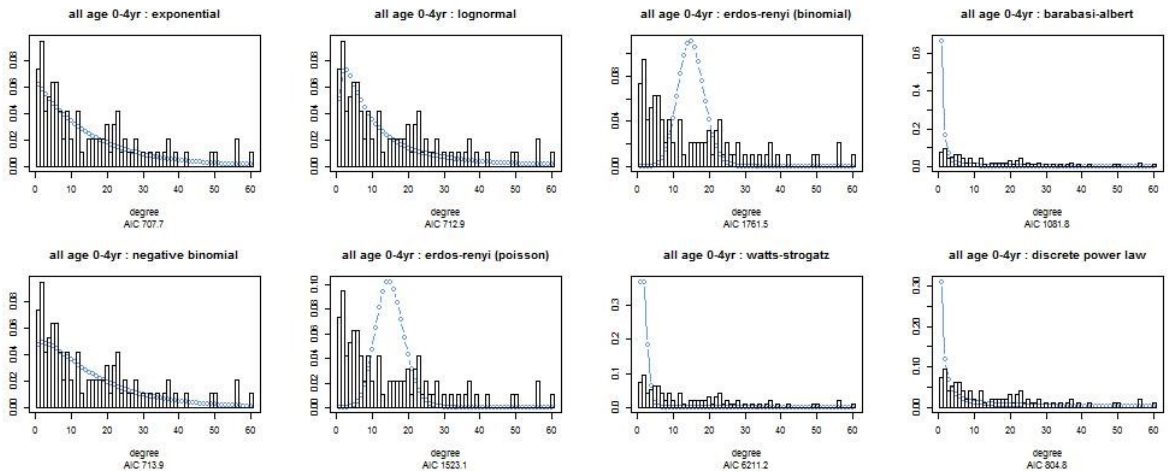
Base: all children with non-sibling contacts

Hence we propose a single fitting for the non-childcare contacts distribution data (see Figure 5-5a). The degree distributions generated by the network generation algorithms (Erdős-Rényi, Watts-Strogatz, Barabási-Albert) offer poorer fits than the other distributions considered, that in turn each perform similarly well on the unstratified zero-censored data (using AIC values). If we relax the zero-contact censorship for these remaining distributions (flooring values to provide support on [0,∞) for the lognormal) the negative binomial provides the best fit (see Figure 5-5b) and so is our preferred mathematical representation for this distribution. Fitted values for the parameters are given in Table 5-2.

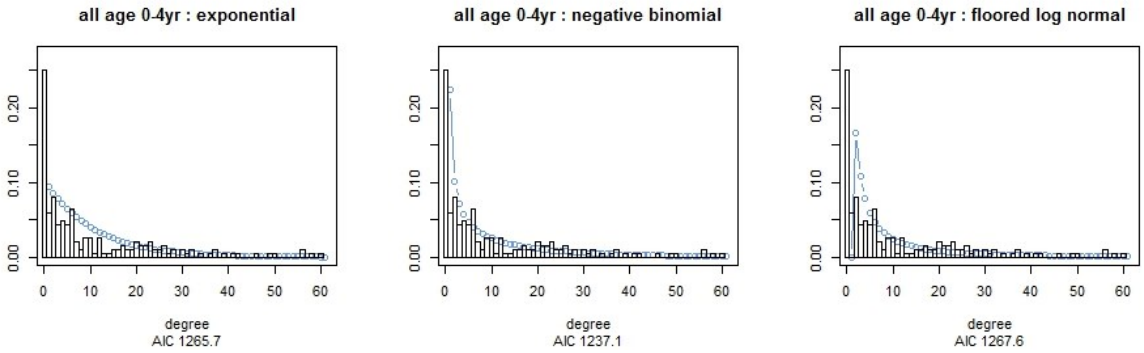
Figure 5-5 Best fits for infection network candidate degree distributions

compared with normalised observations of the non-childcare (non-sibling) contacts survey data

a] all respondents (unstratified data) censored for zero non-childcare contacts



b] all respondents (unstratified data)

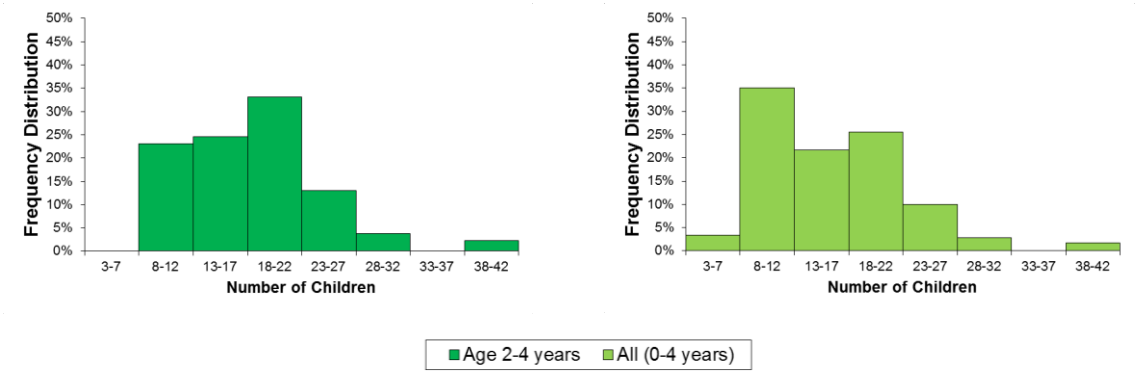


Base: all children with non-sibling contacts

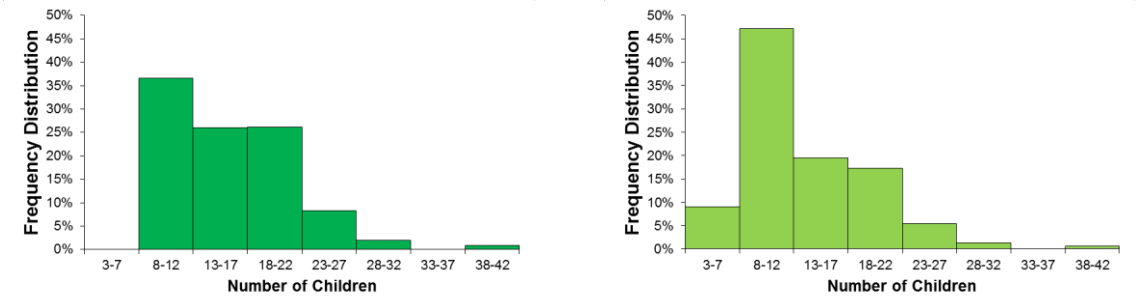
Initial visual inspection of the room-mate data suggests we analyse that binned data, centred on multiples of 5 (Figure 5-6a). The distribution of childcare room contacts is subject to length sampling bias. Hence we adjust the observed frequency by the room-size to estimate the frequency density of room size (attendees) within the population of childcare settings (Figure 5-6b). We fit to the estimated distribution of room sizes for children aged 2-4 years.

Figure 5-6 Childcare room contacts' distribution

a] Room-size as reported by child



b] Estimated distribution of room sizes



Base: all children attending centre (n=180)

Parameter parsimony indicates that approximation my a discrete uniform distribution on {10, 15, 20} is appropriate, and this also outperforms re-scaled geometric and poisson distributions (both latter distributions acting on the domain {2, 3, 4,...} which is then mapped to {10, 15, 20 …}).

The best fit distributions (with corresponding parameter values) for all elements of the network degree distributions fitting are shown in Table 5-2.

Table 5-2 Network structure: best fit degree distributions and parameters

a] Information network:

for each of the alternative assumptions for the couple vertices

Lognormal distribution	Fitted parameter values	
	mean	s.d.
Assuming "solo" network degree for couples:	3.22	1.94
apply same distribution to all vertices		
Assuming "joint" network degree for couples:		
lone and couple vertex types draw from separate degree distributions		
Lone parent estimate	2.88	2.09 }
Joint decision couple estimate	4.35	2.08 }

b] Potential infection network:

for both elements of the overall degree distribution

		Fitted parameter values	
Non-childcare contacts	Size	Prob	(mean)
Negative binomial distribution	0.48	0.04	10.60
Childcare room-mates			
Uniform discrete distribution	Uniform {10, 15, 20}		

5.2.2.3. Decision model response function

From the visual inspection of the observed state vs the quantity of alters with that same state we identify that there are some vertices which preclude fitting by MLE with response functions that pass through the origin (see Appendix). It is possible to adjust model assumptions or response functions to incorporate this (e.g. inclusion of a constant term in the response function, which would correspond to state changes independent of the evidence from alters). Though such changes would enable fitting by MLE, in the interests of parameter and assumption parsimony and given the available data, we do not do so here; we fit using least squares.

The candidate functions and initial random state allocation are fitted to the distributions of observed states, for vertices categorised by alters in states A and H (simultaneously and, for count-based algorithms, for each alters' categorisation separately). The best fit values, based on simultaneous fitting of both sets of marginal proportions, are shown in Figure 5-7.

The fraction-based response functions are poorer fits than the count-based algorithms. We note that the fit of the "majority rules'' function is improved if, instead of a strict majority triggering change (critical value f = 0.50), the critical value is treated as a parameter to be fitted – using a value in [0.75-0.79] produces the best fit for this function format. The "count" response function has a better fit than "threshold" when both marginal distributions are considered simultaneously. The fitted parameters give the following best fit count-based response functions:

"count" algorithm: $r(c) = \exp^{-0.0159c}$ with $a_{init} = 0.979$ †

"threshold" algorithm: $r(c) = \begin{cases} 0 & c < 4 \\ 1 & c \geq 4 \end{cases}$ with $a_{init} = 0.922$ †

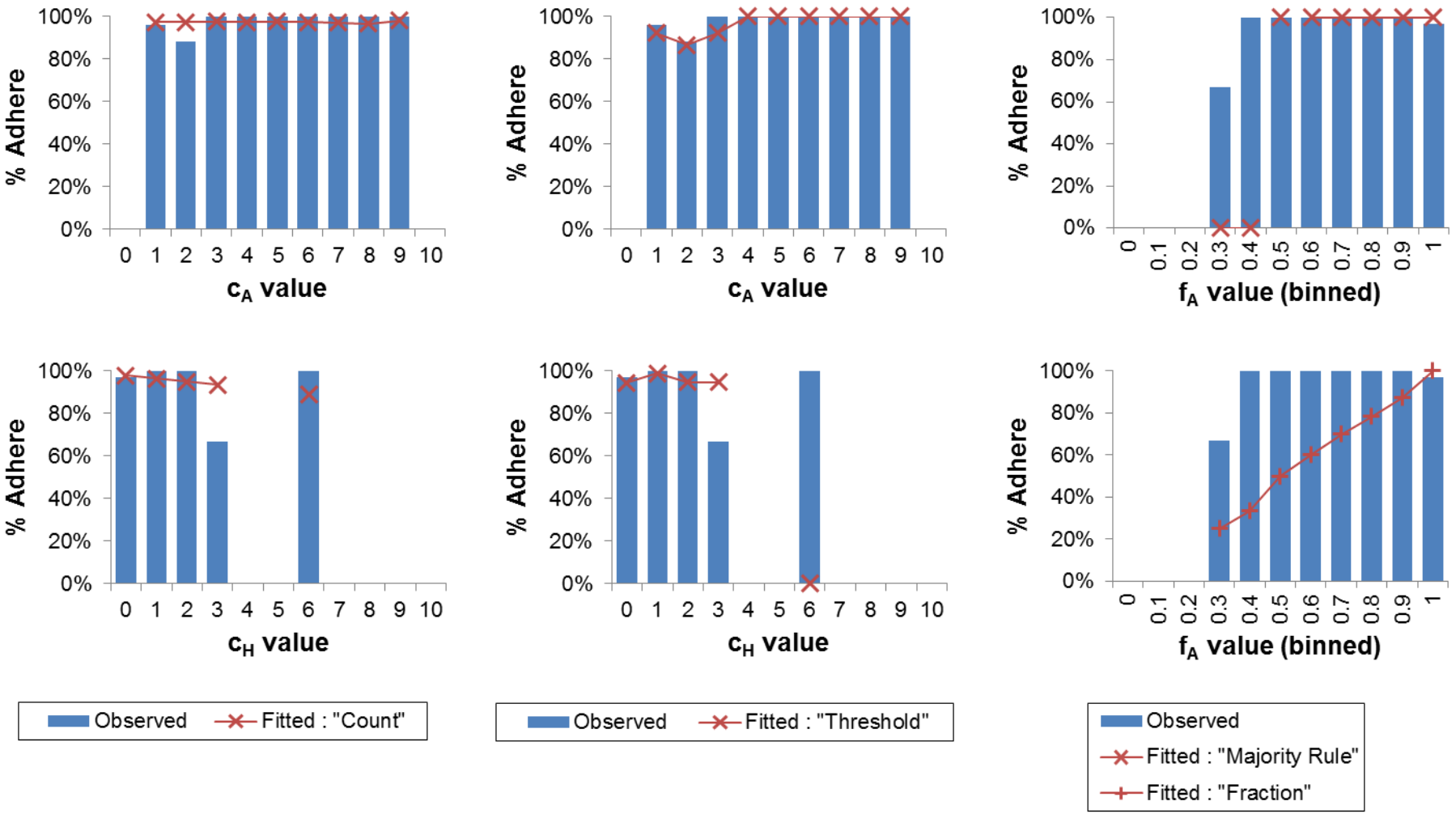
† c.f. Adhere proportion in nodes without alters = 0.867

However, the sample support across the range of the variable c is better for the count of 'Adhere' alters (c_A) than that of the 'Hesitate' alters (c_H), and specifically the value for c_H = 6, is based on a single datapoint. These raise concerns about the robustness of the fit: the former regarding the appropriateness of a simultaneous fitting (c_A and c_H) across the full range of count values, and the latter outlier may exert undue influence. We note that the simultaneously-fitted "threshold" function is a better fit to the c_A marginal distribution that that for the "count" function, and remains the better fit if we instead fit solely for the c_A categorised data. Removing the outlier datapoint (with c_H = 6) causes the "count" response function fit to collapse to a trivial invariant function (e.g. $a_{init} = 1$ or $\beta = 0$), whereas the "threshold" function maintains integrity. Furthermore, the 'threshold' parameter is constant across both marginal fits and the overall fit (with and without the outlier datapoint).

Hence we prefer the "threshold' response function, as the more robust fit to the data.

Figure 5-7 Best fits for response function (expected outcome vs observations)

a] "Count" response function **b] "Threshold" response function** **c] fraction-based response functions**



5.3. Model building

5.3.1. Methods

We build the decision process element of the model following the structure of the previous work, the full set of sequential steps shown in Figure 5-8 (with those of the decision model abridged, the key differences in the decision model element are the restriction of both the network structure options and decision algorithm option, based on the parameter fitting above).

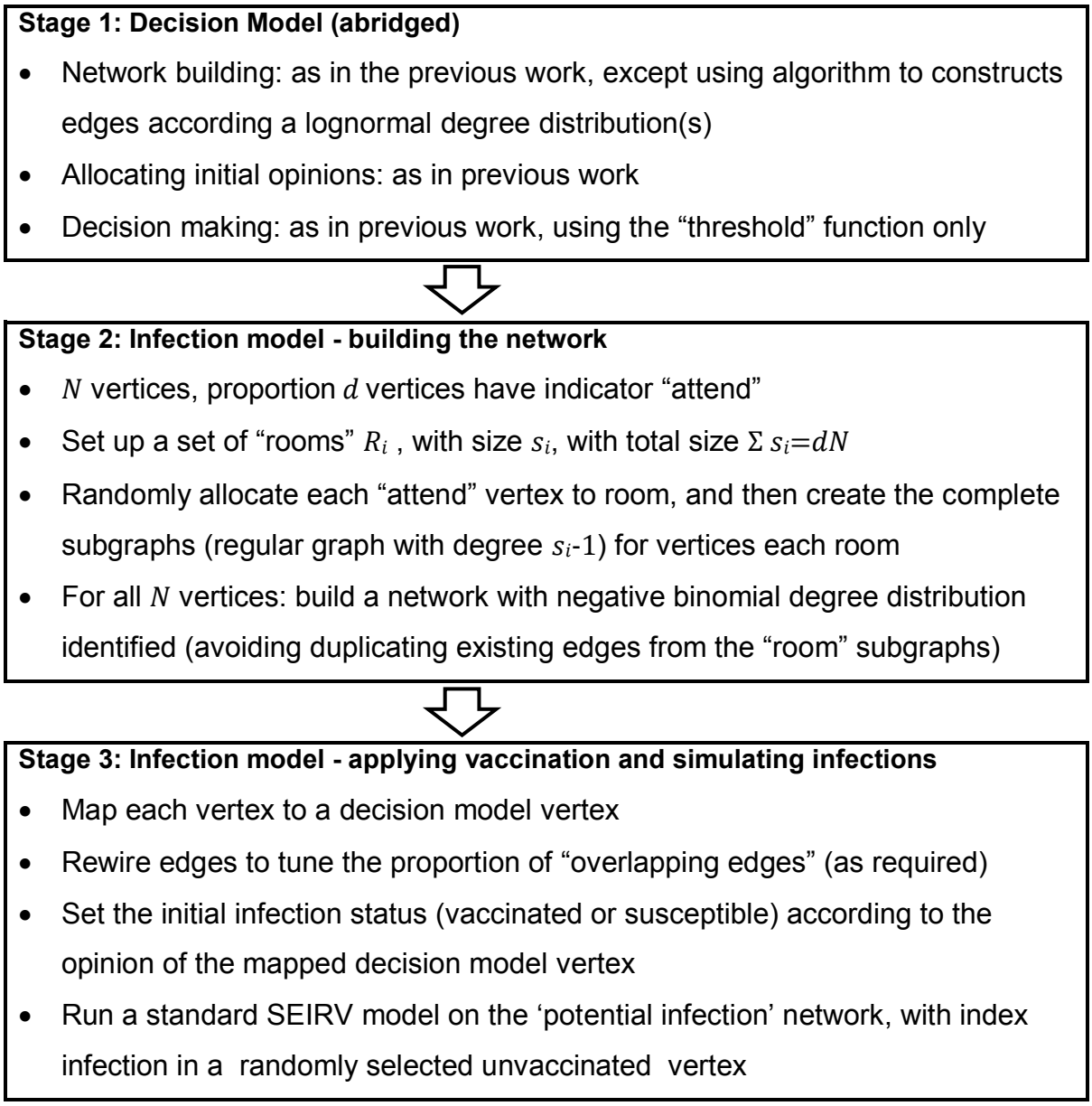
Stage 1: Decision Model (abridged)

We use an algorithm to construct network edges according to a specified degree distribution [288, 289] and use the lognormal distribution identified as appropriate, with two parameters (mean μ and standard deviation σ). Without further data to inform the "net" neighbourhood of a couple, we explore sensitivity between using the single ego type unweighted data (a lognormal distribution) and the weighted data (combining the separately fitted distributions according to the proportion of lone parents, l). From the initial modelling we retain N =4000, and the reporting of the final supporters proportion and intradyad agreement (IDA) as measureable outcomes of this process.

Stage 2: Infection model - building the network

We determine edges in the infection model network in two steps – complete subgraphs representing the dense networks within childcare rooms (proportion d of all vertices included) overlaid with a second degree distribution applied to all vertices. From the parameter fitting in the first part of this chapter, we assume uniform distribution of rooms by size ($\in$ {10, 15, 20}) within the total required room capacity ($\Sigma\ s_i$=dN) and allocate the other set of edges specifying the fitted negative binomial degree distribution. We note that the degree distribution based network generation algorithm [289] has a non-random ordering of vertex degree, so prior to its application, we apply a shuffle to the vertex identifiers (using Fisher-Yates Knuth algorithm [290]) to prevent artefacts relating this degree value and membership of the same "room" subgraph. Additionally we adjust the standard algorithm to avoid doubling edges inherited from the "room" subgraphs.

Figure 5-8 Model stages for one simulation



Stage 3: Infection model – applying vaccination and simulating infections

We put vertices in the two networks in a one-to-one correspondence – representing a child and their parent(s) - and the vaccination opinion arising from the decision model initialises the infection state variable in the infection model (as vaccinated or susceptible). We report the IDA of this initial binary state – calculated for the edges of the infection model network – as a measure of the clustering of susceptible individuals.

The infection dynamics model uses a stochastic SEIRV compartmental model – including a 'Vaccinated' status compartment within a standard epidemiological SEIR (Susceptible, Exposed, Infectious, Recovered) model [34] which acts across the edges of the child

network. The transitions between states are handled using the Gillespie algorithm and use measles natural history parameters [28]. We assume vaccination provides perfect protection (although outbreak reports indicate cases are seen in those with a history of vaccination [53, 174, 175]). A single index infection is introduced and the model runs until all infected individuals have recovered (i.e. any outbreak has run its full course). The proportion of simulations resulting in any secondary infections and the number of secondary infections is reported, as a measure of the outbreak risk within the partially-vaccinated network.

We consider the overlap of edges (pairs of parent and child who both joined by edges in their respective network), prior to the "vaccination". We have been unable to find an algorithm to enable specification of proportion of overlap edges between these two networks with different structures (MVD and degree distribution). Hence we use a pragmatic approach – firstly measuring the overlap proportions "naturally" observed in the simulations. To explore a wider set overlap variables, we identify sets of vertices that may be rewired to increase (or decrease) the numbers of overlapping edges whilst holding the degree of the vertices constant: for example, if vertices {u,v} are joined by a non-overlapping edge, we identify an edge between a second pair {x,y}, such that {u,x} is an overlapping edge, and for which {x,y} and {v,y} share overlap status, see Figure 5-9.

Figure 5-9 Example rewiring to tune the overlap between networks

The red and blue colours represent the binary categorisation of overlap, as determined by the pattern of corresponding edges in the other linked network. This rewiring increases the blue category and reduces the red category (only the rewired edges are shown here, with vertices selected to avoid double edges and self-loops).

.

A list of model parameters is given in Table 5-3 (values taken from the parameter fitting or otherwise stated in the model description). It is noted that we do not include a specific value for the rate of transmission events. Initial exploratory work has been conducted using a generic value, of the order as seen in the models most closely related to this one [168-170]. However having collected, and otherwise sourced, data specific to the MMR and measles in

the UK for all other parameters, it is preferred to conduct the full analysis with the inclusion of a more specific value for this parameter also, ideally calibrating the model vs empirical observations .

Table 5-3 List of model parameters

Decision model

Symbol	Description	Value	
a_{init}	Initial proportion of vaccine support	0.9	♣
α	Threshold response function parameter	4	♦
degree distribution – if couples network unweighted:			
μ	Mean of lognormal distribution	3.22	♦
σ	SD of lognormal distribution	1.94	♦
degree distribution (lognormal) – if couples network weighted:			
μ_L	Mean of lognormal distribution	2.88	♦
σ_L	SD of lognormal distribution	2.09	♦
μ_C	Mean of lognormal distribution	4.35	♦
σ_C	SD of lognormal distribution	2.08	♦
l	Proportion of lone parent vertices	0.80	♥

Network overlap

Symbol	Description	Value
o_{adult}	Proportion of adult-adult edges with overlap	See text
o_{child}	Proportion of child-child edges with overlap	See text

Infection model

Symbol	Description	Value	
s_i	Number of vertices in the subgraph for "room" R_i	$\{10,15,20\}$	♦
w	Size parameter of negative binomial distribution	0.48	♦
p	Probability parameter of negative binomial distribution	0.04	♦
β_{SE}	Rate of transmission event (S→E) (per day)	See text	
β_{SE}	Rate of progression event (E→I) (per day)	$1/\delta_E$	
β_{SE}	Rate of recovery event (I→R) (per day)	$1/\delta_I$	
δ_E	Duration of Exposed state (days)	10	♠
δ_I	Duration of Infected state (days)	7	♠

♣ *estimate from recent trends in COVER data* [42]

♦ *fitted from survey data*

♥ *mid point of range 75% - 85%* [287]-[213]

♠ *PHE green book* [28]

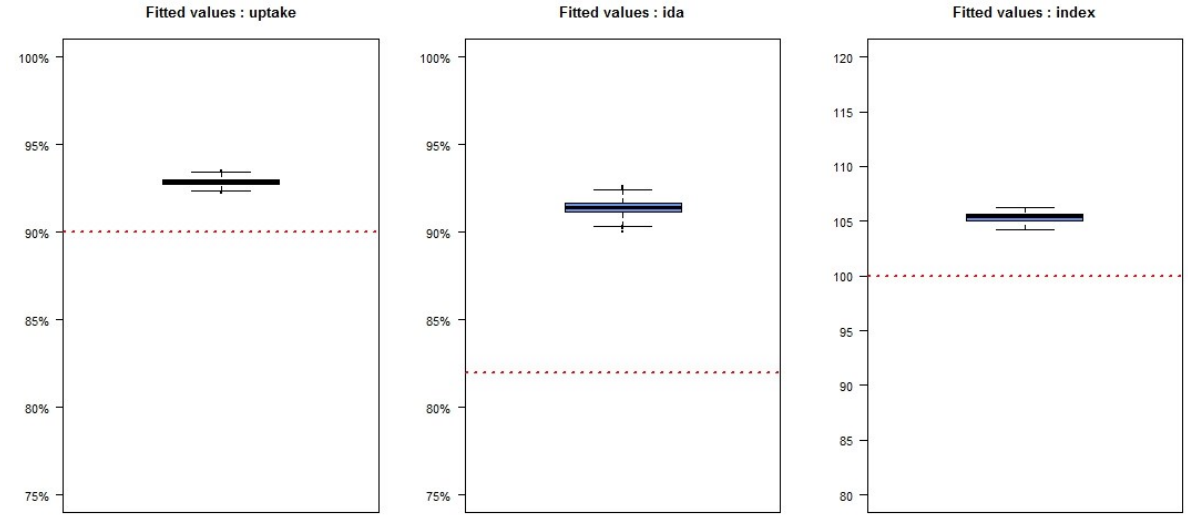
5.3.2. Results

We run the decision model under both possible assumptions for estimating neighbourhood of ‘couple’ type vertices: unstratified degree distribution for all vertices in the information network (assumes ‘solo’ decision network couples) and two stratified degree distributions (assumes ‘joint’ decision network couples). In all cases for each scenario we report results for simulations with 50 network building processes each with 100 decision processes applied (50000 simulations per scenario) extracted from the full model. The parameters values are as specified in Table 5-3. For detailed definitions of the three outcome measures used refer to the initial model (Chapter 3), but in summary we use a measure of proposed vaccine uptake (% vertices supporting vaccination) and two measures of opinion clustering – intradyad agreement (IDA) (proportion of all edges whose vertices have the same opinion stat) and IDA index (as IDA is a function of proportion of supporting vertices, we index the IDA against that expected at the final support levels, simulation-by-simulation). In a scenario where the decision process does not alter the population level measure, we would expect to see values of 90%, 82% and 100 respectively.

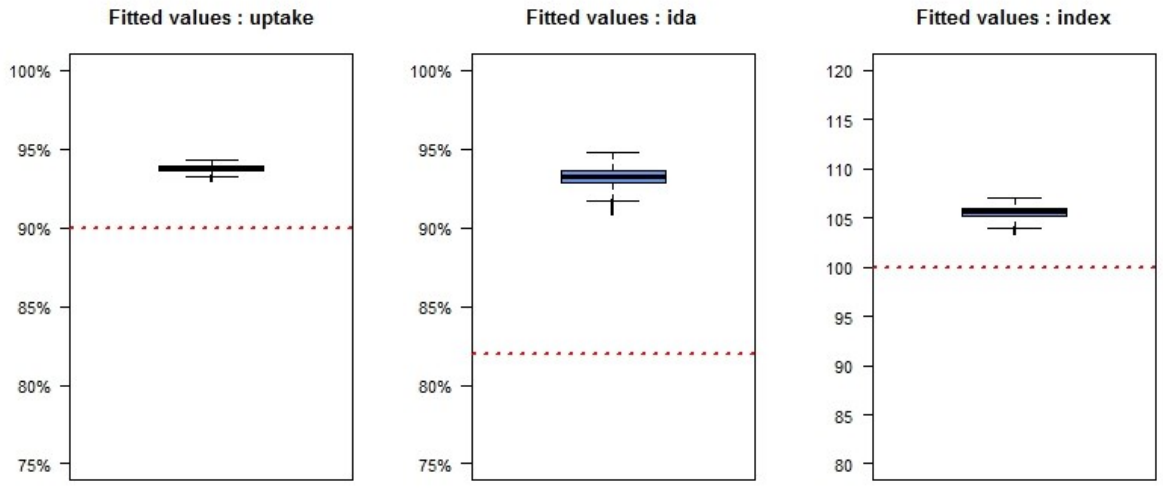
The decision model showed that the decision process increased both the level of support for vaccination and the clustering of opinions, with the clustering increase in excess of that which would be expected for the (post decision) higher support levels in the population (Figure 5-10) This result is independent of the use of stratified or unstratified degree distributions.

Figure 5-10 Summary of post decision process outcomes: vaccine-support (%), intradyad agreement and the index of intradyad agreement vs expected

a] unstratified population (one lognormal degree distribution for all vertices)



b] stratified populations (separate log normal degree distributions, weighted data for couples)



These results were also qualitatively robust under all explored perturbations to the parameter values in the degree distribution to build the network structure. Sensitivity was explored for the "mean" parameter within the set of value containing the fitted value and the whole of the 95% confidence interval (Figure 5-2) of the observed mean of the survey data. We explore this for unstratified distribution (Figure 5-11) and stratified distribution (Figure 5-12) and find the qualitative results are invariant. Additionally in the stratified scenario, varying the proportion of lone parents (within the ranges indicated by the birth and census data in §5.2.1.2) dos not alter the qualitative results (Figure 5-13).

Figure 5-11 Summary of post decision process outcomes – unstratified population, sensitivity to MVD parameter

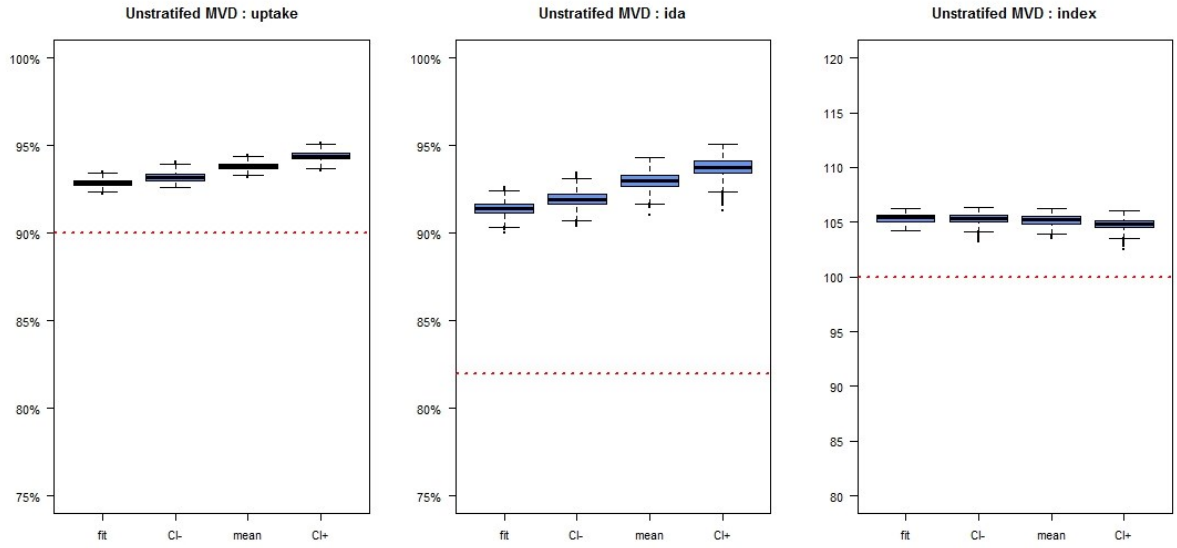
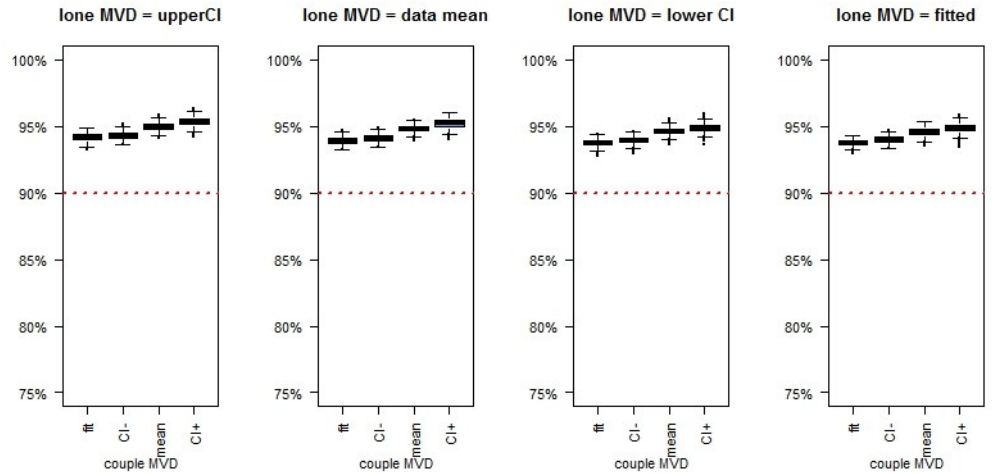
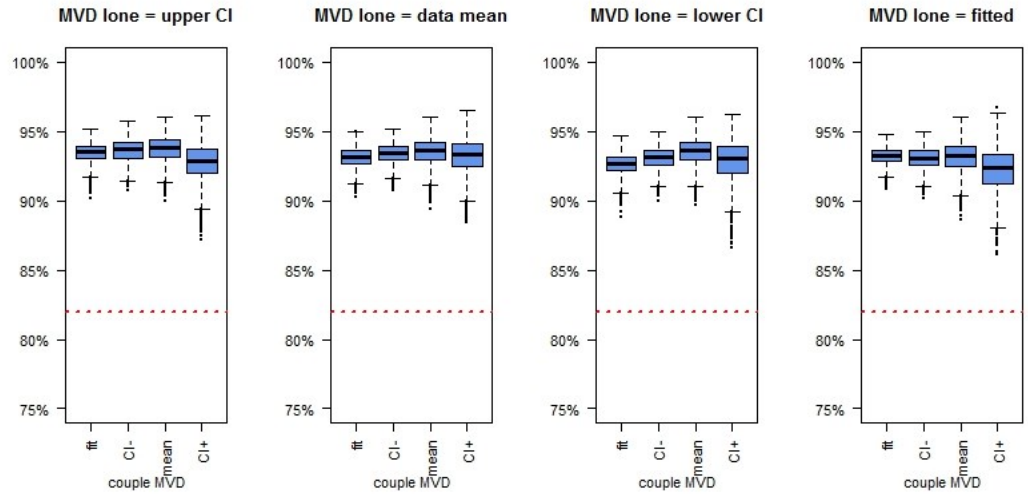


Figure 5-12 Summary of post decision process outcomes – stratified data, sensitivity to MVD in both degree distributions

a[vaccine-support (%)



b] intradyad agreement



c] index of intradyad agreement vs that expected

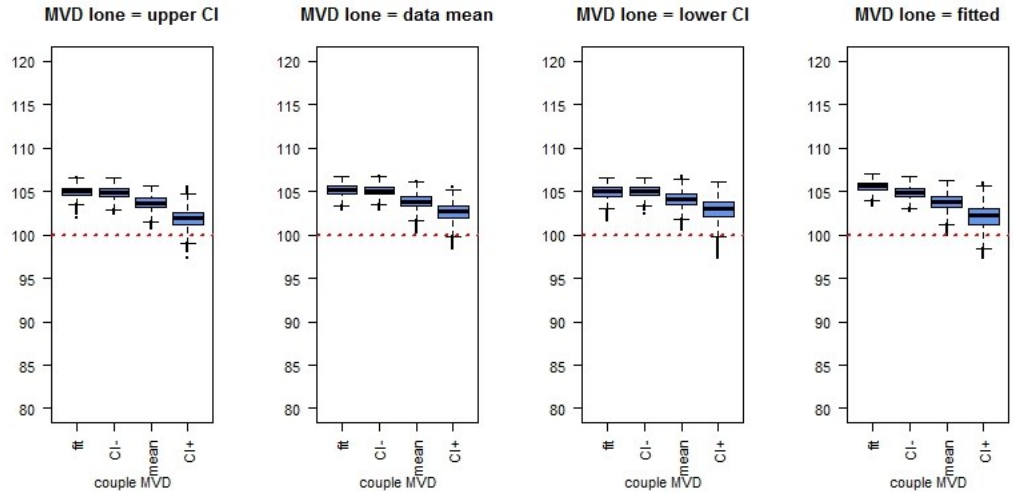
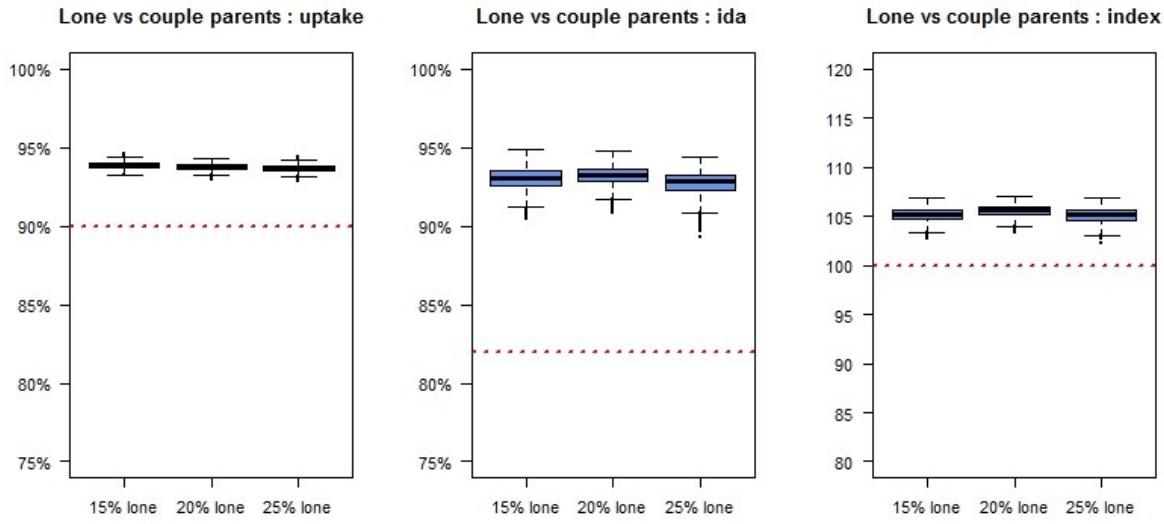


Figure 5-13 Summary of post decision process outcomes – stratified data, under varying proportions of lone and couple parents

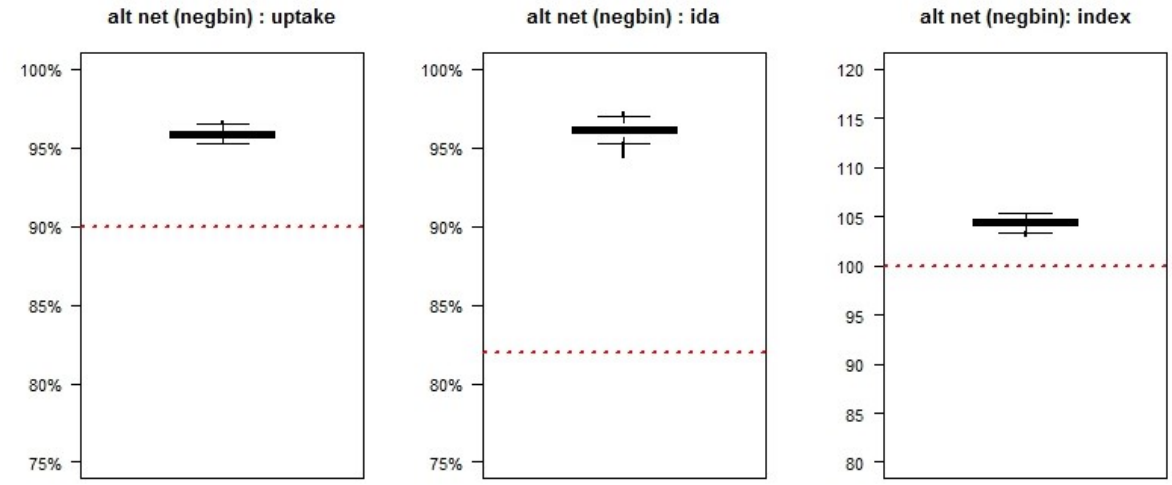


As an additional indication of the strength of these findings we substitute the “second best fitting” adult degree distribution formulation (negative binomial – using the fitted ‘size’ and ‘prob’ parameter values from §5.2.2.1) in the network building, and the “second best fitting” response function (count, using fitted value for β) in the decision rep representation. The qualitative results for all three measures are unaffected under a change of degree distribution, and the clustering measures are also increased under the change of response function, albeit the increase is cover is not observed (Figure 5-14)

Figure 5-14 Summary of post decision-making outcomes – stratified data, under the "second best fitting" candidates for network structure and for decision algorithm

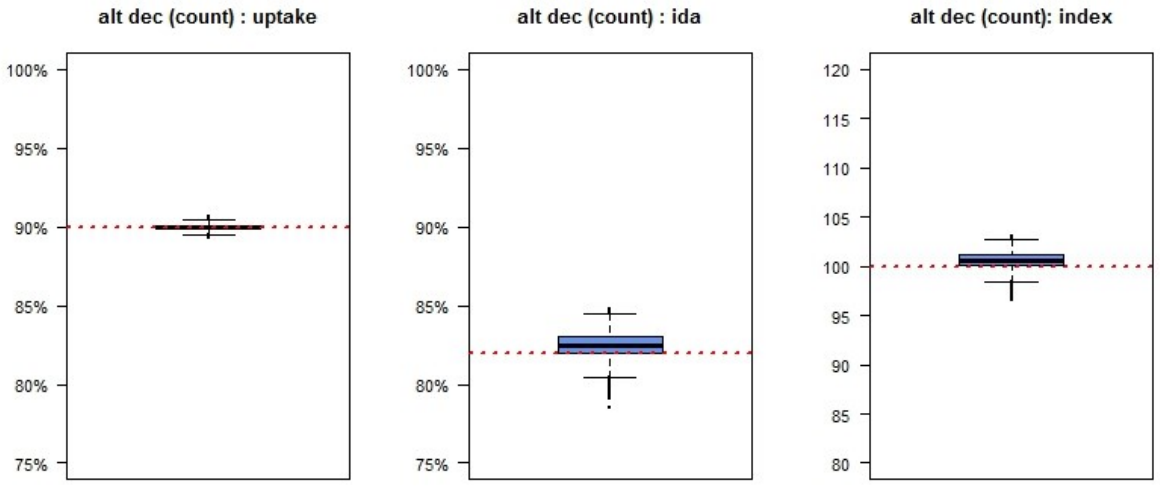
a] "second best fitting" for network structure

using a negative binomial distribution for the network degree distribution



b] "second best fitting" for decision algorithm

using a 'count' algorithm as the response function in the decision-making simulation



5.4. Discussion

Parameter fitting

Unlike the previously published vaccination decision mathematical models identified in Chapter 1 [166-170], we have identified the decision algorithm format and the network structure across which it acts from parameter fitting of empirical data. We have also fitted the network structure for the coupled infection network, unlike the highlighted study with separate information and infection networks [168] they have not been restricted to a shared structure but is also determined by fitting from empirical data.

The response function in the decision modelling that best fits the observed pattern of schedule adherence in the survey data uses the "threshold" decision algorithm, which is a function of the count of adjacent vertices (with the opposing opinion state). This implies the type of complex contagion in this scenario is "uncontested" [119], with the decision depending on the numbers of alters presenting advice to change one's mind (and is independent of the amount of advice received to not change opinion). We note that the "count" function (the investigated function with next best goodness-of-fit) also represents uncontested contagion, and both fraction-based response functions offered poor fits, suggesting that the type of complex contagion is a stronger factor in the fitting than the exact choice of function form.

The lognormal and negative binomial distributions, which are best fits to the observed networks' degree distribution data (information and potential infection networks respectively), both have precedents in the all-age social contact datasets referred to in Chapter 4. Danon et al [235] found the BSCS data was well fitted with a lognormal in the body and power law in the tail when their grouped contact response were included, and a negative binomial when only individually listed contacts were considered, and POLYMOD data was fitted by a negative binomial [234]. Hence the distribution for the Information network is not unlike those previous observed for social contact networks, despite the different nature of the connection represented (physical proximity is not a necessary for information transfer).

The distribution fittings are based on degree distribution, other structural characteristics such as transitivity or shortest path length have not been considered, although the nature of these structures can influence behaviour of dynamical systems operating on the networks. To include other structure characteristics to discriminate between candidate network representations requires more information than that available from ego-centric data. That said, through fitting to subsets of the child contact datasets we are able to retain an

observed feature of the potential infection network (fully connected subgraphs in the child network) which a random allocation of edges within a relatively sparse network would not consistently reproduce.

Model

We demonstrate that a measles vaccination decision process, represented by an empirically informed response function, across an information-sharing network, with structural parameters also fitted to empirical observations, can create qualitative changes in both the population levels of vaccine support and the amount of opinion clustering in the population. , which originally had randomly allocated opinions (and support proportion at 5% below the WHO MMR target uptake [33]). This result is robust across a plausible range of network structure parameterisations. Furthermore, not only are all elements of the model parameterised based on empirical observations, but the structure and decision dynamic choices are made from single-source data.

Our previous modelling (Chapter 3) examined the fitted decision function ("threshold" with $\alpha=4$) but the distribution identified for the information network (lognormal) was not one examined. We note that at MVD close to the mean value for the fitted degree distributions (3.32 or 2.88-4.35 for the range of assumptions on the couple's egocentric decision network) we found increases in coverage, post-decision, for all examined network structures, as is also the case for the empirically fitted network. However there was no agreement on the qualitative behaviour of the opinion clustering measures (IDA and IDA index), with both increases and decreased observed. The empirically fitted network produces clustering measures qualitatively similar to those found for the small world network type. The is also an indication that the cluster-increasing effect(as measured by IDA index) is stronger on the empirically informed network, although direct comparison with the previous work is not possible due to different values of MVD.

Hence for empirically-informed values for mean of advisors and decision function the outcome of the modelled process is sensitive to network structure assumptions. Despite the lognormal's "long tail" the both vaccine support and cluster measure dynamics are opposite to those seen on the power law Barabási-Albert network, nor do we see the type of dynamics reported for the networks generated by the Erdős-Rényi algorithm. Indeed the closest comparison is with the Watts-Strogatz network, which had the poorest fit with the observed information network degree distribution of the 8 candidates examined in the parameter fitting process. This observation further supports the original decision to explore

the sensitivity to network assumptions and use empirical data to inform the model parameters.

The increased support for vaccination would be expected to decrease the risk of outbreaks in a population vaccinated in the same pattern. However the clustering effect (especially as higher than expected for the level of vaccine support) would be expected to increase the outbreaks. These contradicting naïve inferences from the measured results, bolster the research decision to not proceed with the infection stage of the full model using an arbitrary transmission – it is likely that the final outcome in terms of infection outbreaks will be sensitive to this value.

In the fitting process, two areas were identified where assumptions were made in the absence of sufficient data. Firstly, data were unavailable on the combined sources that a co-parenting couple draw upon to make a joint decision. However sensitivity analysis of the decision model outcomes indicates that this does not need to be a priority in further data collection, as the results were qualitatively invariant to plausible adjustments to size of the couples network neighbourhood. Although we caution that we do not yet know how the quantitative differences might affect the impact of the opinion clusters in the infection part of the model.

The second area where the data were insufficient for a confident interpretation was the choice of response function. The ability of the decision process to change population level support is sensitive to the choice made in the parameter fitting stage (between “threshold” and “count” algorithms), but the ability to increase clustering is not. Hence under the alternative functional form for decision algorithm the inferred opportunity for this pattern of opinions to affect outbreak risk remains, albeit less so than under the threshold algorithm.

The collection of more data to revisit the response function formulation remains desirable. The sample sizes (especially of alters proposing non-adherence to the routine vaccination) were smaller for the higher numbers within the ego’s neighbourhood – and it is at these values that one might expect their influence to be felt most strongly. The fits are obtained primarily from low-integer ‘opposition counts’, and we have assumed a monotonic relationship which continues across the higher-valued ‘opposition counts’. Data collection that purposively oversamples the vaccine-hesitant advisors would be necessary to provide suitable samples for the higher values of local anti-vaccination sentiment, based on current population levels of vaccination support. Furthermore these data may be able to inform a two-way response function, dependent on the initial state, and also address the issue of “spontaneous” decisions raised above (§5.2.1.4) in the context of outliers at zero-opposing alters,.

Within the decision model we have assumed a random distribution of initial opinions, further work to explore the affect that pre-existing homophily has on the ability of the modelled decision process to alter opinion clustering would be valuable. We similarly note the inherent assumptions about binary decision states and static networks. The data could be used to generate a 3 category ordinal scale of vaccine support corresponding to the 3-way categorisation used in part of Chapter 4 (non-adhere splits to “delay” and “refuse”), but the sample sizes would be insufficient to support the parameter fitting exercise.

However, the natural priority for further work relating to the full model, is to determine an empirically-valid value for the transmission parameter to investigate the predicted outbreak risk.

6. Concluding remarks

This thesis has explored the subject of information-sharing on social networks and its potential influence on participation in routine child vaccinations using a range of techniques: synthesis of published studies (quantitative and qualitative), statistical analysis of existing data (both published and unpublished), mathematical models (simulating the dynamics of two processes: transmission of information and of infection), and collection of empirical data (to our knowledge, the largest survey of UK preschool social contacts).

These combine to both confirm some proposed aspects within our hypothesis and its framework – such as the variation in MMR 1 uptake on a small spatial scale within the community (Chapter 2) – and to address areas where relevant empirical data was lacking in the literature – such as the networks of both information-sharing parents and of preschool children in required physical proximity for measles transmission (Chapter 4) - and the final decision model (Chapter 5) provides some parametrically robust results directly on the hypothesised effect on vaccination options, notably the increase in the clustering measure. This evidence of vaccine-related status clustering, is consistent with other studies which have observed assortative mixing on networks, by vaccine-related categorisation: message network of twitter users' opinions on the introduction of pandemic influenza (H1 N1) vaccine [255], advice network of households in India on polio vaccine hesitancy [291], close contact network of USA school students by seasonal influenza vaccination.[292]. We are aware of only one other dataset regarding parents opinions on their child's vaccination – from the USA [271] - but to our knowledge there are no other investigations with linked data from both parent and child networks, nor from other voluntary childhood vaccination contexts.

However the results of the initial work with the final mathematical model are not clear-cut in terms of clustering of vaccination status (in children), moderated as that is by the overlap between the two networks (a non-trivial question given they differ in degree distribution function and MVD). Assuming the increased opinion-clustering is sufficiently large or well-patterned – in some way, as yet not understood - to be transferred into the child contact network, there is the aspect of the contradictory effects on infection dynamics of increases vaccine support and increased clustering of susceptible individuals. There remains the opportunity to continue the exploration of the current model construct and to revisit some assumptions and increase the validation with empirical data, and other opportunities to extend the existing work have been identified at each stage.

The strengths, weaknesses and implications of the individual elements of the work have been discussed throughout the thesis. Here we comment on the overall combined strengths and weakness and implications.

The multifaceted nature of the approach to the hypothesis strengthens the phenomenon that was observed in each of the arms of the hypothesis explorations included in these research programme: that of “small scale” clustering (albeit with differing concepts of “small scales” and measuring this phenomenon in one or both populations of interest).

Each of the three main techniques employed to address the overall hypothesis has strengths and weakness, but in some aspects these are balanced out across the piece. For example, we have analysed two datasets of parents and their children’s MMR status: in Chapters 2 and 4. The data used in Chapter 2 are subject to ecological bias, but have the advantage of low sampling, coverage and recall biases through use of, mainly, census data and uptake values ultimately sourced from GP records; by contrast the data collected and analysed in Chapter 4 is a sample (with inherent challenges faced to minimise sampling and coverage biases), and used a recalled measure of vaccine status, but with values for individuals hence without ecological bias.

Some weaknesses remain despite investigation with different tools, for example the intra-dyad agreement in the survey (Chapter 4) and mathematical modelling (Chapters 3 and 5). The survey measurement includes inferred opinion status which is determined by inferences from recalled behaviour (of the vaccination advice received from alter, and presenting one’s own child for vaccination) both indirect measurements subject to recall bias, and the model is a model - designed to investigate the dynamics of a particular theoretical system. It shows the possibility seeing such intra-dyad agreements, but it cannot compensate directly for the biases in the data collection. Vice versa, the survey observations cannot attribute causality to the theorised dynamic, in this regard they are solely evidence that does not contradict the model’s predictions.

What does the combined evidence to support the hypothesis (that peer-to-peer information sharing influences the clustering of opinions) mean in terms of interventions? Two key initial points: we have investigated the medium through which the information is transmitted, much of the intervention work focuses on effective messages to transmit. Beyond comments made previously (Chapter 2) about education levels and communication style, this work is not well suited to contribute to the discussion on the content communicated and its presentation. Secondly, we do not have clear evidence how the information process – if indeed it actually does contribute to opinion-clustering – affects outbreaks of the vaccine-preventable disease. The final stages of the model in Chapter 5 have not been

simulated with realistic infection transmission parameters and the theoretical increase in cover resulting from the decision process is expected to act to protect the population, unlike a rise in clustering which is expected to increase risk. Hence, until we do know the final theoretical outcome on morbidity of the hypothesised information-sharing process, it may be one that we wish to encourage or to discourage, or (in terms of morbidity dynamics) be of little interest, affecting patterns of opinion only. For each of these plausible scenarios techniques will be different, with health promotion vs health protection mindsets.

If it is a process we wish to intervene against, there is an existing body of work on how information (primarily, the proximity of cases) can act across networks to amend behaviours in attempt to avoid infection (action by actors) [157] and also work to identify key individuals in a social network [297] to vaccinate, and so break the chain of infection (or more straightforward techniques such as ring vaccination). Are these concepts transferrable to controlling the spread sentiment in an information network where network members may be less inclined to be proactive/co-operate than with the more concrete effect of disease) and do these techniques developed for simple contagion also work in the case of complex contagion? Alternatively we could target a later stage in the opinion-vaccination-infection process, for example the as-yet unknown overlap would suggest a weak point based on Eames 2009 [168].

Or we may find the process – through higher pro-vaccine sentiment - increases population protection despite the clustering of opinions. Should it therefore be actively encouraged? However, and in common with the neutral morbidity scenario, what are the long-term implications of the remaining clusters of anti-vaccine sentiment?, Perhaps we are already seeing an effect whereby this sentiment, as anti-normative, causes those who are hesitant to act differently: seeking information not from peers in a local network but from other sources (recalling the lower mean advisors for non-adherents in chapter 4 – but the USA study [271] has the opposite finding).

Does the clustering mean that the hesitancy sentiment thus perpetuates, acting as a reservoir for the next scare to exploit? Also the research has quantified the role the grandparents as advisors for current new parents – does the current cohort carry those opinions with them into the future forming a basis for their influence on the process if their child becomes a parent? Also we have only considered the measles outcomes in preschool children, might future health concerns may include the other diseases that MMR protects against (mumps, rubella)? The current models and cross-sectional studies are not designed to be useful for time periods over which the networks may no longer be regarded as static and the at-risk groups no included in the sample base.

However extending and adapting the mathematical model to one with cohorts of new parents joining and leaving as their child ages could explore these questions of longer-term impact of contemporary decisions. Our measures of opinion are also only binary – a more sensitive scale may reveal more subtle dynamics, with borderline adherents potentially making the system less resilient to shocks.

The question of protection against measles via MMR uptake remains topical, and a continuing challenge for health protection in the UK: for example, recent mainstream UK media reports include both the ongoing large measles outbreak in Italy (and introduction of compulsory vaccination [293]) and also the latest activities of the lead author of the now retracted MMR-autism link Lancet paper [294], and during the lifespan of this thesis, the UK has experienced the largest measles outbreak since the full introduction of routine MMR vaccination, and the national MMR uptake rate has recovered from the trough associated with the vaccine scare in the early 2000's but remains below WHO guidelines. It is a subject of interest to parents trying to make good decisions on behalf of their child and to health service professionals working with individuals and the community, and this thesis is concerned with those personal decisions and their communal effect. It is hoped that the work presented here contributes some useful data, analysis and insight to the debate.

7. References

1. Health and Social Care Act 2012, c. 7. London: Her Majesty's Stationery Office; 2012.
2. NHS Confederation. *About CCGs - NHS Clinical Commissioners*; [Date Accessed February 2017]. Available from: https://www.nhscc.org/ccgs/.
3. Public Health England. *About us - Public Health England* London; [Date Accessed February 2017]. Available from: https://www.gov.uk/government/organisations/public-health-england/about.
4. Public Health England. *Cover of Vaccination Evaluated Rapidly (COVER) Change Request*. London: Public Health England; 2014.
5. Public Health England. *UK immunisation schedule: the green book, chapter 11*. 2016. In: Immunisation against infectious disease: the green book [Internet]. London: GOV.UK. Available from: https://www.gov.uk/government/publications/immunisation-schedule-the-green-book-chapter-11.
6. Children immunisation Centre. *Children Immunisation Centre*; [Date Accessed April 2017]. Available from: https://childrenimmunisation.wordpress.com/home/.
7. BabyJabs. *BabyJabs offering you an informed choice of vaccines for your child*; [Date Accessed April 2017]. Available from: http://www.babyjabs.co.uk/.
8. Sonnenberg P, Crowcroft NS, White JM, Ramsay ME. *The contribution of single antigen measles, mumps and rubella vaccines to immunity to these infections in England and Wales*. Arch Dis Child. 2007;92(9):786-9.
9. Public Health England. *COVER Programme A guide to submitting data*. London; 2016.
10. Health Protection Agency. *COVER Methods*; 2011 [Date Accessed April 211]. Available from: http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/VaccineCoverageAndCOVER/COVERMethods/.

11. N. H. S. England. *Data Flow for Direct Commissioning Child immunisations (COVER_ UNIFY) Collections: Guidance*. 2016. 1.54
12. Health Protection Agency. *COVER Request Parameters Q11-1 Apr to Jun 2011.*; 2011 [Date Accessed July 2011]. Available from: www.hpa.org.uk/infections/topics_az/cover/default.htm.
13. Information Services Division NHSNSS. *Immunisation* Edinburgh; [Date Accessed April 2011]. Available from: http://www.isdscotland.org/Health-Topics/Child-Health/Immunisation/.
14. Public Health Wales Health Protection Division. *National immunisation uptake data*; [Date Accessed April 2011]. Available from: http://www.wales.nhs.uk/sites3/page.cfm?orgid=457&pid=54144.
15. Public Health Agency. *Vaccination coverage* Belfast; [Date Accessed April 2011]. Available from: http://www.publichealth.hscni.net/directorate-public-health/health-protection/vaccination-coverage.
16. Public Health England. *COVER: vaccine coverage data submissions and publications schedule* London; 2016 [Date Accessed February 2017]. Available from: https://www.gov.uk/government/publications/vaccine-coverage-statistics-publication-dates/cover-vaccine-coverage-data-submission-and-publication-schedule.
17. Audit Commission. *The national duplicate registration initiative 2009/10 National report*. London.
18. Office for National Statistics. *Patient Register: quality assurance of administrative data used in population statistics, Dec 2016 - Office for National Statistics*. London: GOV.UK; 2016.
19. Pearce A, Law C, Elliman D, Cole TJ, Bedford H. *Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study*. BMJ. 2008;336(7647):754-7.
20. Yarwood J, Noakes K, Kennedy D, Campbell H, Salisbury D. *Tracking mothers attitudes to childhood immunisation 1991-2001*. Vaccine. 2005;23(48-49):5670-87.
21. TNS/BRMB. *Childhood Immunisation Tracking* 2010. Available from: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_126438.pdf.

22. Public Health England. *Vaccine update: issue 240* 2016. Available from: https://www.gov.uk/government/publications/vaccine-update-issue-240-january-2016.
23. Smith A, Yarwood J, Salisbury DM. *Tracking mothers' attitudes to MMR immunisation 1996-2006*. Vaccine. 2007;25(20):3996-4002.
24. Andrews N, Tischer A, Siedler A, Pebody RG, Barbara C, Cotter S, et al. *Towards elimination: measles susceptibility in Australia and 17 European countries*. Bulletin of the World Health Organization. 2008;86(3):197-204.
25. The Health Protection (Notification) Regulations 2010, 2010 No. 659. London: Her Majesty's Stationery Office; 2010.
26. Public Health England. *Notifications of infectious diseases (NOIDs)* London; 2014 [Date. Available from: https://www.gov.uk/government/collections/notifications-of-infectious-diseases-noids.
27. Medicines and Healthcare products Regulatory Agency. *Yellow Card Scheme - MHRA* London; 2017 [Date Accessed February 2017]. Available from: https://yellowcard.mhra.gov.uk/.
28. Public Health England. *Measles: the green book, chapter 21*. In: Salisbury D, Ramsay M, editors. Immunisation against infectious disease: the green book. London: GOV.UK; 2013.
29. Donaldson L, Beasley C, Ridge K. *The MMR catch-up programme*. London: Department of Health; 2008. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_086817.pdf.
30. Public Health England. *National MMR vaccination catch-up programme announced in response to increase in measles cases*: GOV.UK; 2013. Available from: https://www.gov.uk/government/news/national-mmr-vaccination-catch-up-programme-announced-in-response-to-increase-in-measles-cases.
31. B. B. C. News. *New research fuels MMR debate*: BBC; 3 February 2002. Available from: http://news.bbc.co.uk/1/hi/health/1798290.stm.
32. Triggle N. *MMR doctor struck from register*: BBC; 24 May 2010. Available from: http://news.bbc.co.uk/1/hi/health/8695267.stm.

33. World Health Organization Regional Office for Europe. *Operational targets for EPI diseases*. Copenhagen : WHO Regional Office for Europe; 1996.
34. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*: Oxford University Press; 1991.
35. NHS Information Centre. *NHS Immunisation Statistics England 2008-09*; 2009 [Date Accessed July 2011]. Available from: http://www.ic.nhs.uk/webfiles/publications/003_Health_Lifestyles/immstats2008-2009/Final_Imms_Bulletin_2008-09_Updated27Oct10.pdf.
36. NHS Information Centre. *NHS Immunisation Statistics England 2009-10*; 2010 [Date Accessed July 2011]. Available from: http://www.ic.nhs.uk/webfiles/publications/003_Health_Lifestyles/immstats0910/Immunisations_Bulletin_2009-10.pdf.
37. NHS Information Centre. *NHS Immunisation Statistics England 2010-11*; 2011 [Date Accessed September 2011]. Available from: http://www.ic.nhs.uk/webfiles/publications/003_Health_Lifestyles/Immunisation%20Stats%202010-11/Immunisations_Bulletin_2010_11.pdf.
38. NHS Information Centre. *NHS Immunisation Statistics, England - 2011-12* London; 2012 [Date Accessed October 2012]. Available from: http://content.digital.nhs.uk/catalogue/PUB09125.
39. NHS Information Centre. *NHS Immunisation Statistics, England - 2012-13* London; 2013 [Date Accessed October 2013]. Available from: http://content.digital.nhs.uk/catalogue/PUB11665.
40. NHS Information Centre. *NHS Immunisation Statistics, England - 2013-14* London; 2014 [Date Accessed October 2014]. Available from: http://content.digital.nhs.uk/catalogue/PUB14949.
41. NHS Information Centre. *NHS Immunisation Statistics, England - 2014-15* London; 2015 [Date Accessed October 2015]. Available from: http://content.digital.nhs.uk/catalogue/PUB18472.
42. Screening and Immunisations Team ND. *NHS Immunisation Statistics, England - 2015-16* London; 2016 [Date Accessed October 2016]. Available from: http://content.digital.nhs.uk/catalogue/PUB21651.
43. NHS England. *Child Immunisation 2015/16 GP*; 2016 [Date Accessed February 2017]. Available from: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2014/03/1516-Child-Immunisation-by-GP-4.xlsx.

44. McAteer S, Oultram J, J. O. *Routine Childhood Vaccination and Immunisation Programme: Health Equity Audit.* . Halton: NHS Halton & St Helens; 2011.
45. Cowen J. MHRA. Personal communication. 2015.
46. European Centre for Disease Prevention and Control. *Measles once again endemic in the United Kingdom*. Euro Surveill. 2008;13(27).
47. Jansen VA, Stollenwerk N, Jensen HJ, Ramsay ME, Edmunds WJ, Rhodes CJ. *Measles outbreaks in a population with declining vaccine uptake*. Science. 2003;301(5634):804.
48. Health Protection Agency. *Confirmed cases of measles by region and age: 1996-2013*; 2013 [Date Accessed February 2017]. Available from: http://webarchive.nationalarchives.gov.uk/20140505192923/http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1195733778332.
49. Public Health England. *Confirmed cases of measles in England and Wales by region and age: 2012 to 2016* London; 2017 [Date Accessed April 2017]. Available from: https://www.gov.uk/government/publications/measles-confirmed-cases/confirmed-cases-of-measles-in-england-and-wales-by-region-and-age-2012-to-2014.
50. Health Protection Agency. *Measles Notifications England and Wales, by Region, 1989-2012*; 2013 [Date Accessed February 2017]. Available from: http://webarchive.nationalarchives.gov.uk/20140505192923/http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1195733756177.
51. Public Health England. *Measles notifications in England and Wales, by region*; 2015 [Date Accessed February 2017]. Available from: https://www.gov.uk/government/publications/measles-notifications-by-age-group-region-and-sex/measles-notifications-in-england-and-wales-by-region.
52. Office for National Statistics. *UK and regional population estimates 1838 to 2015* 2016. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland.
53. Public Health Wales. *Outbreak of measles in Wales Nov 2012 – July 2013. Report of the agencies which responded to the outbreak*. Cardiff: Public Health Wales NHS Trust; 2013. Available from: http://www.wales.nhs.uk/sitesplus/888/page/66389.

54. Palit N. *Swansea measles epidemic: Confirmed cases jump to 693*. BBC News [Internet]. 2013. Available from: http://www.bbc.co.uk/news/av/uk-wales-22111134/swansea-measles-cases-rise-to-nearly-700.

55. Hodgekiss A. *20 cases of measles a day are now being diagnosed in Swansea epidemic*: Associated Newspapers Ltd; 5 April 2013. Available from: http://www.dailymail.co.uk/health/article-2304578/20-cases-measles-day-diagnosed-Swansea-epidemic.html.

56. Mason BW, Donnelly PD. *Impact of a local newspaper campaign on the uptake of the measles mumps and rubella vaccine*. J Epidemiol Community Health. 2000;54(6):473-4.

57. Public Health England. *MMR Action Plan*. London; 2013. PHE gateway number: 2013067.

58. Simone B, Balasegaram S, Gobin M, Anderson C, Charlett A, Coole L, et al. *Evaluation of the measles, mumps and rubella vaccination catch-up campaign in England in 2013*. Vaccine. 2014;32(36):4681-8.

59. Public Health England. *Evaluation of vaccine uptake during the 2013 MMR catch-up campaign in England Report for the national measles oversight Group* London2014. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/285890/Evaluation_of_the_2013_MMR_catch-up_campaign_in_England.pdf.

60. Ramsay ME, Yarwood J, Lewis D, Campbell H, White JM. *Parental confidence in measles, mumps and rubella vaccine: evidence from vaccine coverage and attitudinal surveys*. Br J Gen Pract. 2002;52(484):912-6.

61. Mixer RE, Jamrozik K, Newsom D. *Ethnicity as a correlate of the uptake of the first dose of mumps, measles and rubella vaccine*. J Epidemiol Community Health. 2007;61(9):797-801.

62. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, et al. *Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children*. Lancet. 1998;351(9103):637-41.

63. Begg N, Ramsay M, White J, Bozoky Z. *Media dents confidence in MMR vaccine*. BMJ. 1998;316(7130):561.

64. Lewis J, Speers T. *Misleading media reporting? The MMR story*. Nat Rev Immunol. 2003;3(11):913-8.

65. Guillaume L, Bath PA. *A content analysis of mass media sources in relation to the MMR vaccine scare*. Health Informatics J. 2008;14(4):323-34.

66. Demicheli V, Rivetti A, Debalini MG, Di Pietrantonj C. *Vaccines for measles, mumps and rubella in children*. Cochrane Database Syst Rev. 2012(2):Cd004407.

67. The Editors of the Lancet. *Retraction--Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children*. Lancet. 375. England2010. p. 445.

68. Middleton E, Baker D. *Comparison of social distribution of immunisation with measles, mumps, and rubella vaccine, England, 1991-2001*. BMJ. 2003;326(7394):854.

69. Wright JA, Polack C. *Understanding variation in measles-mumps-rubella immunization coverage--a population-based study*. Eur J Public Health. 2006;16(2):137-42.

70. Anderberg D, Chevalier A, Wadsworth J. *Anatomy of a health scare: education, income and the MMR controversy in the UK*. J Health Econ. 2011;30(3):515-30.

71. McIntyre P, Leask J. *Improving uptake of MMR vaccine*. BMJ. 2008;336(7647):729-30.

72. Burgess DC, Burgess MA, Leask J. *The MMR vaccination and autism controversy in United Kingdom 1998-2005: inevitable community outrage or a failure of risk communication?* Vaccine. 2006;24(18):3921-8.

73. Stefanoff P, Mamelund SE, Robinson M, Netterlid E, Tuells J, Bergsaker MA, et al. *Tracking parental attitudes on vaccination across European countries: The Vaccine Safety, Attitudes, Training and Communication Project (VACSATC)*. Vaccine. 2010;28(35):5731-7.

74. Hanratty B, Holt T, Duffell E, Patterson W, Ramsay M, White JM, et al. *UK measles outbreak in non-immune anthroposophic communities: the implications for the elimination of measles from Europe*. Epidemiol Infect. 2000;125(2):377-83.

75. Ruijs WL, Hautvast JL, van Ijzendoorn G, van Ansem WJ, van der Velden K, Hulscher ME. *How orthodox protestant parents decide on the vaccination of their children: a qualitative study*. BMC Public Health. 2012;12:408.

76. Henderson L, Millett C, Thorogood N. *Perceptions of childhood immunization in a minority community: qualitative study*. J R Soc Med. 2008;101(5):244-51.
77. Cohuet S, Bukasa A, Heathcock R, White J, Brown K, Ramsay M, et al. *A measles outbreak in the Irish traveller ethnic group after attending a funeral in England, March-June 2007*. Epidemiol Infect. 2009;137(12):1759-65.
78. Stewart-Freedman B, Kovalsky N. *An ongoing outbreak of measles linked to the United Kingdom in an ultra-orthodox Jewish community in Israel*. Euro Surveill. 2007;12(9):E070920.1.
79. Centers for Disease Control and Prevention. *Notes from the field: measles outbreak among members of a religious community - Brooklyn, New York, March-June 2013*. MMWR Morb Mortal Wkly Rep. 2013;62(36):752-3.
80. Pareek M, Pattison HM. *The two-dose measles, mumps, and rubella (MMR) immunisation schedule: factors affecting maternal intention to vaccinate*. Br J Gen Pract. 2000;50(461):969-71.
81. Petrovic M, Roberts RJ, Ramsay M, Charlett A. *Parents' attitude towards the second dose of measles, mumps and rubella vaccine: a case-control study*. Commun Dis Public Health. 2003;6(4):325-9.
82. Smailbegovic MS, Laing GJ, Bedford H. *Why do parents decide against immunization? The effect of health beliefs and health professionals*. Child Care Health Dev. 2003;29(4):303-11.
83. Flynn M, Ogden J. *Predicting uptake of MMR vaccination: a prospective questionnaire study*. Br J Gen Pract. 2004;54(504):526-30.
84. Gellatly J, McVittie C, Tiliopoulos N. *Predicting parents' decisions on MMR immunisation: a mixed method investigation*. Fam Pract. 2005;22(6):658-62.
85. MacDonald M. *Parents' decisions on MMR vaccination for their children were based on personal experience rather than scientific evidence*. Evid Based Nurs. 2005;8(2):60.
86. Wroe AL, Bhan A, Salkovskis P, Bedford H. *Feeling bad about immunising our children*. Vaccine. 2005;23(12):1428-33.
87. Casiday RE. *Children's health and the social theory of risk: insights from the British measles, mumps and rubella (MMR) controversy*. Soc Sci Med. 2007;65(5):1059-70.
88. Casiday R. *Uncertainty, decision-making and trust: lessons from the MMR controversy*. Community Pract. 2006;79(11):354-7.

89. Cassell JA, Leach M, Poltorak MS, Mercer CH, Iversen A, Fairhead JR. *Is the cultural context of MMR rejection a key to an effective public health discourse?* Public Health. 2006;120(9):783-94.

90. Brown KF, Shanley R, Cowley NA, van Wijgerden J, Toff P, Falconer M, et al. *Attitudinal and demographic predictors of measles, mumps and rubella (MMR) vaccine acceptance: development and validation of an evidence-based measurement instrument*. Vaccine. 2011;29(8):1700-9.

91. Evans M, Stoddart H, Condon L, Freeman E, Grizzell M, Mullen R. *Parents' perspectives on the MMR immunisation: a focus group study*. Br J Gen Pract. 2001;51(472):904-10.

92. McMurray R, Cheater FM, Weighall A, Nelson C, Schweiger M, Mukherjee S. *Managing controversy through consultation: a qualitative study of communication and trust around MMR vaccination decisions*. Br J Gen Pract. 2004;54(504):520-5.

93. Casiday R, Cresswell T, Wilson D, Panter-Brick C. *A survey of UK parental attitudes to the MMR vaccine and trust in medical authority*. Vaccine. 2006;24(2):177-84.

94. Poltorak M, Leach M, Fairhead J, Cassell J. *'MMR talk' and vaccination choices: an ethnographic study in Brighton*. Soc Sci Med. 2005;61(3):709-19.

95. Hilton S, Petticrew M, Hunt K. *'Combined vaccines are like a sudden onslaught to the body's immune system': parental concerns about vaccine 'overload' and 'immune-vulnerability'*. Vaccine. 2006;24(20):4321-7.

96. Skea ZC, Entwistle VA, Watt I, Russell E. *'Avoiding harm to others' considerations in relation to parental measles, mumps and rubella (MMR) vaccination discussions - an analysis of an online chat forum*. Soc Sci Med. 2008;67(9):1382-90.

97. Brown KF, Long SJ, Ramsay M, Hudson MJ, Green J, Vincent CA, et al. *U.K. parents' decision-making about measles-mumps-rubella (MMR) vaccine 10 years after the MMR-autism controversy: a qualitative analysis*. Vaccine. 2012;30(10):1855-64.

98. Lamden KH, Gemmell I. *General practice factors and MMR vaccine uptake: structure, process and demography*. J Public Health (Oxf). 2008;30(3):251-7.

99. Hawker JI, Olowokure B, Wood AL, Wilson RC, Johnson R. *Widening inequalities in MMR vaccine uptake rates among ethnic groups in an urban*

area of the UK during a period of vaccine controversy (1994-2000). Vaccine. 2007;25(43):7516-9.

100. Hilton S, Petticrew M, Hunt K. *Parents' champions vs. vested interests: who do parents believe about MMR? A qualitative study*. BMC Public Health. 2007;7:42.
101. Petts J, Niemeyer S. *Health risk communication and amplification: learning from the MMR vaccination controversy*. Health, Risk & Society. 2004;6(1):7-23.
102. Gardner B, Davies A, McAteer J, Michie S. *Beliefs underlying UK parents' views towards MMR promotion interventions: a qualitative study*. Psychol Health Med. 2010;15(2):220-30.
103. Tickner S, Leman PJ, Woodcock A. *Parents' views about pre-school immunization: an interview study in southern England*. Child Care Health Dev. 2010;36(2):190-7.
104. Samad L, Tate AR, Dezateux C, Peckham C, Butler N, Bedford H. *Differences in risk factors for partial and no immunisation in the first year of life: prospective cohort study*. BMJ. 2006;332(7553):1312-3.
105. Friederichs V, Cameron JC, Robertson C. *Impact of adverse publicity on MMR vaccine uptake: a population based analysis of vaccine uptake records for one million children, born 1987-2004*. Arch Dis Child. 2006;91(6):465-8.
106. May T, Silverman RD. *'Clustering of exemptions' as a collective action threat to herd immunity*. Vaccine. 2003;21(11-12):1048-51.
107. Richard JL, Masserey Spicher V. *Large measles epidemic in Switzerland from 2006 to 2009: consequences for the elimination of measles in Europe*. Euro Surveill. 2009;14(50).
108. Sugerman DE, Barskey AE, Delea MG, Ortega-Sanchez IR, Bi D, Ralston KJ, et al. *Measles outbreak in a highly vaccinated population, San Diego, 2008: role of the intentionally undervaccinated*. Pediatrics. 2010;125(4):747-55.
109. Parent du Chatelet I, Floret D, Antona D, Levy-Bruhl D. *Measles resurgence in France in 2008, a preliminary report*. Euro Surveill. 2009;14(6).
110. Muscat M, Vinner L, Christiansen AH, Glismann S, Bottiger BE. *The benefit of molecular characterization during a measles upsurge in Denmark*. Vaccine. 2007;25(33):6232-6.

111. Antona D, Levy-Bruhl D, Baudon C, Freymuth F, Lamy M, Maine C, et al. *Measles elimination efforts and 2008-2011 outbreak, France*. Emerg Infect Dis. 2013;19(3):357-64.
112. Wallinga J, Heijne JC, Kretzschmar M. *A measles epidemic threshold in a highly vaccinated population*. PLoS Med. 2005;2(11):e316.
113. Fine PE, Clarkson JA. *Measles in England and Wales--I: An analysis of factors underlying seasonal patterns*. Int J Epidemiol. 1982;11(1):5-14.
114. Olsen LF, Schaffer WM. *Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics*. Science. 1990;249(4968):499-504.
115. Schenzle D. *An age-structured model of pre- and post-vaccination measles transmission*. IMA J Math Appl Med Biol. 1984;1(2):169-91.
116. Keeling MJ, Eames KT. *Networks and epidemic models*. J R Soc Interface. 2005;2(4):295-307.
117. Smith KP, Christakis NA. *Social Networks and Health*. Annu Rev Sociol. 2008;34:405-29.
118. Volz EM, Miller JC, Galvani A, Ancel Meyers L. *Effects of heterogeneous and clustered contact patterns on infectious disease dynamics*. PLoS Comput Biol. 2011;7(6):e1002042.
119. Centola D, Macy M. *Complex contagions and the weakness of long ties 1*. American Journal of Sociology. 2007;113(3):702-34.
120. Centola D. *The spread of behavior in an online social network experiment*. Science. 2010;329(5996):1194-7.
121. Von Neumann J, Morgenstern O. *Theory of games and economic behavior*: Princeton university press; 2007.
122. Baron J. *Thinking and deciding*. 2nd ed. Cambridge: Cambridge University Press; 1994.
123. Rosenstock IM. *Why people use health services*. Milbank Mem Fund Q. 1966;44(3):Suppl:94-127.
124. Fishbein M, Ajzen I. *Belief, attitude, intention, and behavior: An introduction to theory and research*. 1977.
125. Ajzen I. *Attitudes, personality, and behavior*: McGraw-Hill Education (UK); 2005.
126. MacDonald NE. *Vaccine hesitancy: Definition, scope and determinants*. Vaccine. 2015;33(34):4161-4.

127. Tversky A, Kahneman D. *Judgment under Uncertainty: Heuristics and Biases*. Science. 1974;185(4157):1124-31.
128. Simon HA. *The behavioral and social sciences*. Science. 1980;209(4452):72-8.
129. Nickerson RS. *Confirmation bias: A ubiquitous phenomenon in many guises*. Review of General Psychology. 1998;2(2):175.
130. Douglas M, Wildavsky A. *Risk and culture: An essay on the selection of technological and environmental dangers*: Univ of California Press; 1983.
131. Song G. *Understanding Public Perceptions of Benefits and Risks of Childhood Vaccinations in the United States*. Risk Anal. 2014;34(3):541-55.
132. Fischhoff B, Slovic P, Lichtenstein S, Read S, Combs B. *How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits*. Policy Sciences. 1978;9(2):127-52.
133. Banerjee AV. *A simple model of herd behavior*. The Quarterly Journal of Economics. 1992;107(3):797-817.
134. Plous S. *Psychology of Judgement and Decision Making*. New York: McGraw-Hill; 1993.
135. Cross R, Borgatti SP, Parker A. *Beyond answers: dimensions of the advice network*. Social Networks. 2001;23(3):215-35.
136. Creswick N, Westbrook JI, Braithwaite J. *Understanding communication networks in the emergency department*. BMC Health Serv Res. 2009;9:247.
137. Christakis NA, Fowler JH. *The spread of obesity in a large social network over 32 years*. N Engl J Med. 2007;357(4):370-9.
138. Christakis NA, Fowler JH. *The collective dynamics of smoking in a large social network*. N Engl J Med. 2008;358(21):2249-58.
139. Hill AL, Rand DG, Nowak MA, Christakis NA. *Infectious disease modeling of social contagion in networks*. PLoS Comput Biol. 2010;6(11):e1000968.
140. Scherer CW, Cho H. *A social network contagion theory of risk perception*. Risk Anal. 2003;23(2):261-7.
141. Bauch CT, Galvani AP, Earn DJ. *Group interest versus self-interest in smallpox vaccination policy*. Proc Natl Acad Sci U S A. 2003;100(18):10564-7.
142. Bauch CT, Earn DJ. *Vaccination and the theory of games*. Proc Natl Acad Sci U S A. 2004;101(36):13391-4.

143. Bauch CT. *Imitation dynamics predict vaccinating behaviour*. Proc Biol Sci. 2005;272(1573):1669-75.
144. Bhattacharyya S, Bauch CT. *"Wait and see" vaccinating behaviour during a pandemic: a game theoretic analysis*. Vaccine. 2011;29(33):5519-25.
145. Reluga TC, Bauch CT, Galvani AP. *Evolving public perceptions and stability in vaccine uptake*. Math Biosci. 2006;204(2):185-98.
146. Voinson M, Billiard S, Alvergne A. *Beyond Rational Decision-Making: Modelling the Influence of Cognitive Biases on the Dynamics of Vaccination Coverage*. PLoS One. 2015;10(11):e0142990.
147. Oraby T, Thampi V, Bauch CT. *The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases*. Proc Biol Sci. 2014;281(1780):20133172.
148. Del Valle S, Hethcote H, Hyman JM, Castillo-Chavez C. *Effects of behavioral changes in a smallpox attack model*. Math Biosci. 2005;195(2):228-51.
149. Fu F, Rosenbloom DI, Wang L, Nowak MA. *Imitation dynamics of vaccination behaviour on social networks*. Proc Biol Sci. 2011;278(1702):42-9.
150. Codeco CT, Luz PM, Coelho F, Galvani AP, Struchiner C. *Vaccinating in disease-free regions: a vaccine model with application to yellow fever*. J R Soc Interface. 2007;4(17):1119-25.
151. Shim E, Chapman GB, Townsend JP, Galvani AP. *The influence of altruism on influenza vaccination decisions*. J R Soc Interface. 2012;9(74):2234-43.
152. Gross T, D'Lima CJ, Blasius B. *Epidemic dynamics on an adaptive network*. Phys Rev Lett. 2006;96(20):208701.
153. Epstein JM, Parker J, Cummings D, Hammond RA. *Coupled contagion dynamics of fear and disease: mathematical and computational explorations*. PLoS One. 2008;3(12):e3955.
154. Shaw LB, Schwartz IB. *Fluctuating epidemics on adaptive networks*. Phys Rev E Stat Nonlin Soft Matter Phys. 2008;77(6 Pt 2):066101.
155. Zanette DH, Risau-Gusman S. *Infection spreading in a population with evolving contacts*. J Biol Phys. 2008;34(1-2):135-48.
156. Van Segbroeck S, Santos FC, Pacheco JM. *Adaptive contact networks change effective disease infectiousness and dynamics*. PLoS Comput Biol. 2010;6(8).

157. Funk S, Gilad E, Watkins C, Jansen VA. *The spread of awareness and its impact on epidemic outbreaks*. Proc Natl Acad Sci U S A. 2009;106(16):6872-7.
158. Funk S, Gilad E, Jansen VA. *Endemic disease, awareness, and local behavioural response*. J Theor Biol. 2010;264(2):501-9.
159. Kiss IZ, Cassell J, Recker M, Simon PL. *The impact of information transmission on epidemic outbreaks*. Math Biosci. 2010;225(1):1-10.
160. Hatzopoulos V, Taylor M, Simon PL, Kiss IZ. *Multiple sources and routes of information transmission: Implications for epidemic dynamics*. Math Biosci. 2011;231(2):197-209.
161. Bagnoli F, Lio P, Sguanci L. *Risk perception in epidemic modeling*. Phys Rev E Stat Nonlin Soft Matter Phys. 2007;76(6 Pt 1):061904.
162. Perisic A, Bauch CT. *A simulation analysis to characterize the dynamics of vaccinating behaviour on contact networks*. BMC Infect Dis. 2009;9:77.
163. Perisic A, Bauch CT. *Social contact networks and disease eradicability under voluntary vaccination*. PLoS Comput Biol. 2009;5(2):e1000280.
164. Cardillo A, Reyes-Suárez C, Naranjo F, Gómez-Gardeñes J. *Evolutionary vaccination dilemma in complex networks*. Physical Review E. 2013;88(3):032803.
165. Fukuda E, Tanimoto J, Akimoto M. *Influence of breaking the symmetry between disease transmission and information propagation networks on stepwise decisions concerning vaccination*. Chaos, Solitons & Fractals. 2015;80:47-55.
166. Ndeffo Mbah ML, Liu J, Bauch CT, Tekel YI, Medlock J, Meyers LA, et al. *The impact of imitation on vaccination behavior in social contact networks*. PLoS Comput Biol. 2012;8(4):e1002469.
167. Xia S, Liu J. *A computational approach to characterizing the impact of social influence on individuals' vaccination decision making*. PLoS One. 2013;8(4):e60373.
168. Eames KT. *Networks of influence and infection: parental choices and childhood disease*. J R Soc Interface. 2009;6(38):811-4.
169. Campbell E, Salathe M. *Complex social contagion makes networks more vulnerable to disease outbreaks*. Sci Rep. 2013;3:1905.

170. Salathe M, Bonhoeffer S. *The effect of opinion clustering on disease outbreaks*. J R Soc Interface. 2008;5(29):1505-8.

171. Tickner S, Leman PJ, Woodcock A. *The Immunisation Beliefs and Intentions Measure (IBIM): predicting parents' intentions to immunise preschool children*. Vaccine. 2010;28(19):3350-62.

172. International Society for Infectious Diseases. *Parents urged to get children vaccinated after measles outbreak strikes in North Devon*. PRO/EDR> Measles update [Internet]. 2013 [January 2013]; 2013 (02). Available from: http://www.promedmail.org/.

173. International Society for Infectious Diseases. *Measles outbreak in Morecambe and Heysham*. PRO/EDR> Measles update [Internet]. 2013 [013 013]; 2013 (3). Available from: http://www.promedmail.org/ <p class="MsoNormalCxSpFirst" style="mso-margin-top-alt:auto.

174. Pegorie M, Shankar K, Welfare WS, Wilson RW, Khiroya C, Munslow G, et al. *Measles outbreak in Greater Manchester, England, October 2012 to September 2013: epidemiology and control*. Euro Surveill. 2014;19(49).

175. Vivancos R, Keenan A, Farmer S, Atkinson J, Coffey E, Dardamissis E, et al. *An ongoing large outbreak of measles in Merseyside, England, January to June 2012*. Euro Surveill. 2012;17(29).

176. Pearce A, Mindlin M, Cortina-Borja M, Bedford H. *Characteristics of 5-year-olds who catch-up with MMR: findings from the UK Millennium Cohort Study*. BMJ Open. 2013;3(7).

177. Raworth E. Great Yarmouth & Waveney PCT,. Personal communication. 2013.

178. Office of National Statistics *Neighbourhood Statistics (NeSS)* London; 2016 [Date Accessed April 2016]. Available from: http://neighbourhood.statistics.gov.uk/dissemination/.

179. Office for National Statistics. *Statistical disclosure control for 2011 Census*. Available from: https://www.ons.gov.uk/census/2011census/2011censusdata/2011censususerguide/qualityandmethods.

180. Association of Public Health Observatories. *Using small area data in public health intelligence*. York: Association of Public Health Observatories; 2009.

181. Health Protection Agency. *COVER data Q10-1 Apr to Jun 2010*; 2010 [Date Accessed July 2011]. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/.

182. Health Protection Agency. *COVER data Q10-2 Jul to Sep 2010*; 2010 [Date Accessed July 2011]. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1287146854796.

183. Health Protection Agency. *COVER data Q10-3 Oct to Dec 2010*; 2011 [Date. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1296683967715.

184. Health Protection Agency. *COVER data Q10-4 Jan to Mar 2011*; 2011 [Date Accessed July 2011]. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1296689222082.

185. Office for National Statistics. *Part 4 Derived variables*. 2011 Census Variable and Classification Information:. London: Office for National Statistics; 2014.

186. Office for National Statistics. *DC1115EW (Age of youngest dependent child by family type by marital status by age of Family Reference Person (FRP)) - Nomis - Official Labour Market Statistics*; 2013 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/dc1115ew.

187. Office for National Statistics. *LC5102EW (Highest level of qualification by age) - Nomis - Official Labour Market Statistics*; 2014 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/lc5102ew.

188. Office for National Statistics. *LC6107EW (Economic Activity by sex by age) - Nomis - Official Labour Market Statistics*; 2014 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/lc6107ew.

189. Office for National Statistics. *LC1114EW (Dependent children by family type) - Nomis - Official Labour Market Statistics*; 2014 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/lc1114ew.

190. Office for National Statistics. *KS201UK (Ethnic group) - Nomis - Official Labour Market Statistics*; 2014 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/ks201uk.

191. Office for National Statistics. *QS103EW (Age by single year) - Nomis - Official Labour Market Statistics*; 2013 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/qs103ew.

192. Office for National Statistics. *LC2105EW (Proficiency in English by age) - Nomis - Official Labour Market Statistics*; 2013 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/lc2105ew.
193. Office for National Statistics. *2011 Census* London; undated [Date. Available from: https://www.ons.gov.uk/census/2011census.
194. Office for National Statistics. *QS102EW (Population density) - Nomis - Official Labour Market Statistics*; 2013 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/qs102ew.
195. Department for Communities and Local Government. *Guidance*. 2015 30 September 2015. In: English indices of deprivation 2015 [Internet]. London: GOV.UK. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf.
196. Department for Communities and Local Government, editor. *English Indices of Deprivation 2010 Guidance document*. London: GOV.UK; 2011.
197. Health and Social Care Information Centre. *Quality and Outcomes Framework (QOF)*. NHS England; 2012.
198. Health and Social Care Information Centre. *QOF data Child Health Surveillance (CH501)*. Quality and Outcomes Framework - 2011-12, Practice level. 2012.
199. Health and Social Care Information Centre. *QOF data Patient Experience (PE01)*. 2012.
200. Department for Communities and Local Government. *English indices of deprivation 2015: research report*. London: Department for Communities and Local Government; 2015. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/464597/English_Indices_of_Deprivation_2015_-_Research_Report.pdf.
201. Open Geography Portal. *Output areas (2011) to wards (2011) to local authority districts (2011) E+W lookup* London; [Date Accessed April 2016]. Available from: https://geoportal.statistics.gov.uk/geoportal/catalog/main/home.page }.
202. Open Geography Portal. *Output areas (2011) to lower layer super output areas (2011) to middle layer super output areas (2011) to local authority*

districts (2011) E+W lookup; [Date Accessed April 2016]. Available from: https://geoportal.statistics.gov.uk/geoportal/catalog/main/home.page.

203. Department for Communities and Local Government. *File 6: population denominators*. 2015. In: English indices of deprivation [Internet]. London: GOV.UK. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/467773/File_6_ID_2015_Population_denominators.xlsx.
204. IBM Corp. *IBMSPSS Statistics for Windows, Version 22.0.* IBM Corp, Armonk, NY. 2013.
205. Evans JD. *Straightforward statistics for the behavioral sciences*: Brooks/Cole; 1996.
206. Aitchison J. *Principal component analysis of compositional data*. Biometrika. 1983:57-65.
207. Filzmoser P, Hron K, Reimann C. *Principal component analysis for compositional data with outliers*. Environmetrics. 2009;20(6):621-32.
208. Templ M, Hron K, Filzmoser P. *robCompositions: an R-package for robust statistical analysis of compositional data.* 2011.
209. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. 2016. 3.2.4 Revised. http://www.R-project.org/
210. Linting M, Meulman JJ, Groenen PJF, van der Koojj AJ. *Nonlinear principal components analysis: introduction and application*. Psychological Methods. 2007;12(3):336.
211. Office for National Statistics. *LC1117EW (Sex by age) - Nomis - Official Labour Market Statistics*; 2015 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/lc1117ew.
212. Marascuilo LA, McSweeney M. *Nonparametric post hoc comparisons for trend*. Psychological Bulletin. 1967;67(6):401.
213. Office for National Statistics. *Births by Parents' Characteristics in England and Wales: 2014*. London: Office for National Statistics; 2015.
214. Grossman M. *On the concept of health capital and the demand for health*. Journal of Political Economy. 1972;80(2):223-55.
215. Tabacchi G, Costantino C, Napoli G, Marchese V, Cracchiolo M, Casuccio A, et al. *Determinants of European parents' decision on the vaccination of their*

children against measles, mumps and rubella: A systematic review and meta-analysis. Hum Vaccin Immunother. 2016;12(7):1909-23.

216. Office for National Statistics. *Graduates in the UK Labour Market 2013*. London: Labour Force Survey; 2013. Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/graduatesintheuklabourmarket/2013-11-19.
217. Jakab Z, Salisbury DM. *Back to basics: the miracle and tragedy of measles vaccine*. Lancet. 2013;381(9876):1433-4.
218. Lyratzopoulos G, Aston R, Bailey K, Flitcroft J, Clarke H. *Accuracy of routine data on MMR vaccination coverage and validity of parental recall of vaccination*. Commun Dis Public Health. 2002;5(4):305-10.
219. Department for Communities and Local Government. *English indices of deprivation 2015: technical report*. London: GOV.UK; 2015. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/464485/English_Indices_of_Deprivation_2015_-_Technical-Report.pdf
220. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. *Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012*. Vaccine. 2014;32(19):2150-9.
221. European Centre for Disease Prevention and Control. *European monthly measles monitoring (EMMO) February 2012*. Stockholm: ECDC; 2012.
222. European Centre for Disease Prevention and Control. *Measles and rubella monitoring February 2013*. Stockholm: ECDC; 2013.
223. European Centre for Disease Prevention and Control. *European Centre for Disease Prevention and Control. Measles and rubella monitoring, February 2014 – Reporting on January–December 2013 surveillance data and epidemic intelligence data to the end of February 2014*. Stockholm; 2014.
224. European Centre for Disease Prevention and Control. *Measles and rubella monitoring January 2015 Reporting on January – December 2014 surveillance data*. 2015.
225. d'Onofrio A, Manfredi P, Poletti P. *The impact of vaccine side effects on the natural history of immunization programmes: an imitation-game approach*. J Theor Biol. 2011;273(1):63-71.

226. Zhang H, Fu F, Zhang W, Wang B. *Rational behavior is a 'double-edged sword'when considering voluntary vaccination*. Physica A: Statistical Mechanics and its Applications. 2012;391(20):4807-15.
227. Prettejohn BJ, Berryman MJ, McDonnell MD. *Methods for generating complex networks with selected structural properties for simulations: a review and tutorial for neuroscientists*. Front Comput Neurosci. 2011;5:11.
228. Erdös P, Rényi A. *On random graphs, I*. Publicationes Mathematicae (Debrecen). 1959;6:290-7.
229. Watts DJ, Strogatz SH. *Collective dynamics of 'small-world' networks*. Nature. 1998;393(6684):440-2.
230. Barabasi AL, Albert R. *Emergence of scaling in random networks*. Science. 1999;286(5439):509-12.
231. Marsden PV. *Core discussion networks of Americans*. American Sociological Review. 1987:122-31.
232. Hanneman R, Riddle M. *Introduction to social network methods*. Riverside, Ca: University of California Riverside; 2005. Available from: http://faculty.ucr.edu/~hanneman/.
233. Keeling MJ, Rohani P. *Modeling infectious diseases in humans and animals*: Princeton University Press; 2008.
234. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. *Social contacts and mixing patterns relevant to the spread of infectious diseases*. PLoS Med. 2008;5(3):e74.
235. Danon L, House TA, Read JM, Keeling MJ. *Social encounter networks: collective properties and disease transmission*. J R Soc Interface. 2012;9(76):2826-33.
236. TNS/BMRB. *COI/DH: Childhood Immunisation Wave 33 (England) report*. London; 2010.
237. Bond L, Nolan T. *Making sense of perceptions of risk of diseases and vaccinations: a qualitative study combining models of health beliefs, decision-making and risk perception*. BMC Public Health. 2011;11:943.
238. Brown KF, Kroll JS, Hudson MJ, Ramsay M, Green J, Vincent CA, et al. *Omission bias and vaccine rejection by parents of healthy children: implications for the influenza A/H1N1 vaccination programme*. Vaccine. 2010;28(25):4181-5.

239. Nyhan B, Reifler J. *Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information*. Vaccine. 2015;33(3):459-64.
240. Bhattacharyya S, Bauch CT, Breban R. *Role of word-of-mouth for programs of voluntary vaccination: A game-theoretic approach*. Math Biosci. 2015;269:130-4.
241. McPherson M, Smith-Lovin L, Cook JM. *Birds of a feather: Homophily in social networks*. Annual Review of Sociology. 2001;27(1):415-44.
242. McPherson M, Smith-Lovin L, Brashears ME. *Social isolation in America: Changes in core discussion networks over two decades*. American Sociological Review. 2006;71(3):353-75.
243. Health Protection Agency. *HPA National Measles Guidelines Local & Regional Services*. Health Protection Agency; 2010.
244. Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DA. *Close encounters of the infectious kind: methods to measure social mixing behaviour*. Epidemiol Infect. 2012;140(12):2117-30.
245. Salathe M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. *A high-resolution human contact network for infectious disease transmission*. Proc Natl Acad Sci U S A. 2010;107(51):22020-5.
246. Smieszek T, Barclay VC, Seeni I, Rainey JJ, Gao H, Uzicanin A, et al. *How should social mixing be measured: comparing web-based survey and sensor-based methods*. BMC Infect Dis. 2014;14:136.
247. Fournet J, Barrat A. *Contact patterns among high school students*. PLoS One. 2014;9(9):e107878.
248. Stehlé J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton J-F, et al. *High-resolution measurements of face-to-face contact patterns in a primary school*. PLoS One. 2011;6(8):e23176.
249. Conlan AJ, Eames KT, Gage JA, von Kirchbach JC, Ross JV, Saenz RA, et al. *Measuring social networks in British primary schools through scientific engagement*. Proc Biol Sci. 2011;278(1711):1467-75.
250. Mastrandrea R, Fournet J, Barrat A. *Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys*. PLoS One. 2015;10(9):e0136497.

251. Glass LM, Glass RJ. *Social contact networks for the spread of pandemic influenza in children and teenagers*. BMC Public Health. 2008;8:61.
252. Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. *Life in the network: the coming age of computational social science*. Science. 2009;323(5915):721.
253. Eagle N, Pentland AS, Lazer D. *Inferring friendship network structure by using mobile phone data*. Proceedings of the National Academy of Sciences. 2009;106(36):15274-8.
254. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. *Detecting influenza epidemics using search engine query data*. Nature. 2009;457(7232):1012-4.
255. Salathe M, Khandelwal S. *Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control*. PLoS Comput Biol. 2011;7(10):e1002199.
256. Brunson EK. *How parents make decisions about their children's vaccinations*. Vaccine. 2013;31(46):5466-70.
257. Allan N, Harden J. *Parental decision-making in uptake of the MMR vaccination: a systematic review of qualitative literature*. J Public Health (Oxf). 2015;37(4):678-87.
258. Swanson V, Power KG. *Initiation and continuation of breastfeeding: theory of planned behaviour*. J Adv Nurs. 2005;50(3):272-82.
259. Willis GB, Lessler JT. *Question appraisal system QAS-99*1999.
260. Office for National Statistics. *Families and households in the UK, 2001 to 2010*. Newport Wales: Office for National Statistics; 2011.
261. Suarez L, Simpson DM, Smith DR. *Errors and correlates in parental recall of child immunizations: effects on vaccination coverage estimates*. Pediatrics. 1997;99(5):E3.
262. Eames KT, Tilston NL, Edmunds WJ. *The impact of school holidays on the social mixing patterns of school children*. Epidemics. 2011;3(2):103-8.
263. OECD Family Database. *Enrolment in childcare and pre-school*. Organisation for Economic Co-operation and Development; 2016. PF3.2.
264. Brant R. *Web-based Sample Size/Power Calculations* Calgary; [Date Accessed April 2011]. Available from: https://www.stat.ubc.ca/~rollin/stats/ssize/index.html.

265. Public Health England. *Statutory Notifications of Infectious Diseases in England and Wales Week 2013/21 week ending 26/05/2013*. London; 2013.
266. Department for Education. *Childcare and early years providers survey: 2013*. London; 2014.
267. Office for National Statistics. *LC5104EW (Highest level of qualification by number of children) - Nomis - Official Labour Market Statistics*. 2014(April 2016).
268. Public Health England. *Postcode Lookup Widget*; 2014 [Date Accessed November 2014]. Available from: https://www.ndtms.org.uk/emids/cgi-bin/ons_locale.cgi.
269. Hong Y. *poibin: The Poisson Binomial Distribution.* 2013. R package version 1.2. http://CRAN.R-project.org/package=poibin
270. R. Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. 2012. 2.15.2. http://www.R-project.org/
271. Brunson EK. *The impact of social networks on parents' vaccination decisions*. Pediatrics. 2013;131(5):e1397-404.
272. Prevention CfDCa. *SchoolVaxView School Vaccination Requirements and Exemptions* Atlanta, GA; 2017 [Date Accessed April 2017]. Available from: https://www2a.cdc.gov/nip/schoolsurv/schImmRqmt.asp.
273. Brewer D.D. *Forgetting in the recall-based elicitation of personal and social networks*. Social Networks. 2000;22(1):29-43
274. Tourangeau R, Yan T. *Sensitive questions in surveys*. Psychological Bulletin. 2007;133(5):859.
275. Kreuter F, Presser S, Tourangeau R. *Social desirability bias in CATI, IVR, and Web surveys the effects of mode and question sensitivity*. Public Opinion Quarterly. 2008;72(5):847-65
276. Paulhus DL, Braun HI, Jackson DN, Wiley DE. *Socially desirable responding: The evolution of a construct*. The Role of Constructs in Psychological and Educational Measurement. 2002;49459.
277. Shih T-H., Fan X. *Comparing response rates from web and mail surveys: A meta-analysis*. Field Methods. 2008;20(3):249-71
278. Mather M, Shafir E, Johnson MK. *Misremembrance of options past: Source monitoring and choice*. Psychological Science. 2000;11(2):132-8.

279. Brotherton R, French CC, Pickering AD. *Measuring belief in conspiracy theories: The generic conspiracist beliefs scale*. Frontiers in Psychology. 2013;4:279.
280. Health Protection Agency. *Health Protection Report*. London; 2013. 7 (6).
281. Woudenberg T, van Binnendijk RS, Sanders EA, Wallinga J, de Melker HE, Ruijs WL, et al. *Large measles epidemic in the Netherlands, May 2013 to March 2014: changing epidemiology*. Euro Surveill. 2017;22(3).
282. Barrat A, Weigt M. *On the properties of small-world network models*. The European Physical Journal B-Condensed Matter and Complex Systems. 2000;13(3):547-60.
283. Dorogovtsev SN, Mendes JFF. *Exactly solvable small-world network*. EPL (Europhysics Letters). 2000;50(1):1.
284. Bolker B, R. Core Team. *bbmle: Tools for General Maximum Likelihood Estimation.* 2016. R package version 1.0.16. url https://CRAN.R-project.org/package=bbmle
285. Gillespie CS. *Fitting Heavy Tailed Distributions: The poweRlaw Package*. 2015;64(2):16.
286. Clauset A, Shalizi CR, Newman ME. *Power-law distributions in empirical data*. SIAM review. 2009;51(4):661-703.
287. Office for National Statistics. *DC1118EW (Youngest dependent child in family by family type) - Nomis - Official Labour Market Statistics*; 2014 [Date Accessed April 2016]. Available from: https://www.nomisweb.co.uk/census/2011/dc1118ew.
288. Chung F, Lu L. *Connected components in random graphs with given expected degree sequences*. Annals of Combinatorics. 2002;6(2):125-45.
289. Miller JC, Hagberg A, editors. *Efficient generation of networks with given expected degrees*. International Workshop on Algorithms and Models for the Web-Graph; 2011: Springer.
290. Knuth DE. *The art of computer programming: sorting and searching*: Pearson Education; 1998.
291. Onnela JP, Landon BE, Kahn AL, Ahmed D, Verma H, O'Malley AJ, et al. *Polio vaccine hesitancy in the networks and neighborhoods of Malegaon, India*. Soc Sci Med. 2016;153:99-106.

292. Barclay VC, Smieszek T, He J, Cao G, Rainey JJ, Gao H, et al. *Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity*. PLoS One. 2014;9(2):e87042.

293. B. B. C. News. *Italy makes 12 vaccinations compulsory for children - BBC News*: BBC; 2017. Available from: http://www.bbc.co.uk/news/world-europe-39983799.

294. Trump, The Doctor and the Vaccine Scandal: Channel 4 Dispatches [press release]. Channel 42017.

295. Department for Education. *Foundation Stage Profile Attainment by Pupil Characteristics in England, 2009/10*2010 (SFR 39/2010). Available from: http://www.education.gov.uk/researchandstatistics/statistics/allstatistics/a00196625/attainment-by-pupil-characteristics-foundation-st.

296. NHS Information Centre. *Compendium of Clinical & Health Indicators Interactive Atlas.*; 2011 [Date Accessed July 2011]. Available from: http://www.nchod.nhs.uk/NCHOD/IASingle.nsf/atlas.html.

8. Appendices

8.1. Relating to Chapter 1: General Introduction

8.1.1. HPA Q11-1 request parameters

2011/2012- 1 REQUEST PARAMETERS FOR COVER DATA: EVALUATION QUARTER 01/04/11 to 30/06/11

The following groups of children are to be included as PCT responsible population for COVER data.
Children for whom the PCT is responsible are:

- all children registered with a GP whose practice forms part of the PCT, regardless of where the child is resident, plus
- any children not registered with a GP, who are resident within the PCT's statutory geographical boundary

Note that children resident within the PCT geographical area, but registered with a GP belonging to another PCT, are the responsibility of that other PCT.

Request 1: 12 MONTH COHORT

1. Total number of children for whom the PCT is responsible on 30/06/11 reaching their 1st birthday during the above evaluation quarter.

2. Total number and percentage vaccinated (to one decimal place) included in line 1 completing a course* at any time up to their 1st birthday for each of the following:

DTaP/IPV/Hib	MenC	PCV
%	%	%

Request 2: 24 MONTH COHORT

3. Total number of children for whom the PCT is responsible on 30/06/11 reaching their 2nd birthday during the above evaluation quarter.

4. Total number and percentage vaccinated (to one decimal place) included in line 3 completing a course** at any time up to their 2nd birthday and also total number and percentage included in line 3 receiving boosters for each of the following:

DTaP/IPV/Hib	MMR	MenC *infant*	Hib/MenC** *Booster*	PCV** *Booster*
%	%	%	%	%

***at 12 months completed courses are defined as:**

DTaP/IPV/Hib is 3 doses before 1st birthday; if child received primary immunisations outside UK then 3 doses of each: DTP or DTaP, IPV or OPV, Hib before 1st birthday

MenC and PCV is 2 doses before 1st birthday (PCV can be either PCV7 or PCV13, given in any combination)

****at 24 months completed courses are defined as:**

DTaP/IPV/Hib is 3 doses before 2nd birthday; if child received primary immunisations outside UK then 3 doses of each: DTP or DTaP, IPV or OPV

MMR is 1 dose on or after 1st birthday and before 2nd birthday (i.e. **excludes** MMR given before 1st birthday)

MenC infant is 2 doses before 1st birthday

Hib/MenC booster is either (i) one dose of combined Hib/MenC vaccine on or after 12 months and before 2nd birthday
or (ii) 1 dose of DTaP/IPV/Hib and 1 dose of MenC, both given on or after 1st birthday and before 2nd birthday (i.e. children completing primary course after 1st birthday)

PCV booster is one dose on or after 12 months (irrespective of the number of doses before that age) and before 2nd birthday

8.2. Relating to Chapter 3: Initial Modelling

8.2.1. Results for N=400

Model results for N=400, as comparison with vs N=4000 (with constraint that MVD>$ln\ (N)$ relaxed) are given in Figure 8-1, Figure 8-2 and Figure 8-3

Figure 8-1 Distribution of vaccine-support (%) post decision process – N=400

Proportion of all network vertices which have of final opinion-status “support”.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.

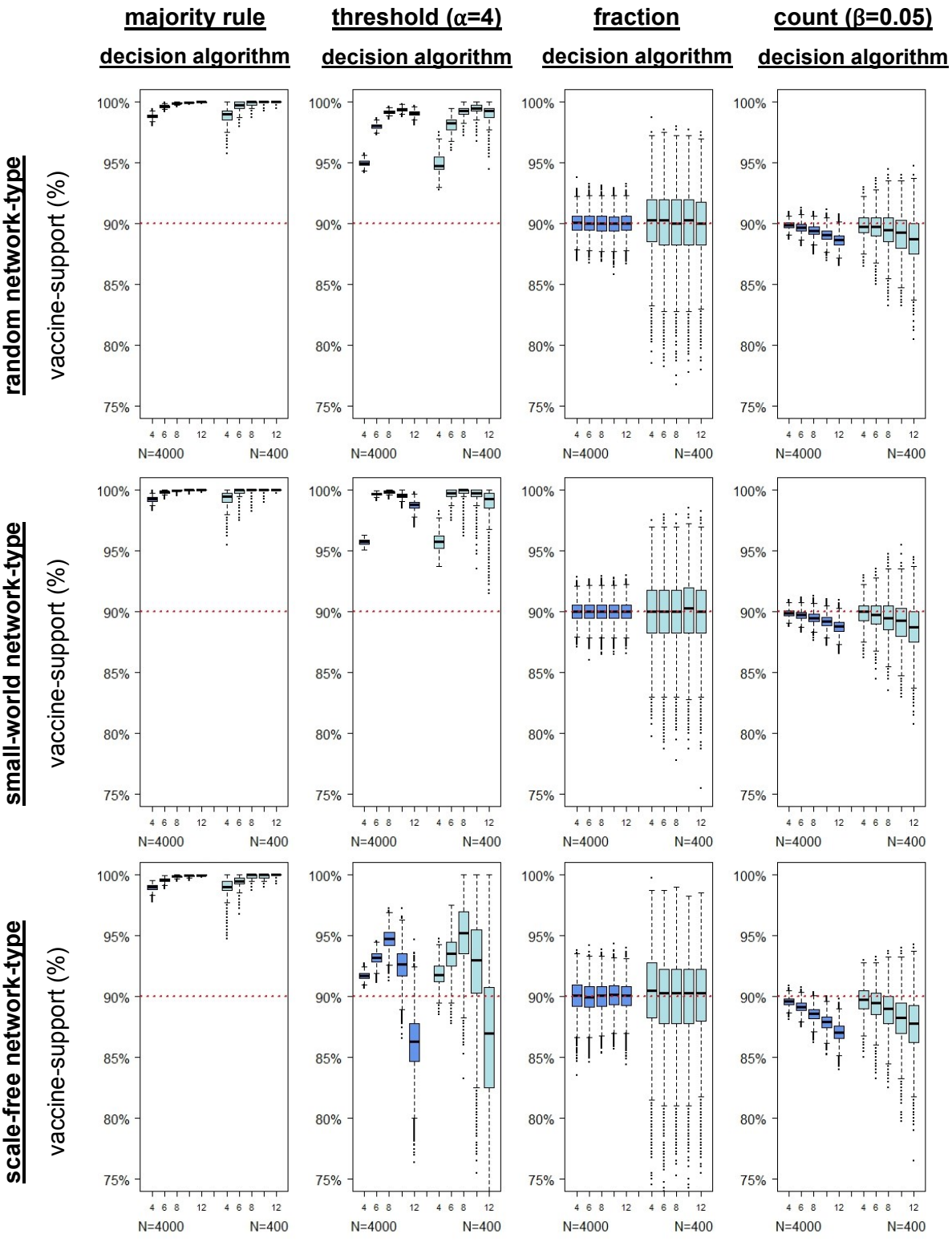


Figure 8-2 Intra-dyad agreement (%) post decision process – N=400

Proportion of all network edges which connect vertices of same final opinion-status.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.

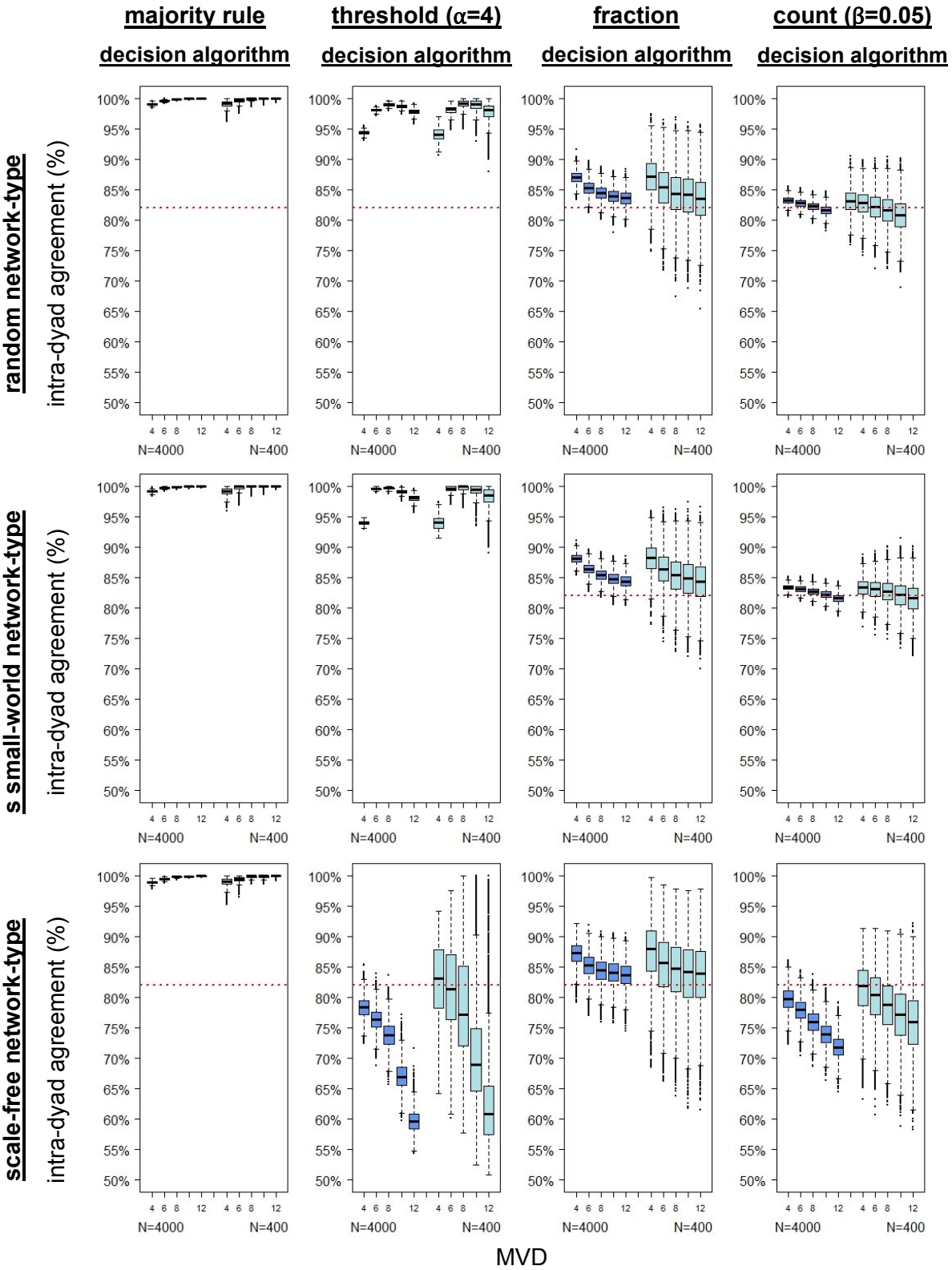
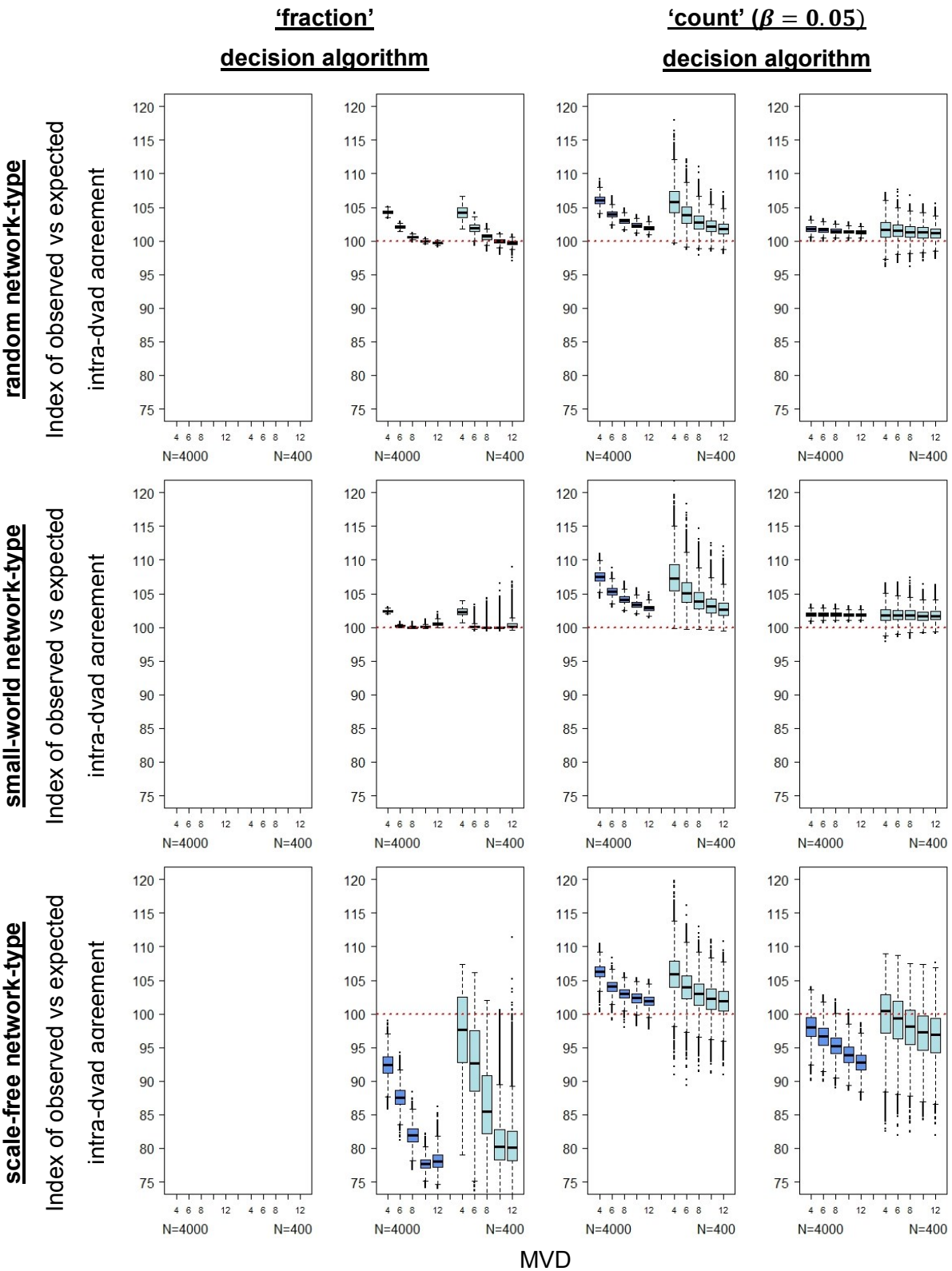


Figure 8-3 Intra-dyad agreement post decision making process, observed vs expected value – N=400

Index: expected value = 100, calculated by simulation.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and (specified) decision algorithm.



8.2.2. Results for count decision-algorithm, sensitivity to β

Model results for the count decision-algorithm with selected β values from range [0.0125, 0.1] are given in Figure 8-4, Figure 8-5 and Figure 8-6.

Figure 8-4 Distribution of vaccine-support (%) post decision process – vary β

Proportion of all network vertices which have of final opinion-status "support".

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'count' decision-making algorithm with selected β values.

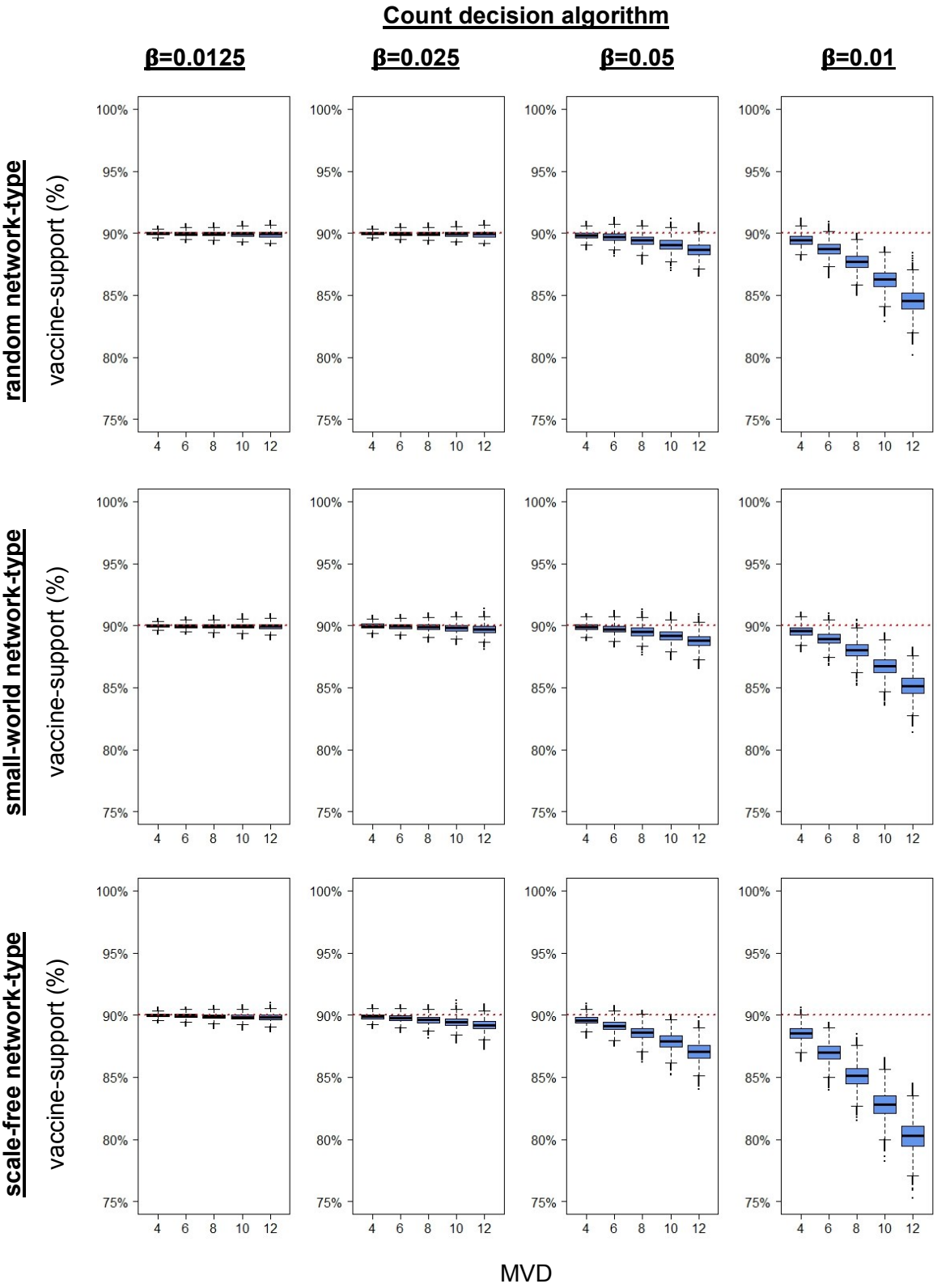


Figure 8-5 Intra-dyad agreement (%) post decision process – vary β

Proportion of all network edges which connect vertices of same final opinion-status.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'count' decision-making algorithm with selected β values

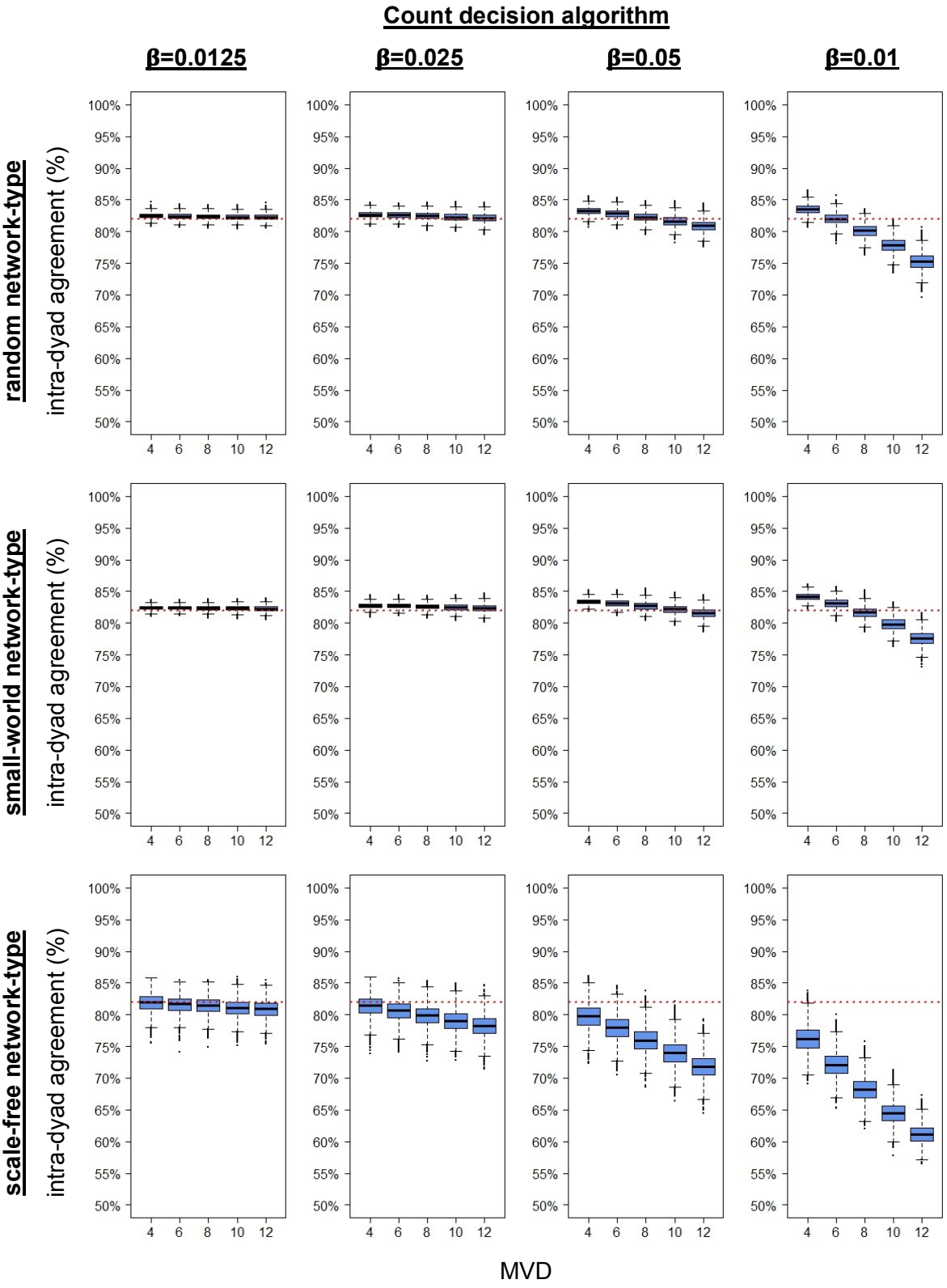
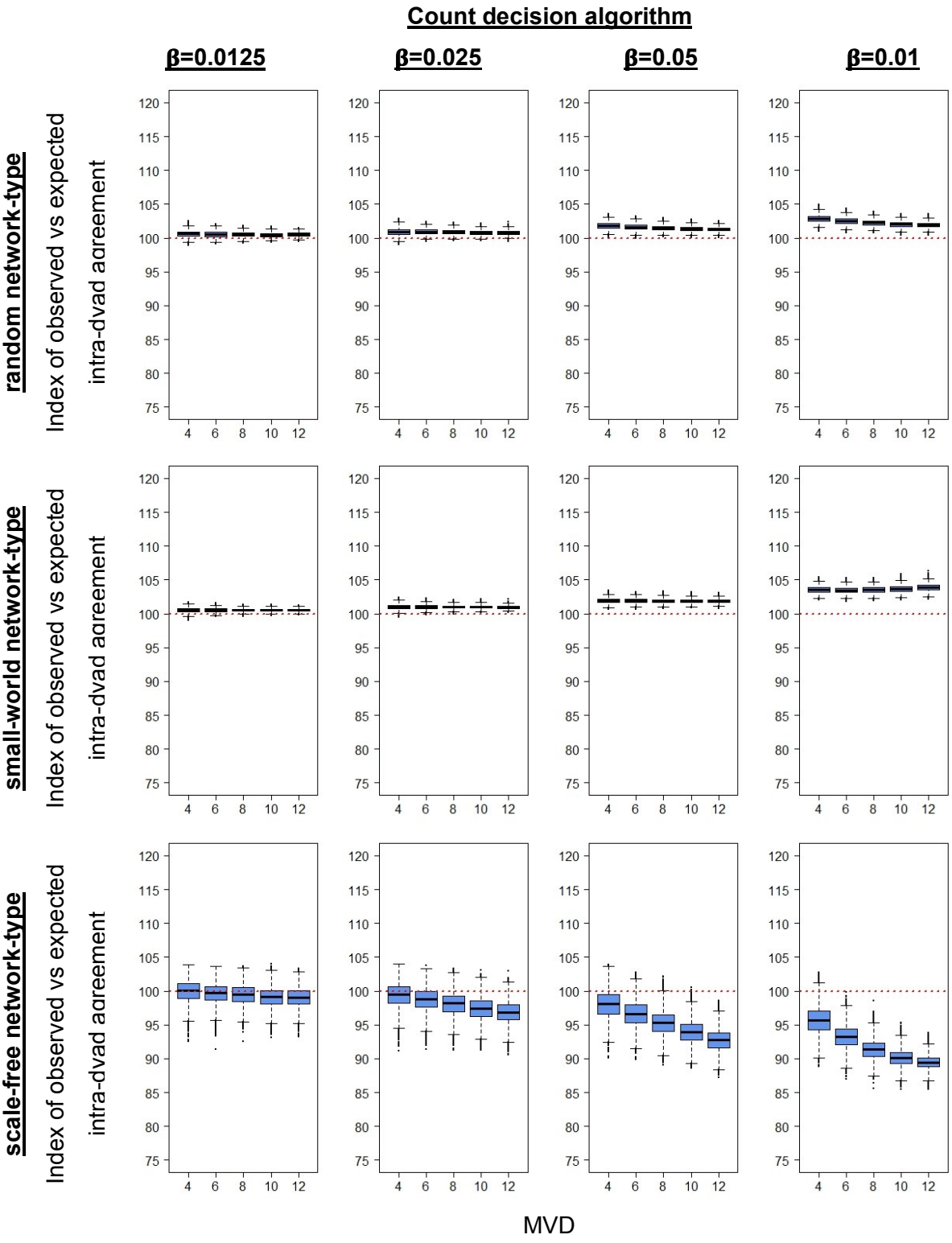


Figure 8-6 Intra-dyad agreement post decision process, observed vs expected value – vary β

Index: expected value = 100, calculated by simulation.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'count' decision-making algorithm with selected β values



8.2.3. Results for threshold decision-algorithm, sensitivity to α

Model results for the threshold decision-algorithm with selected α values from range [3,6] are given in Figure 8-7, Figure 8-8 and Figure 8-9.

Figure 8-7 Distribution of vaccine-support (%) post decision-process – vary α

Proportion of all network vertices which have of final opinion-status "support".

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'threshold' decision-making algorithm with selected α values.

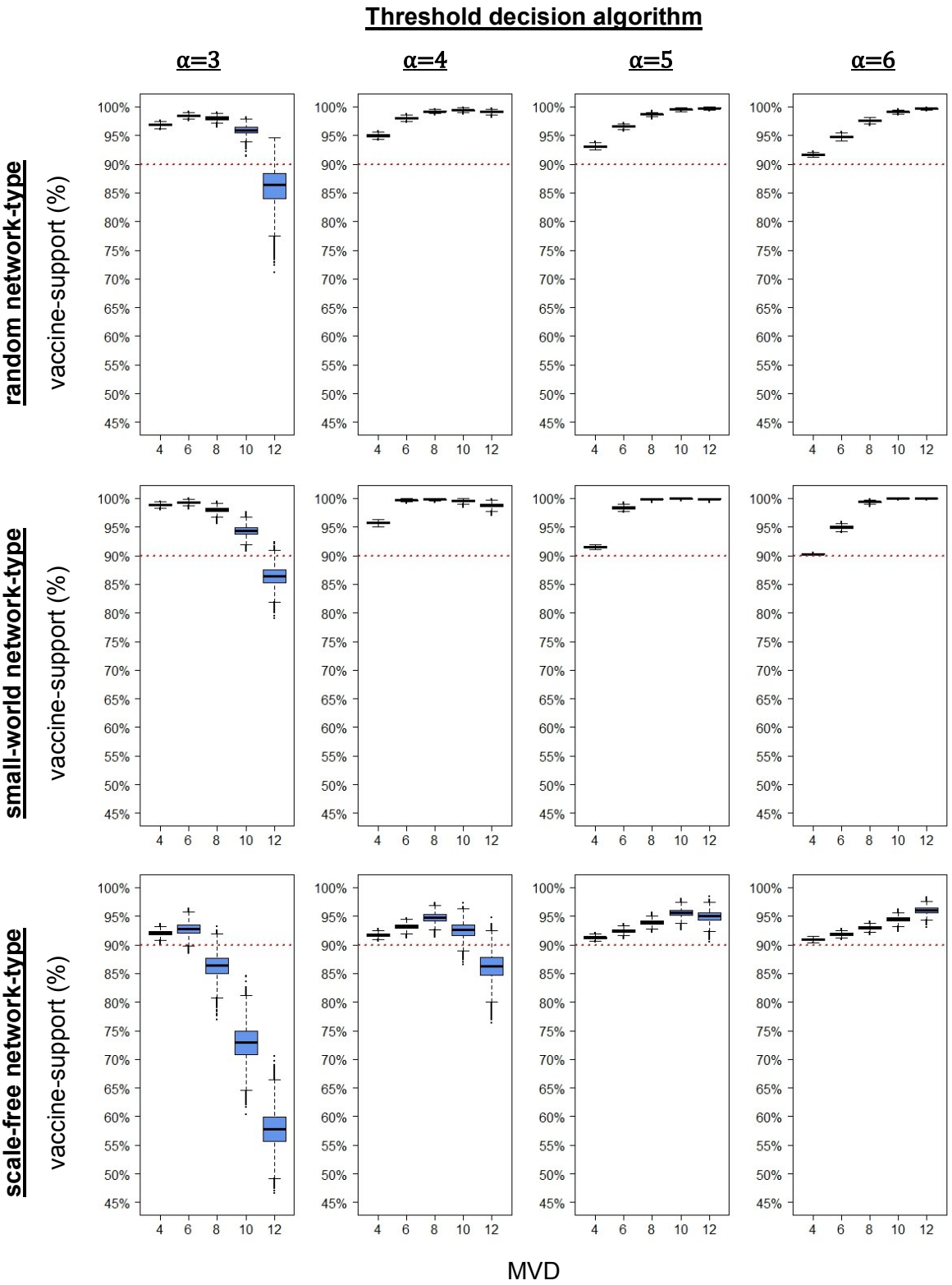


Figure 8-8 Intra-dyad agreement (%) post decision process – vary α

Proportion of all network edges which connect vertices of same final opinion-status.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'threshold' decision-making algorithm with selected α values

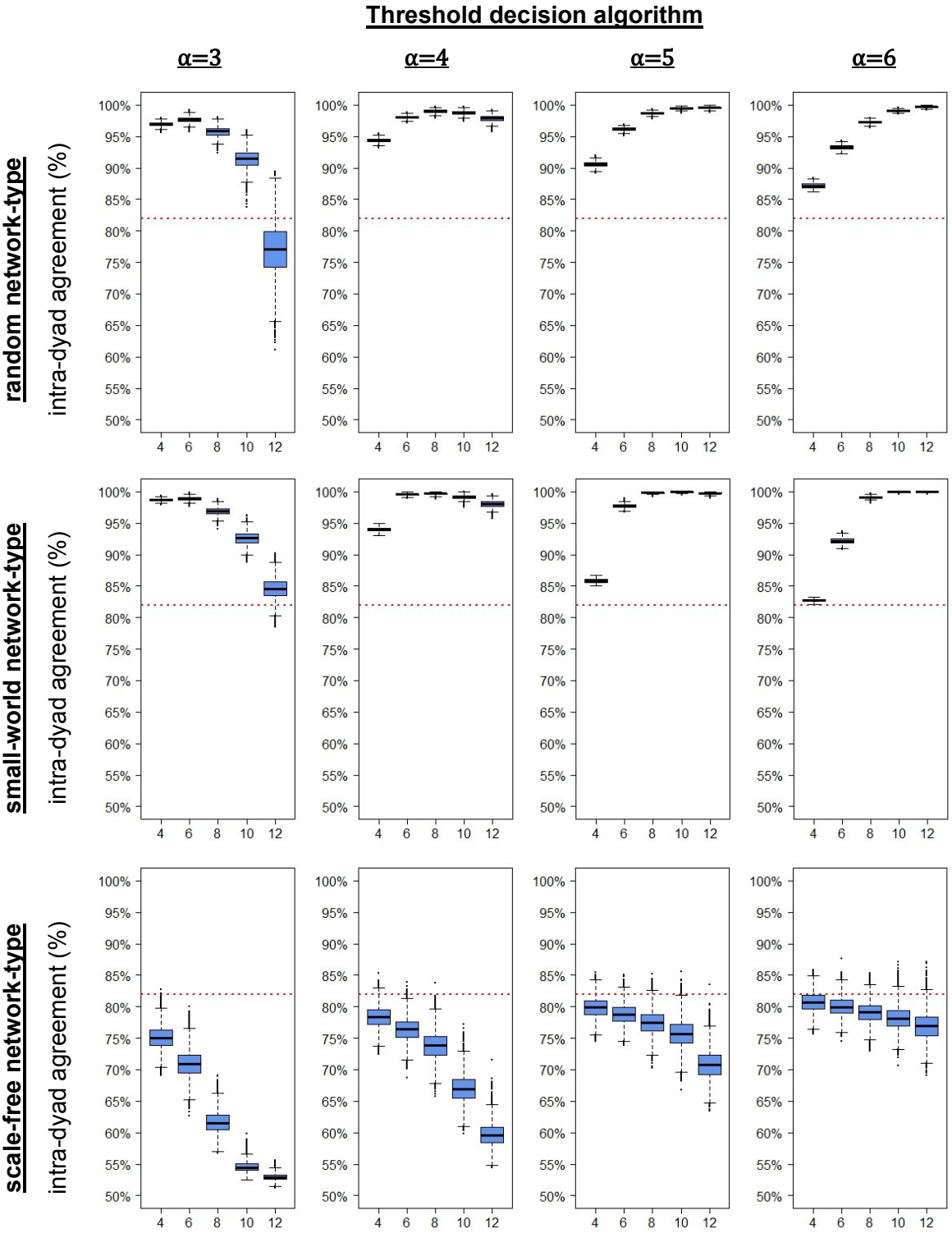
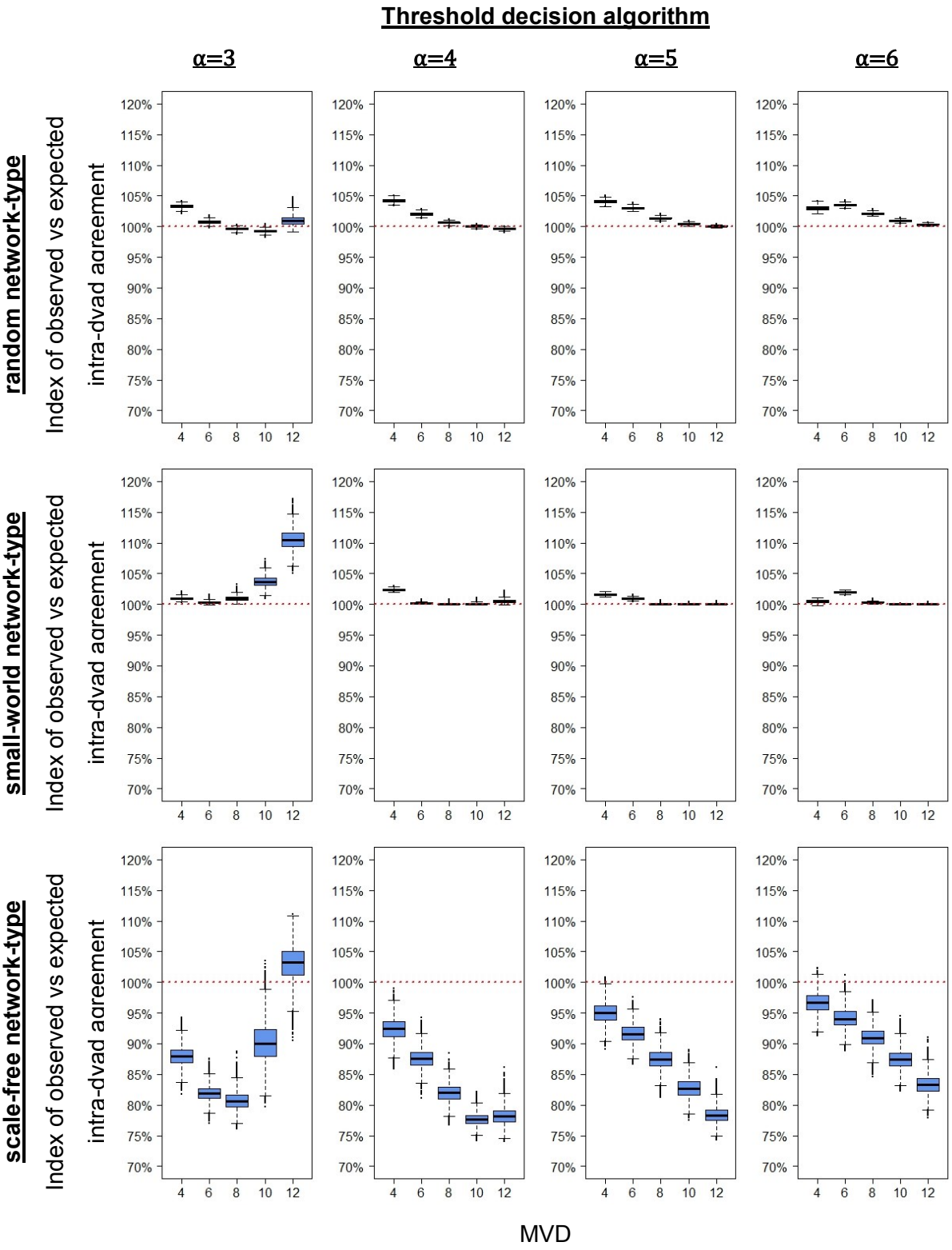


Figure 8-9 Intra-dyad agreement post decision process, observed vs expected value – vary α

Index: expected value = 100, calculated by simulation.

Box-plot of observed values across 10,000 simulations for each combination of network structure (network-type and MVD) and the 'threshold' decision-making algorithm with selected α values



8.3. Relating to Chapter 4: Survey

8.3.1. Survey ethical procedures and safeguards

As the questionnaire is self-completed consent is implicit with the return of a completed questionnaire to the researcher. The childcare setting provided written approval that they consented to act as recruitment centre, and to act within the ethical guidelines provided (§8.3.2.1). This approval was gained before the survey materials were supplied by the researchers for forwarding to parents. These materials included the survey's "Invitation to Participate" (§8.3.2.2) and "Participant Information Sheet" (§8.3.2.3), which were also on the landing page of the online survey. Recruitment centres did not provide researchers with any contact details for potential participants.

In order to be able to link the responses from individuals into a connected network, we required real names for people's contacts. For child contacts, as minors, the questionnaire instructions both requested that full names not be supplied and explicitly stated the use that was being made of the names (to identify duplicates within the survey responses). At the end of the survey fieldwork period, all links between participants were made before, and the records were then anonymised (as names are not required at any later stage in the processing) before the remaining analysis was conducted. All real names were permanently removed from the database (and original records destroyed) and there is no coding key (for participants or non-participating contacts) that could be used to identify individuals. (This necessitated a time limit after which data from a respondent withdrawing consent can no longer be removed from the dataset.)

The survey was conducted under the regulations specified by the Data Protection Act 1998. The questionnaire responses were confidential and all response data was encrypted. Encryption of those data obtained electronically was conducted at source. Paper questionnaires were supplied with an envelope that could be used by respondents to send directly back to the researchers in confidence.

Centres using paper questionnaires (paper or mixed presentation) were supplied with a ballot-box style collection box to be kept in a place accessed by parents; researchers retrieved these boxes, unopened, at the end of fieldwork. Paper survey packs also included an envelope to provide confidentiality of completed questionnaires, as either reassurance that the contents of the collection box were not accessible by the centre, or to be used to send the response directly back to the researchers in confidence. Completed paper questionnaires were transcribed electronically and the paper versions securely destroyed.

Steps have been taken to avoid situations where participants, who happen to be health care professionals (HCP), and who give vaccination advice in this role are asked to give details of the recipients of this advice and thus potentially compromise patient confidentiality. If, by chance, an HCP wished to participate, they were instructed to consider their replies as they relate to the vaccination of their own child and not in their professional capacity (as with all participants they are under no obligation to participate and are free to later withdraw consent). Note that the vaccination status of the children of participating parents is represented solely by parental recall (if volunteered as a response to a survey questionnaire) – this information is not sourced from, nor checked against, personal medical records.

Also, it was hoped that the existing awareness of a historical scare involving MMR would minimise the risk that the act of asking questions would inadvertently raise false suspicions amongst participants about MMR or other vaccines.

8.3.2. Survey materials

8.3.2.1. Centre confirmation form

Imperial College London

Vaccination Information Survey

CENTRE PARTICIPATION CONSENT

Name of Childcare Setting: _____

Address of Setting: _____

_____

Name of Authoriser:
(Manager/Senior Staff Member) _____

Signature: _____

Date: _____

Permission is given for the Research Team to approach parents and guardians of children at this setting to request their participation in the above Survey (covered by Ethical Approval IC_REC_12_2_2).

All parents and guardians of pre-school children will be invited to participate – with the exception of members of vulnerable groups (prisoners, mentally ill or under 16s) where this status is already known to setting management or staff.

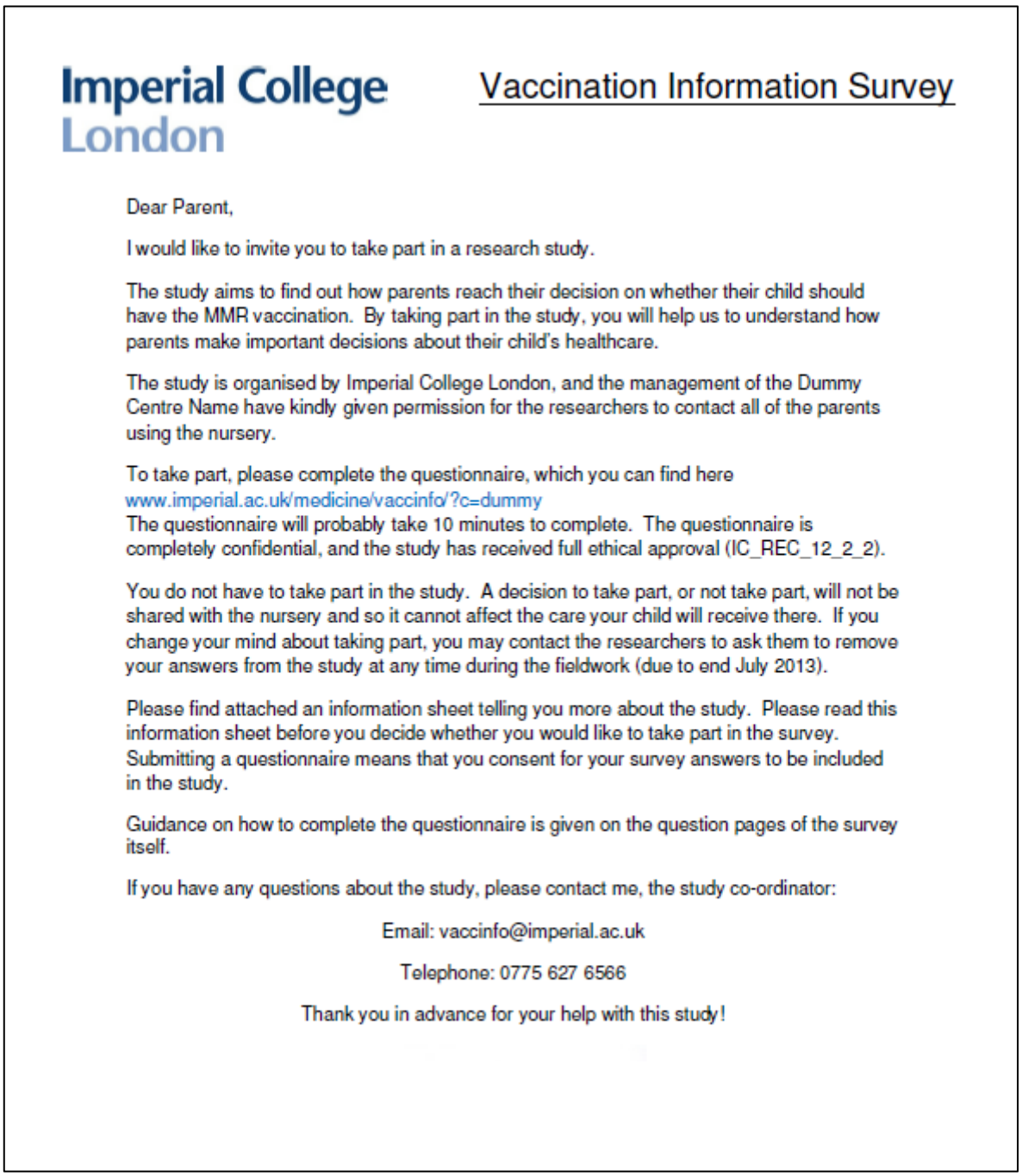
If the setting staff or management should become aware whether a specific person has participated, or not participated, in the survey then this will not affect the care given to their child.

If you have any questions about this study, please contact Diane Pople, the study co-ordinator:

Email: vaccinfo@imperial.ac.uk

Telephone : 0775 627 6566

8.3.2.2. Invitation to participate



Imperial College London

Vaccination Information Survey

Dear Parent,

I would like to invite you to take part in a research study.

The study aims to find out how parents reach their decision on whether their child should have the MMR vaccination. By taking part in the study, you will help us to understand how parents make important decisions about their child's healthcare.

The study is organised by Imperial College London, and the management of the Dummy Centre Name have kindly given permission for the researchers to contact all of the parents using the nursery.

To take part, please complete the questionnaire, which you can find here
www.imperial.ac.uk/medicine/vaccinfo/?c=dummy
The questionnaire will probably take 10 minutes to complete. The questionnaire is completely confidential, and the study has received full ethical approval (IC_REC_12_2_2).

You do not have to take part in the study. A decision to take part, or not take part, will not be shared with the nursery and so it cannot affect the care your child will receive there. If you change your mind about taking part, you may contact the researchers to ask them to remove your answers from the study at any time during the fieldwork (due to end July 2013).

Please find attached an information sheet telling you more about the study. Please read this information sheet before you decide whether you would like to take part in the survey. Submitting a questionnaire means that you consent for your survey answers to be included in the study.

Guidance on how to complete the questionnaire is given on the question pages of the survey itself.

If you have any questions about the study, please contact me, the study co-ordinator:

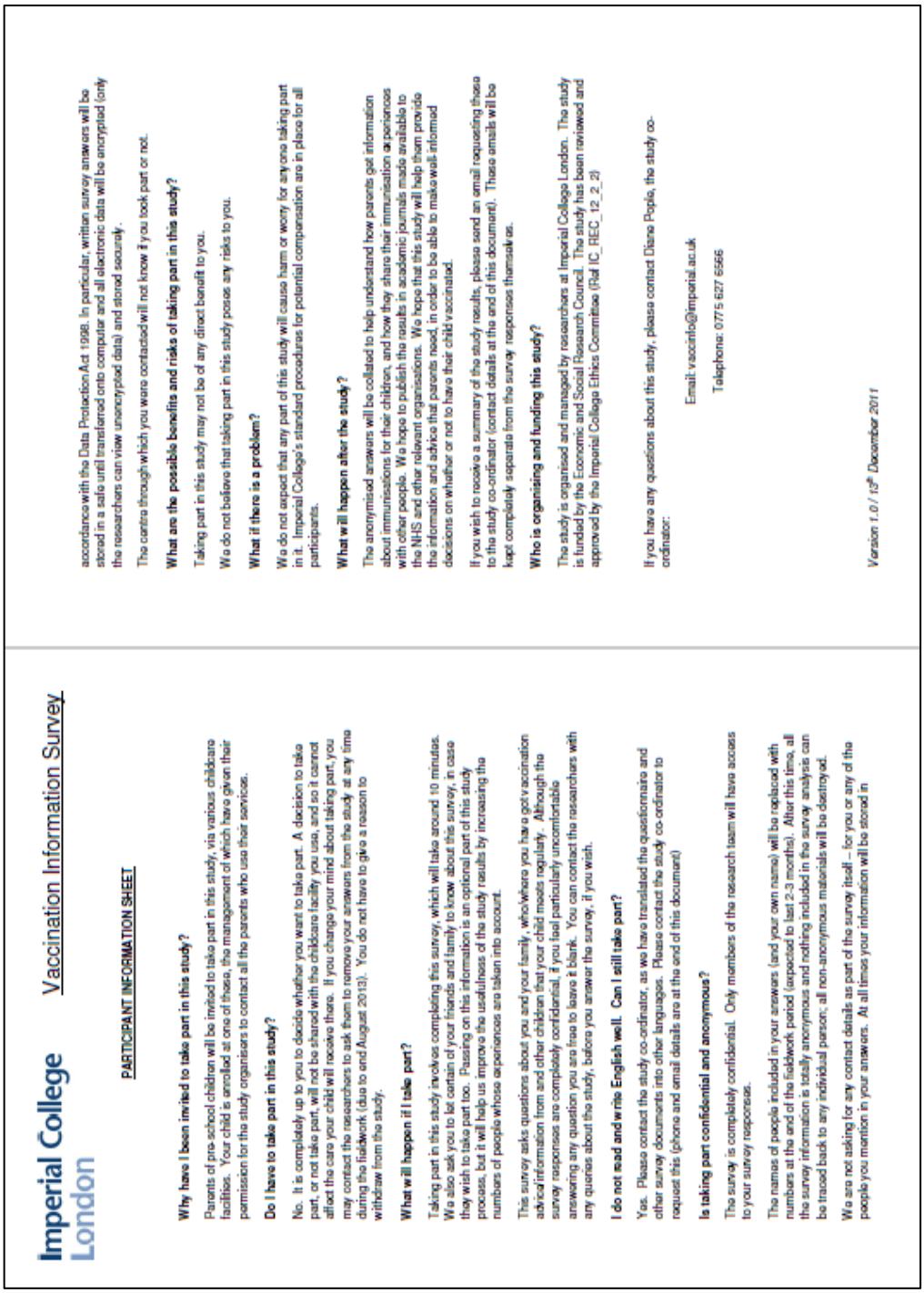
Email: vaccinfo@imperial.ac.uk

Telephone: 0775 627 6566

Thank you in advance for your help with this study!

This was customised with the name of the centre (in place of the dummy details here). It was included in the printed survey pack or as pdf attachment to an email from the centre management, as appropriate to the delivery.

8.3.2.3. Participant information sheet



Imperial College London — Vaccination Information Survey

PARTICIPANT INFORMATION SHEET

Why have I been invited to take part in this study?

Parents of pre-school children will be invited to take part in this study, via various childcare facilities. Your child is enrolled at one of these, the management of which have given their permission for the study organisers to contact all the parents who use their services.

Do I have to take part in this study?

No. It is completely up to you to decide whether you want to take part. A decision to take part, or not take part, will not be shared with the childcare facility you use, and so it cannot affect the care your child will receive there. If you change your mind about taking part, you may contact the researchers to ask them to remove your answers from the study at any time during the fieldwork (due to end August 2013). You do not have to give a reason to withdraw from the study.

What will happen if I take part?

Taking part in this study involves completing this survey, which will take around 10 minutes. We also ask you to let certain of your friends and family to know about this survey, in case they wish to take part too. Passing on this information is an optional part of this study process, but it will help us improve the usefulness of the study results by increasing the numbers of people whose experiences are taken into account.

This survey asks questions about you and your family, who/where you have got vaccination advice/information from and other children that your child mixes regularly. Although the survey responses are completely confidential, if you feel particularly uncomfortable answering any question you are free to leave it blank. You can contact the researchers with any queries about the study, before you answer the survey, if you wish.

I do not read and write English well. Can I still take part?

Yes. Please contact the study co-ordinator, as we have translated the questionnaire and other survey documents into other languages. Please contact the study co-ordinator to request this (phone and email details are at the end of this document)

Is taking part confidential and anonymous?

The survey is completely confidential. Only members of the research team will have access to your survey responses.

The names of people included in your answers (and your own name) will be replaced with numbers at the end of the fieldwork period (expected to last 2-3 months). After this time, all the survey information is totally anonymous and nothing included in the survey analysis can be traced back to any individual person; all non-anonymous materials will be destroyed.

We are not asking for any contact details as part of the survey itself – for you or any of the people you mention in your answers. At all times your information will be stored in accordance with the Data Protection Act 1998. In particular, written survey answers will be stored in a safe until transferred onto computer and all electronic data will be encrypted (only the researchers can view unencrypted data) and stored securely.

The centre through which you were contacted will not know if you took part or not.

What are the possible benefits and risks of taking part in this study?

Taking part in this study may not be of any direct benefit to you.

We do not believe that taking part in this study poses any risks to you.

What if there is a problem?

We do not expect that any part of this study will cause harm or worry for anyone taking part in it. Imperial College's standard procedures for potential compensation are in place for all participants.

What will happen after the study?

The anonymised answers will be collated to help understand how parents get information about immunisations for their children, and how they share their immunisation experiences with other people. We hope to publish the results in academic journals made available to the NHS and other relevant organisations. We hope that this study will help them provide the information and advice that parents need, in order to be able to make well-informed decisions on whether or not to have their child vaccinated.

If you wish to receive a summary of the study results, please send an email requesting these to the study co-ordinator (contact details at the end of this document). These emails will be kept completely separate from the survey responses themselves.

Who is organising and funding this study?

The study is organised and managed by researchers at Imperial College London. The study is funded by the Economic and Social Research Council. The study has been reviewed and approved by the Imperial College Ethics Committee (Ref IC_REC_12_2_2)

If you have any questions about this study, please contact Diane Popla, the study co-ordinator:

Email: vaccinfo@imperial.ac.uk

Telephone: 0775 627 6566

Version 1.0 / 13th December 2011

Actual size was double-sided A4, included in the printed survey pack or as pdf attachment to email

8.3.2.4. Paper questionnaire

Imperial College London

Vaccination Information Survey

This study aims to find out how parents reach their decision on whether their child should have a routine childhood vaccination. By taking part in the study, you will help us to understand how parents make important decisions about their child's healthcare. However, you do not have to be a parent to take part in this study.
To take part, please complete this questionnaire, which will probably take about 10 minutes to complete.
The questionnaire is confidential and the study has received full ethical approval.
If you would rather answer this questionnaire online you can find it here:
www.imperial.ac.uk/medicine/vaccinfo /?c=dummy
Please read the attached "Participant Information Sheet" which may answer some questions you have about the study.

NB if you are a health professional, please answer this from a personal perspective; do not include any people that you meet or advise in a purely professional capacity.

Please tell us about yourself

Q1 What is your name?
Tell us the name which your friends use for you (given name and surname)

Q2 What is the postcode of your home?
Please tell us at least the first 5 characters

Q3 What is your sex?
Tick one
Female ☐ Male ☐

Q4 How old are you?
Tick one
17 years or younger ☐ 18-24 years ☐ 25-34 years ☐ 35-44 years ☐ 45 years or older ☐

Q5 What is your ethnic group?
Tick the one that is the closest match to how you would describe yourself
White: ..... English/Welsh/Scottish/Northern Irish/ British ☐
White: ..... Any other white background ☐
Asian/Asian British: ..... Indian / Pakistani / Bangladeshi ☐
Asian/Asian British: ..... Chinese ☐
Asian/Asian British: ..... Any other Asian background ☐
Black/Black British: ..... African ☐
Black/Black British: ..... Caribbean ☐
Black/Black British: ..... Any other black background ☐
Mixed / multiple ethnic background ☐
Any other ethnic group ☐

Q6 Which of these academic qualifications do you have?
Tick all that apply
If you have an academic qualifications that is not listed (e.g. from outside England & Wales), tick the nearest equivalent(s) shown
Postgraduate qualifications (eg PhD, MBA) ☐
First degree (eg BA, BSc) *or* degree-level Professional qualification *or* other qualification requiring A levels for entry ☐
One or more A Levels *or* equivalent ☐
5 or more GCSE (grade A*-C) *or* O levels (passes) *or* CSEs (grade 1) ☐
1-4 GCSE (grade A*-C) *or* O levels (passes) *or* CSEs (grade 1) ☐
No academic qualifications ☐

Please tell us about your family

Q7 How many children (born in 1997 or later) do you have?
Include only those that were living with you when they turned 1 year old (or who are under 1 now) ☐

Q8 Do you have any children aged under 5?
Tick one, include only children that live with you
Yes ☐ No ☐
If "No" then GO TO ● (last page)

Please answer the following 4 questions for each of your under 5's (one column per child)

Write the answer for each child in the column under their name (as given in question 8)
(if you have more than four under 5's then you can use the online version of this survey)

Q9 What is their name?

Tell us the (given) name your friends use for them

			example
			OLIVIA

Q10 How old are they?

Fill in the years and months for each child, in the column under their name

			example
yrs mths	yrs mths	yrs mths	2 yrs 10 mths

MMR is a vaccination that protects against measles, mumps and rubella (german measles) and is normally available for children aged 13 months and older.

Q11 Have they ever received an MMR jab?

Tick all that apply for each child, in the column under their name

				example
Yes (before 2013)				
Yes (in 2013)				
No				
Don't know / Can't remember				✓

Think of all the other pre-school child(ren) that your child mixes with - this could be at childcare, whilst visiting friends and relatives, at a playground etc

Q12 Please tell us the names of the pre-school children that your child mixes with in a typical week (during school term-time) - include weekdays and weekends.

It may help to think back over the last week, day-by-day, to help remember everyone.

Please include all pre-school children that are either in the same room as your child for more than about 15minutes, or who have face-to-face contact with your child.

Please tell us the other child's given name and the initial of their surname (if you know it).

NB The study has been reviewed and approved by an ethics committee: for details of the data protection measures being enforced, please see the Participant Information Sheet.
At the end of the fieldwork period, all names are removed and only anonymous information kept.
This information is useful only to confirm whether your child knows those of other people who are taking part in the study - we will not attempt to identify or contact the children themselves.

Children who your child mixes with at childcare:

In the first row, estimate the number of children in your child's room, and
on the following rows, list by name those that you know are friends with your child

			example
Children in room	Children in room	Children in room	11 Children in room
			EMILY W
			MO K
			JESS B
			MIA S

Children your child meets in other places:

If there is a group of children that your part of, but you don't know (all) their names:
name the place/activity and put the approximate number of children your child mixes with in the small box

						example	
						OLLY G	
						SOPHIE G	
						LILY	
						SWIMMING	8
						PARK	1

If there are not enough spaces, then you can use the online version of the survey or attach a photocopy to your reply

If your child does not mix with any other under 5's (except those who live in the same house) then tick here ☐

Please answer the question on this page about the youngest of your children who is at least 1 year old
(or you only have one pre-school child, please answer about them)

Q13 Please confirm their name

Use the same name for them as you wrote in Q9

You may have already made a decision on whether to get this child the MMR jab, or you may be thinking about it now

Q14 Please tell us the names of all the people with whom you can remember discussing vaccinations to protect children against measles (e.g. MMR) including giving/receiving advice, information or opinions on this subject.

It may help to look at the types of people in Q15, to remember everyone that you may have discussed this with.

For older children, please think back to when you were first making a decision.

- *Please include people regardless if the information/advice shared came from you or was given to you.*
- *Please tell us the given name and surname for each person (or as much as you know - if you only know them by the job in which you encounter them then please write this instead)*
- *Please say roughly where they live, e.g. the name of a town or city suburb, or the first part of postcode*

Q15 How do you know this person?

Tick one (per person).
If more than one applies, please tick the one that is furthest to the left

Q16 Which of these descriptions is the closest match to this person's advice or opinion?

Tick one (per person).

Q17 Is this person the parent of a pre-school child?

Given Name	Surname	Where they live	My partner/husband/wife	My mother/father (or step-parent)	My partner's parents	Other family	My friend	Nursery/Playgroup/Childcare Staff	Health Visitor	GP/My doctor	Practice Nurse/ Nurse in GP surgery	Nurse in Children's Clinic	Midwife	Other healthcare professional	Other	Overall, they thought you should get the MMR jab done when it is due	Overall, they thought you should get the MMR jab, but wait until the child is much older	Overall, they thought you should refuse to give your child the MMR jab	It really wasn't clear what they thought overall	I can't remember what they thought/ they did not tell me their opinion	Yes	No	Don't Know

If you did not tick "Yes" to anyone in Q17 then GO TO ✱ (last page)

If you have not discussed this subject (vaccinating to protect against measles) with anyone, then tick here ☐ *and GO TO ✱ (last page)*

For every person you marked as having pre-school child(ren) in Q17 :

Write in the initials of the people in Q17 in the first column (in the same order)

Q18 Do they have a child (or children) who was included in your answer to Q12?

- *In the blue box(es) in each row, write in the given name of their child(ren) who mix with yours*
- *Write the names of your child(ren) in the pink boxes next to the child(ren) that they know.*

Adult in Q14 Initials	Their child	Your child	Their child	Your child

- *If there are not enough spaces, then you can use the online version of the survey or attach a photocopy to your reply*

If you were instructed to skip questions, then please re-start here

Thinking about your personal experience with measles and the MMR jab

Q19 Have you, or anyone you know personally, had measles recently? Yes ☐ No ☐
Tick one

Q20 Have you, or anyone you know personally, had a serious adverse reaction attributed to MMR jab?
Tick one Yes ☐ No ☐

Finally, to help us increase the number of people whose experiences contribute to this research we would appreciate it if you could ask the people that you've listed if they would be prepared to take part in the survey too. So please pass on survey details to these people you've listed:

- the adults in Q14 (it doesn't matter if they don't have young children of their own)
- the parents of the children in Q12 (if you know them and they don't have children at the same childcare)

except please do not pass onto healthcare professionals unless they are family members or personal friends

They can get their own copy of the questionnaire from this website:
www.imperial.ac.uk/medicine/vaccinfo/?c=dummy
This is not compulsory; it is OK to just fill-in your own questionnaire and not mention it to anyone else.

Thank you

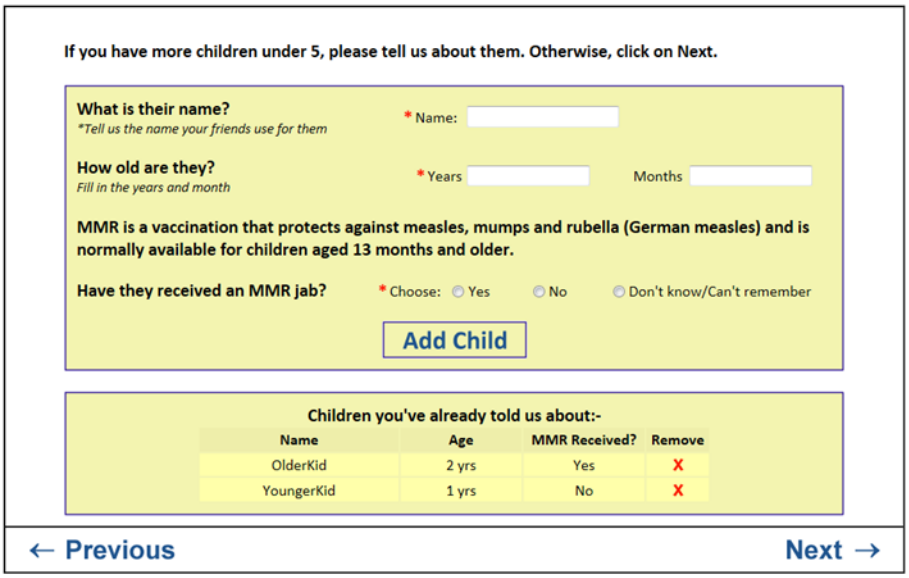
Please return this questionnaire to: Diana Pople
Department of Infectious Disease Epidemiology
Imperial College, St Mary's Hospital campus,
Norfolk Place
LONDON W2 1PG

or: leave it in the box at the nursery (you can use this envelope for confidentiality) from where the researchers will collect them

Version 4a dummy (dummy pct) / June 2013

This was customised with the centre's name (in place of the dummy details here). .

8.3.2.5. Web questionnaire, example screengrab



This screengrab shows the MMR question options as used in the pilot.
In April 2013, this was amended at the same time as the paper questionnaire to categorise "Yes" answers by time of vaccination.

8.3.3. EAL in pre-school households in the shortlisted PCTs

The National Pupil Database [295] records the use of English as first language by schoolchildren as a variable within records of Early Years Foundation Stage Profile achievement by local authority. A statistic derived from this data, as detailed below, is used as a proxy for the status of English language (specifically EAL) within the households of pre-school children.

It is noted that households with EAL status are not necessarily thus impeded from participation in a survey with an English-language questionnaire, but a source of appropriate English-language literacy skills has not been identified. Hence these EAL values are regarded as an upper bound on the proportion of potential respondents that may be affected by EAL-derived coverage bias, and has been used primarily to firstly identify that this is a matter of concern and secondly, to prioritise areas for further investigations.

The National Pupil Database records have been used to estimate a proportion of children assessed at the end of Early Years Foundation Stage (the year the child turns 5-years-old) who do not use English as first language. The value of this statistic for the shortlisted PCTs is given in Table 8-1 . The analysis indicates that the shortlisted regions include those

where the potential for EAL to create coverage and sample bias should be addressed in survey recruitment and presentation.

Table 8-1 EAL in shortlisted PCT

PCT	EAL[2]	PCT	EAL[2]
Barking & Dagenham	40%	Haringey Teaching	50%
Barnet	41%	Hartlepool	3%
Bexley Care Trust	13%	Havering	9%
Brent Teaching	64%	Herefordshire	3%
Bristol	16%	Hounslow	56%
Bromley	8%	Islington	41%
Camden	59%	Kensington & Chelsea	54%
City & Hackney Teaching[1]	54%	Kingston	30%
Coventry Teaching	27%	Lambeth	46%
Croydon	29%	Lewisham	32%
Dorset	2%	Newham	74%
Ealing	59%	Nottingham City	24%
Enfield	45%	Richmond & Twickenham	16%
Great Yarmouth & Waveney[1]	5%	Southwark	39%
Greenwich Teaching	37%	Surrey[1]	9%
Hammersmith & Fulham	47%	Wandsworth	40%

[1] English language usage is reported by local authority area; these following PCT borders to not match those of the overlapping local authority, hence population weighted average of the following local authorities have been used
- City & Hackney Teaching PCT uses City of London and Hackney
- Great Yarmouth & Waveney PCT uses Norfolk and Suffolk
- Surrey PCT uses Surrey
- Sutton & Merton PCT uses Sutton and Merton

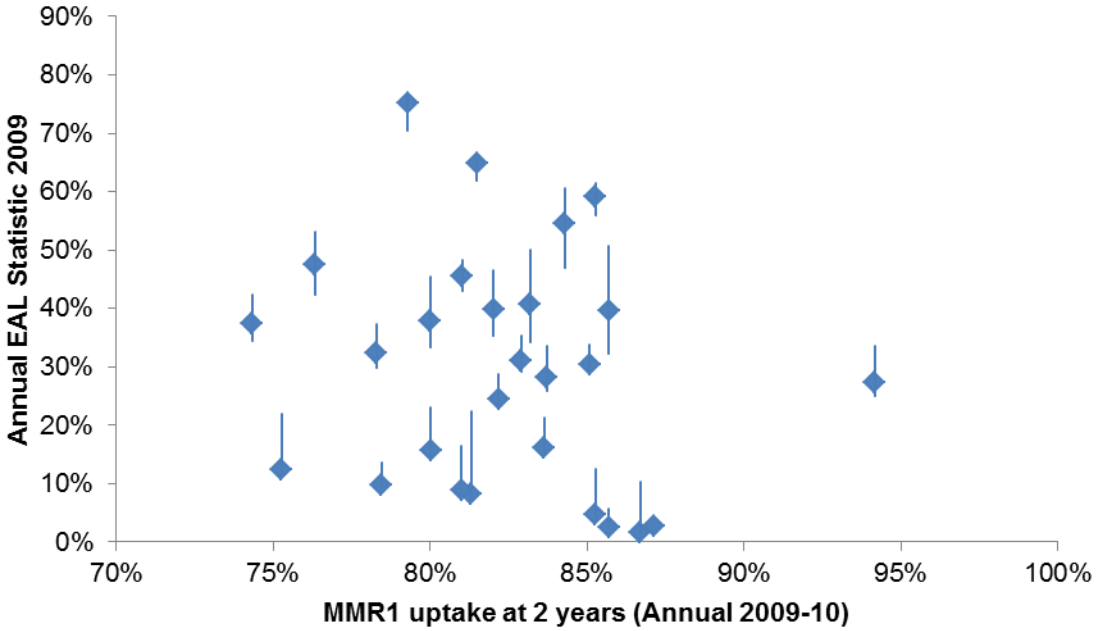
[2] Numerator incudes children whose first language is 'unknown but believed to be other than English' and the denominator is children (in the year they turn 5 years old) whose English language usage status was reported.
The moving average value is calculated from the population-weighted mean across 2007-2010, with annual figures excluded where less than 80% of children's status was reported.

EAL data derived from National Pupil Database [295] Geographic data from Compendium of Clinical & Health Indicators Interactive Atlas. [296]

Furthermore, given a key objective of the survey is the discovery of the information network across which MMR-related information is transmitted, we wish to understand the potential effect of EAL-derived bias on the participation by MMR-acceptor and MMR-rejecters. Figure 8-10 shows the 2009 annual EAL statistic for the nominated PCTs plotted against the reported 2009-10 annual MMR1 uptake. To indicate uncertainty in the EAL statistic due to incomplete records, upper and lower bounds where non-respondents are included and excluded within this category, respectively, have also been calculated, and 7 of the shortlisted PCTs are excluded as less than 80% of children's EAL status is reported.

A simple linear regression of MMR1 uptake on EAL statistic found a non-significant result ($p>0.05$), albeit a result which is subject to ecological bias.

Figure 8-10 EAL vs MMR1 uptake



EAL data derived from National Pupil Database [295]

MMR1 uptake from COVER Annual Report [36]

8.3.4. Reported MMR uptake for survey sample

Table 8-2 Stratified MMR uptake for survey sample

% Vaccinated		Current			Pre-catch-up			"Timely"
Censoring Children aged under:		14mth	24mth		14mth	24mth		
Total		97.8%	97.8%		92.1%	94.3%		94.1%
Age of Child (pre-catch-up: age at that time)	2 year old	100.0%			88.4%			95.9%
	3 year old	92.9%	**		100.0%	*		88.1%
	4 year old	100.0%			95.0%			97.8%
Ordinal of Child	1st	97.5%	97.9%		92.5%	95.8%		94.7%
	2nd	100.0%	100.0%		90.3%	90.0%		92.9%
	3rd or higher	93.3%	91.7%		92.3%	90.9%		91.7%
Centre PCT	Camden PCT	100.0%	100.0%		94.1%	100.0%		92.3%
	Enfield PCT	100.0%	100.0%		95.5%	100.0%		95.2%
	GYW PCT	97.4%	96.7%		90.9%	90.7%		91.7%
	Wandsworth PCT	96.6%	97.6%		91.5%	96.7%		97.6%
Child Attends Childcare	Yes	97.7%	97.7%		92.5%	94.1%		93.9%
	No	100.0%	100.0%		100.0%	100.0%		100.0%
Age of Respondent	18-24 years	100.0%	100.0%		100.0%	100.0%		100.0%
	25-34 years	98.1%	97.5%		95.5%	93.1%		95.0%
	35-44 years	97.3%	97.6%		90.6%	95.4%		94.1%
	45+ years	100.0%	100.0%		80.0%	80.0%		80.0%
Ethnicity of Respondent	White / White British	98.7%	99.2%		93.2%	95.7%		95.1%
	Black / Black British	80.0%	75.0%	**	80.0%	75.0%		75.0%
	Asian / Asian British	92.3%	87.5%	*	81.8%	85.7%		87.5%
	Other including Mixed	100.0%	100.0%		100.0%	100.0%		100.0%
Education of Respondent	Postgraduate	97.9%	100.0%		89.5%	100.0%		96.8%
	Graduate	98.8%	98.5%		94.5%	93.8%		95.5%
	A-Level	96.8%	96.2%		92.6%	95.8%		92.3%
	GCSE/None	94.4%	92.3%		85.7%	83.3%		84.6%

Fisher Exact homogeneity in proportions: * p<0.1** p<0.05

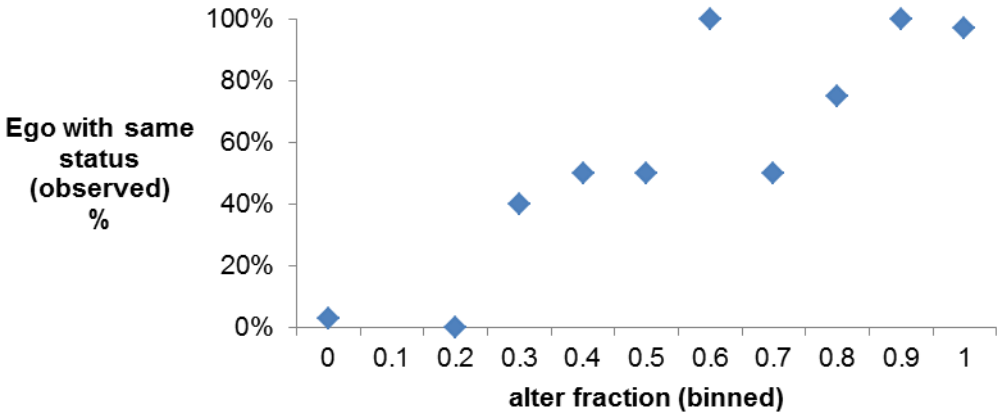
Base: all respondents with informative MMR status

8.4. Relating to Chapter 5: Revisiting the MMR1 Decision Model

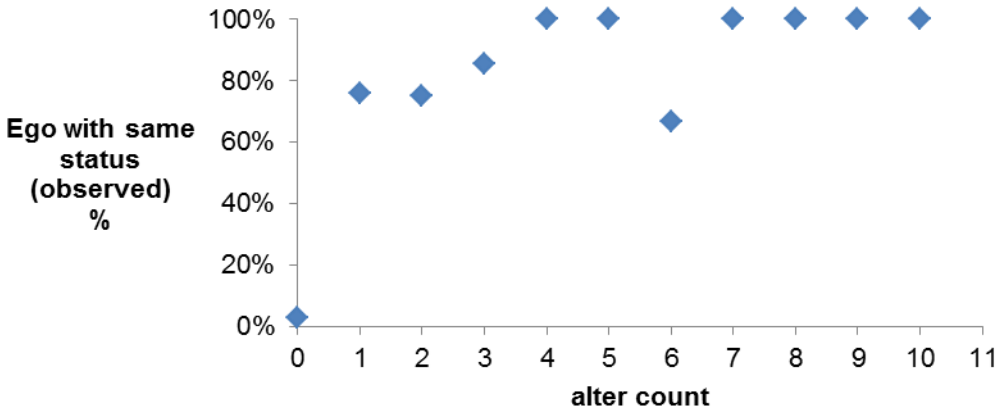
8.4.1. Initial visual inspection of decision-process data

Figure 8-11 Observed opinion status vs alters with same status

a] fraction measure of 'opposition'



b] count measure of 'opposition'



8.5. Permissions

Relating to Figure 2-1

30th May 2017

Dear Ordnance Survey,

I am completing my PhD thesis at Imperial College London entitled The effect of social networks on the participation by new mothers in the baby immunisation programme in the UK'.

I seek your permission to reprint, in my thesis an extract from: Vivancos et al "An Ongoing Large Outbreak of Measles in Merseyside, England, January to June 2012" published in Eurosurveillance, Volume 17, Issue 29, 19 July 2012 pii=20226. Available online@ http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20226

The extract to be reproduced is: **Figure 3.** Geographical distribution of confirmed (n=359) and probable (n=157) measles cases, Merseyside, England, January–June 2012

The article itself is covered by the journal's the Creative Commons Attribution (CC BY) licence, but I see this figure additionally carries "(c) Crown Copyright and database right 2010. All rights reserved, Ordnance Survey Licence number 100016969".

I would like to include the extract in my thesis which will be added to Spiral, Imperial's institutional repository http://spiral.imperial.ac.uk/ and made available to the public under a Creative Commons Attribution-NonCommercial-NoDerivs licence.

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Diane Pople

Permission granted for the use requested above:

I confirm that I am the copyright holder of the extract above and hereby give permission to include it in your thesis which will be made available, via the Internet, for non-commercial purposes under the terms of the user licence.

[please edit the text above if you wish to grant more specific permission]

Signed:

Name:

Organisation:

Job title: