Modelling Uncertainty in Environmental Health Impact Assessment

Marco Mesa-Frias

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy in Public Health & Policy

University of London

2015

Department of Social and Environmental Health Research

Faculty of Public Health & Policy

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by the Engineering and Physical Sciences Research Council (ESPRC) Grant no. EP/F007132/1)

I, Marco Mesa-Frias, confirm that the work presented in this thesis is my own. When information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: Date: 25/05/2015

Acknowledgements

First and foremost I would like to thank my supervisor Dr Zaid Chalabi and my co-supervisor Dr Anna M. Foss for all the support and inspiration. They have provided me with excellent guidance and immense practical support throughout my doctoral studies. I am thankful for the support not only professionally but also on a personal level. Also I would like to thank all members of my PhD Advisory Group Prof Mark Petticrew, Dr Shakoor Hajat and Dr Richard White for their input and advice throughout my studies. I am very grateful for the EPSRC funded project "PUrE Intrawise" for awarding me with a PhD studentship. Furthermore, I thank all the members of the Faculty of Public Health and Policy at the London School of Hygiene and Tropical Medicine (LSHTM) for accepting me into their faculty.

Also I thank all the members of the British Women Heart & Health study at the Faculty of Epidemiology and Population Health for giving me the opportunity to be part of their group. I thank Prof Juan Pablo-Casas, Prof Shah Ibrahim, Dr David Prieto-Merino and Dr Eveline Nuesch for their advice while working as a Research Assistant in Epidemiology. Warm thanks also go to Tim Marsh at the UK Health Forum Modelling Department for his support with time off from duties in order to work on my thesis. I would like to thank colleagues at LSHTM for stimulating discussions: Dr Tazio Vanni, Triantafyllos Pliakas, Dr Oscar Ocho and Dr Manuel Gomes. My last words of gratitude go to my family and my parents for their advice and profound example and Noelle for her love and encouragement. Thank You!

Abstract

Quantifying uncertainty in environmental health impact assessment models is important, particularly if the models are to be used for decision support. This thesis develops a new non-probabilistic framework to quantify uncertainty in environmental health impact assessment models. The framework takes into account two different perspectives of uncertainty: conceptual and analytical in terms of where uncertainty occurs in the model. The first perspective is concerned with uncertainty in the framing assumptions of health impact assessment, whereas the second perspective is concerned with uncertainty in the parameters of a model. The construction of the framework was achieved by focusing on five specific objectives: (i) to describe the complexity of how uncertainty arises in environmental health impact assessment and classify the uncertainty to be amenable for quantitative modelling;(ii) to critically appraise the strengths and limitations of current methods used to handle the uncertainty in environmental health impact assessment; (iii) to develop a novel quantitative framework for quantifying uncertainty from the conceptual and analytical perspectives; (iv) to formulate two detailed case-study examples on health impact assessment of indoor housing interventions; (v) to apply the framework to the two case-studies. After critiquing the uncertainty quantification methods that are currently applied in environmental health impact assessment, the thesis develops the framework for quantifying uncertainty, starting with the conceptual uncertainty (uncertainty associated with the framing assumptions or formulation of the model), then quantifying the analytical uncertainty (uncertainty associated with the input parameters and outputs of the model). The first case-study was concerned with the health impact assessment of improving housing insulation. Using fuzzy cognitive maps, the thesis identifies key indoor factors and their

pathways highly sensitive to the framing assumptions of the health impact assessment. The second case-study was concerned with estimating the uncertainty in the health burdens in England, associated with three ventilation exposure scenarios using fuzzy sets and interval analysis. The thesis presents a wider uncertainty framework as a first step forward in quantifying conceptual and analytical uncertainty in environmental health impact assessment when dealing with limited information.

.

Abbreviations

HIA	Health impact assessment
EHIA	Environmental health impact assessment
MC	Monte Carlo
BMA	Bayesian model averaging
MCMC	Markov chain Monte Carlo
UQ	Uncertainty quantification
DALY	Disability-adjusted life year
MP	Mathematical programming
CM	Cognitive maps
FCM	Fuzzy cognitive maps
BN	Bayesian networks
PAF	Population attributable fraction
RR	Relative risk
ACH	Air changes per hour
AMB	Annual morbidity burden

Table of Contents

List of Tables

List of Figures

List of Appendices

1. Introduction

1.1. Background

Uncertainty is central to every human health impact assessment exercise. It can arise due to our lack of understanding or knowledge of the "system" comprising the affected population and its surrounding environment. Uncertainty can also arise from random variations due to the stochastic nature of most real-life situations. Uncertainty is often broadly classified in the literature by two dimensions: its *nature* and its *location.* [1-3] The nature of uncertainty relates to its underlying causes in the assessment. Walker et al. (2003) used the term "variability uncertainty" to describe the nature of uncertainty due to a random process, and the term "epistemic uncertainty" to describe the nature of uncertainty due to incomplete information or impartial understanding of the system's underlying processes. Some authors also categorised uncertainty in terms of its *location.* [2-4] The location of uncertainty relates to where it occurs within the assessment model, for example in the parameters, the model structure or the input data.

Various qualitative frameworks used to classify and define uncertainty have been combined to report uncertainty more explicitly. [2, 5]Questions have been raised, such as "when" and "how" to deal with uncertainty and whether or not uncertainty needs to be reduced. Also, debates have arisen on how best to tackle the greatest source of uncertainty in the assessment, including the order in which to address uncertainties, given their magnitude. [6] Some authors have adopted a broader definition of

uncertainty in terms of its *location* in an attempt to "include" rather than "exclude" all sources of uncertainty. [7]

Concepts relating to uncertainty in risk assessment are derived from existing theoretical frameworks. [2-4] Knol et al. (2009) offered a typology of uncertainty to help structure, assess and reduce the uncertainties. One author reviewed and adapted uncertainty concepts from other authors. [3] Of particular interest is the concept of "contextual uncertainty", defined as the type of uncertainty that arises in choosing the boundaries of the system or defining the scope of the assessment. Knol et al. (2009) argue that any assessment outcome can be sensitive to the definitions of the system boundaries. They concluded that there is no single approach to deal with "contextual uncertainty", but rather a general guideline should be provided to consistently report the chosen boundaries and definitions. One example of a general set of guidelines based on expert judgement is the Dutch National Agency for Public Health and The Environment, and the Netherland Environmental Agency (RIVM/MNP) guidance for uncertainty. [7] The RIVM/MNP guidance provides a systematic approach to document and communicate uncertainty, starting with problem framing. However, one limitation of approaches, such as that of RIVM/MNP, is that they rely heavily on qualitative assessment and expert judgement. Nevertheless their idea of assessment outcomes being highly sensitive to the choice of system boundaries or defining the scope of the assessment overlaps with the focus of this research, which is to provide a quantitative framework to deal with uncertainty in the field of environmental health impact assessment.

In addition, Briggs et al. (2009) provided a framework to address the need to report uncertainty more systematically in health risk and impact assessments. The authors classified uncertainty in terms of its *location* in three stages. One of the stages is "conceptualisation", which is the type of uncertainty associated with the "framing of the problem" or formulation of the model. Briggs et al. (2009) argue that the best strategy to protect against the complications of conceptual uncertainty is by sharing the assumptions made with others, and by using more participatory approaches at the problem formulation stage. Briggs et al. (2009) are concerned about the lack of attention paid to uncertainty relating to the framing assumptions made in the assessment model, such as problem framing and boundary definition (this overlaps with one of the concerns of the current research project). Problem framing relates to uncertainty as to how to conceptualise the issues at hand in the assessment.

Walker et al. (2003) synthesised numerous contributions from other fields in order to provide an interdisciplinary framework for dealing with uncertainty. In attempting to synthesise and characterise all different types of uncertainty in the literature, the authors found that many "uncertainty experts" from other disciplines agreed on one aspect: that of distinguishing between the view of uncertainty held by the modeller and that of the decision-maker. There are many other views of uncertainty held by other stakeholders such as campaigning non-governmental organisations, citizens groups, managers or owners of polluting industries. However, the focus of this thesis is primarily on the modeller's view of uncertainty conducting an environmental health impact assessment.

1.2. Aims and Objectives

The overall aim of this research is to determine how best to deal with quantitative measures of uncertainty, in a more explicit and systematic way than current best practice, in the context of environmental health impact assessment. This will be achieved by focusing on five specific objectives, to:

1. Describe the complexity of how uncertainty arises in environmental health impact assessment and classify the uncertainty to be amenable for quantitative modelling.
2. Critically appraise the strengths and limitations of current methods used to handle the uncertainty in environmental health impact assessment.
3. Develop a novel quantitative framework as a step forward for quantifying the uncertainty in environmental health impact assessment
4. Formulate two detailed case-study examples on health impact assessment of indoor housing interventions.
5. Apply the framework to the two case-studies

1.3. Thesis structure

The remaining chapters of the thesis are structured as follows.

The **second** chapter describes the quantitative framework of the thesis. The framework is derived from different sets of concepts from other disciplines based on a literature review on uncertainty. These have been adapted for environmental health impact assessment. The chapter also defines the key uncertainty "perspectives" considered in the thesis.

The **third** chapter comprises a research paper which provides a systematic review of uncertainty quantification methods applied in environmental health impact assessment. It presents a critical appraisal of the literature describing the strengths and limitations of current methods used to deal with uncertainty.

The **fourth** chapter comprises a second research paper based on a case study example, which contains a framework, alongside a novel method detailing five main steps to deal with uncertainty associated with framing assumptions (conceptual uncertainty). In addition, the chapter provides a supplementary material of the method proposed in the case study to deal with uncertainty in a non-probabilistic domain. It will be argued that it is necessary to assess the sensitivity of the assessment to the framing assumptions (i.e. the mapping of the assumed causal pathways or structure to health outcomes), prior to conducting a detailed quantitative HIA of an environmental intervention. The framework was applied to determine the key pathways which have the most influence on health. This was done by analysing the causal map, associated with the framing assumptions, which links the potential effect of an environmental intervention of improving housing insulation to health outcomes.

The **fifth** chapter comprises the third research paper. It provides a quantitative framework to deal with the uncertainty associated with the parameters of a HIA model in a non-probabilistic domain (analytical uncertainty). It highlights the processes in the quantification of an HIA followed and it describes the potential health impacts and uncertainty in a case-study example of housing ventilation

exposure scenarios for England. Supplementary materials on the method proposed to deal with uncertainty in the case study are also provided.

The **sixth** chapter provides an overview of the findings of the thesis and main contributions, drawing results from all the research papers. The chapter acknowledges the limitations of the thesis and it identifies areas for further research. The conclusion of the chapter highlights the implications for applied researchers and policy makers.

1.1. Overall contribution of the research

The contribution of this PhD research is to expand the way uncertainty is taken into account in current environmental health impact assessment practice, particularly in relation to the handling of uncertainty when modelling environmental interventions or potential exposure scenarios. New sets of tools are introduced across two uncertainty "perspectives" (conceptual and analytical) in a non-probabilistic domain. The thesis identifies and handles different sources of uncertainty in relation to a health impact assessment of indoor housing and health. As a contribution to current consideration of uncertainty, this thesis aims to move debates from a narrow definition to a broader perspective of uncertainty. This broader perspective is required to define a wider assessment of impacts in environmental health impact assessment. An assumed causal structure can be quantified within the wider issues surrounding housing and health so that uncertainty can be dealt with in an explicit way. It is worth noting that a decision-maker might not be interested only in the analytical sources of uncertainties, but more fundamentally in the conceptual sources associated with the framing assumptions used to define the assessment of impacts

Providing an analytical framework that extends beyond handling the usual uncertainties, by incorporating conceptual sources of uncertainties in an explicit and systematic way can be an initial step forward in quantitative HIA. The added value of this research is in the attempt to include, rather than exclude, the framing assumptions quantitatively or semi-quantitatively in the appraisal of uncertainty. This can be used for decision support and could be a contribution to current environmental health impact assessment modelling practice. In addition, this research attempts to deal with analytical sources of uncertainty in a non-probabilistic domain. Under the proposed framework, the thesis provides a conceptual framework to deal explicitly with conceptual and analytical sources of uncertainty in a non-probabilistic domain.

1.2. Contribution of the candidate to the thesis

The candidate conceptualised the ideas for all the research papers included in the thesis and conducted most of the investigations himself, hence he is the first author of the resulting papers. Research paper 1 was designed by the candidate. He conducted the systematic review, synthesised the strengths and limitations of each method, discussed the findings and drafted the manuscript. Zaid Chalabi and Anna Foss reviewed the methods, interpretation and discussion. Tazio Vanni contributed to the section on Bayesian methods. This paper has been published in the *International Journal of Environmental Health Research.*

In research paper 2, the mathematical method was conceived and adapted by the candidate. He developed the model and the case study example through a literature

review on the relationship between housing insulation and health. The candidate designed the analytical framework, discussed the results and drafted the manuscript. Zaid Chalabi provided advice on the main manuscript and suggestions for the development of the mathematical details of the appendix. Anna Foss also reviewed this paper. The candidate wrote the code to implement the mathematical details of the method. This paper has been published by the journal *Environment International*.

The candidate led on the conception and conceptualisation of the methods for research paper 3 in collaboration with his supervisor, Zaid Chalabi. The candidate developed the health impact model of housing ventilation exposure scenarios and conducted the systematic review alongside a meta-analysis in relation to housing ventilation and health. The candidate was responsible for designing, conducting and interpreting the analysis. Also, the candidate wrote the code to implement the mathematical details of the method. Zaid Chalabi and Anna Foss reviewed this paper. This paper has been published by the journal *Environment International*. In addition a corrigendum of the paper has been published by the same journal.

Further details on the specific contributions of the candidate are shown in the cover page of each research paper.

2. Quantitative framework

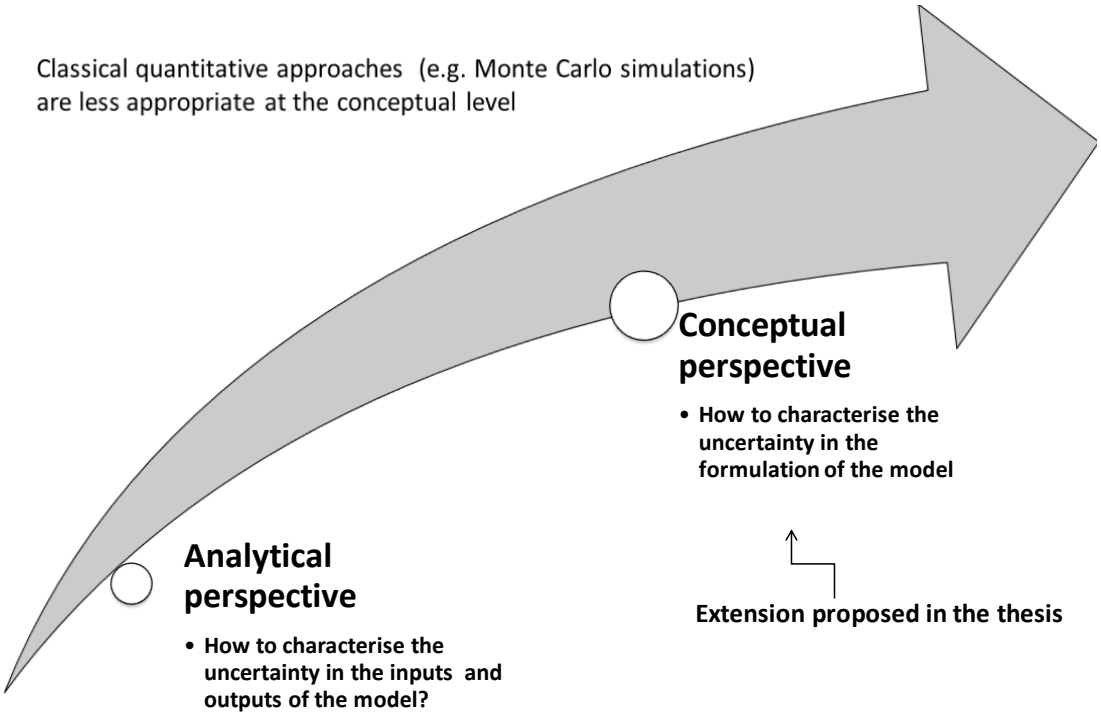
2.1. Perspectives of uncertainty

The need for a shared and coherent understanding of uncertainty has been deemed important in health risk and environmental health impact assessment. In this chapter, some of the concepts for dealing with uncertainty in terms of its *location* [2-4] are combined to form two main "perspectives" of uncertainty. These two main perspective are "conceptual`' and "analytical" and are incorporated in the thesis framework (as shown in Figure 1). The conceptual perspective combines Briggs et al.'s (2009) own definition of "conceptual uncertainty" and Knol et al.'s (2009) definition of "contextual uncertainty". The analytical perspective derives from the three sets of definitions of analytical uncertainty and the conventional modeller's views of uncertainty, such as those associated with the analytical outcome of the models. [2-4]

The assumption of the proposed framework in the thesis is that uncertainty is often difficult to examine from all possible perspectives simultaneously. Analysts often see uncertainty from different viewpoints. As a result, two main perspectives are chosen as a way of narrowing the examination of uncertainty so that some aspects of uncertainty can be emphasised at one stage while others are dealt with at another. The proposed framework is provided in this PhD, to handle the uncertainty in an explicit and systematic way, so that the actions when conducting a quantitative HIA

are clear, particularly when attempting to quantify and propagate the uncertainty in the assessment.

Figure 1: The different perspectives of uncertainty



In relation to uncertainty in HIA, it is important to distinguish between a "conceptual learning process" and an "analytical learning process". [8] A conceptual learning process can be defined as the process of conceptualising the model by refining the issues at hand, the questions, or refocusing the problem in order to gain new insights; it is thus not a linear process. In other words, a conceptual learning process does not progress from one stage to another with a clear starting point and an ending point. The "analytical learning process" on the other hand is defined as the process of finding solutions for a particular problem in the assessment so that the uncertainty can be looked at in relation to its defined purpose. It is also important to note that at the "conceptual level", the uncertainty is often characterised by a lack of knowledge.

As we move down to the “analytical level”, the uncertainty can often be characterised by variability. Lack of knowledge can also be present as part of the “analytical perspective” to some degree, but is less pervasive than at the conceptual level. The same applies to variability, which can also be present to some degree at the conceptual level due to the stochastic nature of the impact of the environment on health.

2.1.1. Conceptual perspective

Conceptual uncertainty deals with the uncertainty associated with the formulation of the problem to be investigated. It is more concerned with the characteristics of the system that is being modelled, which can involve uncertainty in the formulation of the assumed causal pathways. Conceptual uncertainty is not concerned with the analytical aspect of the model, such as the type and statistical associations between variables, but rather on defining and formulating the assessment. Classical quantitative methods such as Monte Carlo (MC) simulations are less appropriate at this level since the uncertainty resides at a more fundamental conceptual level than in the model parametrisation. The need to enhance the problem and model formulation phase of any HIA makes it worth closely investigating conceptual sources of uncertainty. At the start of any assessment exercise, it is important to identify the framing assumption that defines the assessment as part of the model formulation. Including this conceptual perspective of uncertainty in the framework will help to move away from a narrow focus and definition of uncertainty towards a systems-based approach to uncertainty, focusing on the framing aspects of the assessment.

2.1.2. Analytical perspective

The analytical uncertainty perspective views uncertainty from a modeller's point of view, where the main emphasis is on quantifying and characterising the uncertainty in the input parameters and the model outputs. Most of the methods for dealing with uncertainty in environmental health impact assessment are based on this analytical model-based view. Current methods in environmental health impact assessment, as identified in research paper 1, deal with parametric uncertainty in environmental health impact assessment. Analytical perspective specifically refers to uncertainty propagation methods used for estimating and quantitatively propagating the impact of errors or biases in the model outputs. In other words, the analytical perspective quantifies and propagates the impact of parametric sources of uncertainty in the outcome of the assessment. Methods are used under the analytical perspective to perform sensitivity analyses or uncertainty analysis to explore parametric uncertainty.

2.2. Quantitative tools proposed in the framework

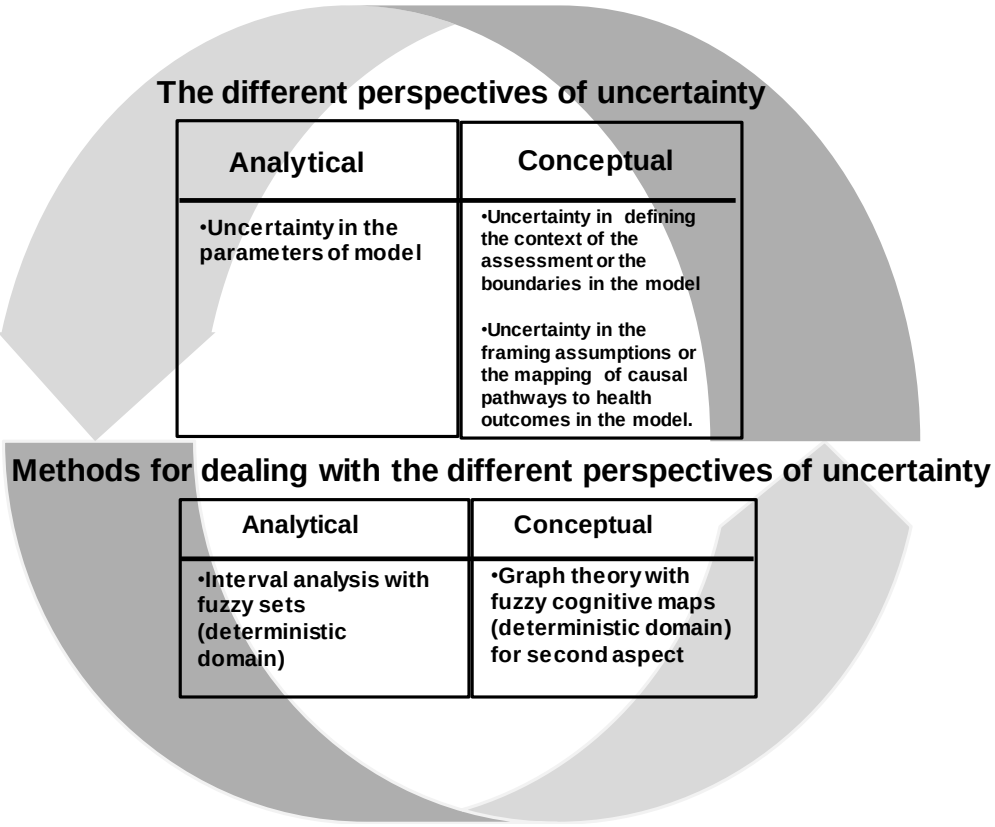
In this framework, a set of tools is introduced across the two uncertainty perspectives operating in a non-probabilistic domain. The two uncertainty perspectives characterise the uncertainty in terms of its *location.* The set of tools handles uncertainty in terms of its *nature.* They assume lack of knowledge rather than random variations in the propagation of uncertainty. One of the reasons for choosing to work in a deterministic domain is that constructing some probabilistic models may require unrealistically detailed data information (as described in research paper 1 in the following chapter). Another reason is that suitable assumptions (given the limited

information) regarding the statistics of the variability of an uncertain parameter cannot be made. Limited quantitative information is often encountered when modelling environmental impacts on health. [9, 10, 11] In developing this uncertainty framework, insufficient data for probabilistic characterisation is assumed such as specific assumptions about the correlations of variables and random variations in probability distributions. Techniques that other disciplines use as alternatives to probabilistic methods are proposed under this framework. Fuzzy cognitive map techniques, as found in graph theory, are proposed to deal with conceptual uncertainty in HIA. Other analytical methods, not commonly applied in environmental health such as fuzzy set theory, using interval analysis, are proposed under the analytical perspective of the framework. A summary of the tools proposed for the framework for dealing with the two perspectives of uncertainty is shown in (Figure 2). It is worth noting that conceptual uncertainty perspective is sub-divided by two aspects: (i) uncertainty in defining the context of the assessment or the boundaries of the model to deal with contextual issues of time and space, and (ii) the uncertainty in the framing assumptions to deal with aspect of problem framing relating to the mapping of the causal pathways to health outcomes. The proposed tool to deal with conceptual uncertainty addresses the second aspect.

Fuzzy cognitive maps are proposed to deal with conceptual uncertainty in relation to the formulation of a health impact model, particularly in the framing or the mapping of the assumed causal pathways to health outcomes. A fuzzy cognitive map is a causal diagram that can be used for the interpretation of causal assumptions. [12] It can map the interpreted causal pathways explicitly in the health impact model. The framing assumptions of a model can be conceptualised by constructing a graphical

model for representing perceptions about the assumed pathways of the problem to be formulated. By using fuzzy cognitive maps, quantitative and qualitative information can be combined into a single diagram. [13] Some authors have suggested coding the graphical representation of cognitive maps into matrices using graph theoretical indices. [14, 15]Fuzzy cognitive maps can ground a causal network structure on a mathematical foundation by using graph theoretic indices such as the centrality index. The added value of representing a structure in a mathematical foundation is that sources of uncertainty at the model formulation level can be explored in the assessment, particularly in the mapping of the causal pathways. The fuzzy cognitive map method has been shown, in other areas of research, to have advantages over other methods used for conceptual or causal modelling such as Bayesian networks. This is particularly useful in situations where data is limited, and the elicitation of probabilities has proven to be difficult. [14]

Figure 2: The different perspectives of uncertainty and proposed methods



As an analytical perspective tool, fuzzy sets using interval analysis can deal with parametric uncertainty in a deterministic domain. Interval analysis using fuzzy sets is used to compute interval bounds in the outputs of a model, assuming input values are fuzzy or imprecise, and bounded in intervals. By using fuzzy sets, an input value can be assumed to lie within an imprecise interval (fuzzy set). [16-19] The uncertainty in the parameter values expressed as a fuzzy interval is propagated through the model using interval arithmetic. This method allows for imprecise definitions or errors in the measurements to be incorporated in the propagation of uncertainty. Interval analysis with fuzzy sets does not depend on sampling, statistical variance or probability elicitation. In addition, this method can represent uncertainty in relation to "lack or limitation in knowledge", [20] which is one of the reasons it is included in the framework to handle parametric uncertainty. When there is limited information about parameter values, lack of knowledge can be represented in the form of fuzzy interval and “nothing else” is assumed. This is done without the need to assume uniform distributions and therefore assert that each value can be “equally likely to occur”. [21] The scope of the proposed methods and how to what extent they address the uncertainty are discussed in the supplementary material of chapter 3. Further mathematical details and definitions of the proposed methods and techniques are shown in supplementary material of chapter 4 and chapter 5 of the thesis.

3. Uncertainty quantification methods in an environmental health impact assessment

3.1. Preamble to research paper 1 – Literature review

The conceptual review provides a background to the use of uncertainty frameworks and indicates the need for a coherent understanding of uncertainty quantification methods. Uncertainty can be a substantial element in environmental health impact models due to the complex associations between environmental exposures and health outcomes; however more research is needed to quantify its impact. Research paper 1 aims to highlight the research gaps in the literature by providing a systematic review and critically appraising current methods used to quantify uncertainty in an environmental health impact assessment (HIA). Research paper 1 considers current quantitative methods and tools for characterising and handling uncertainty and the ways in which they are applied in environmental health impact assessment. It provides a detailed discussion of the strengths and limitations of these methods and tools and also looks at the theoretical frameworks that have been employed in the past to deal with uncertainty. The findings from research paper 1 provide a justification for the methodological approaches chosen for the framework, which are described in detail in chapters 4 and 5 of the thesis.

3.2. Research paper 1

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Marco Mesa Frias
Principal Supervisor	Zaid Chalabi
Thesis Title	Modelling Uncertainty in Environmental Health Impact Assessment

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	International Journal of Environmental Health Research		
When was the work published?	Apri 2012		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	See attached cover sheet

Student Signature: _____ **Date:** 5/10/2015

Supervisor Signature: _____ **Date:** 5/10/2015

Improving health worldwide **www.lshtm.ac.uk**

Research paper 1

Uncertainty in environmental health impact assessment: Quantitative methods and perspectives

Marco MESA-FRIAS[1]; Zaid CHALABI[1]; Tazio VANNI[2]; Anna M. FOSS[2]

[1]Department of Social and Environmental Health Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.
[2]Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.

Status: Published in International Journal of Environmental Health Research, doi: 10.1080/09603123.2012.678002

Contributions: The candidate wrote the first draft of the manuscript. He managed each round of comments and suggestions from co-authors in collaboration with Zaid Chalabi. The candidate conducted the systematic literature review and designed the research question. Anna Foss contributed to the review providing comments and suggestions. Tazio Vanni also contributed and provided comments in the interpretation of the results. All authors read and approved the final draft prior to journal submission and inclusion in the dissertation.

The candidate The supervisor

Abstract

Environmental health impact assessment models are subjected to great uncertainty due to the complex associations between environmental exposures and health. Quantifying the impact of uncertainty is important if the models are used to support health policy decisions. We conducted a systematic review to identify and appraise current methods used to quantify the uncertainty in environmental health impact assessment. In the 19 studies meeting the inclusion criteria, several methods were identified. These were grouped into random sampling methods, second-order probability methods, Bayesian methods, fuzzy sets and deterministic sensitivity analysis methods. All 19 studies addressed the uncertainty in the parameter values but only 5 of the studies also addressed the uncertainty in the structure of the models. None of the articles reviewed considered conceptual sources of uncertainty associated with the framing assumptions or the conceptualisation of the model. Future research should attempt to broaden the way uncertainty is taken into account in environmental health impact assessments.

Keywords: Uncertainty; Environmental health models; Quantitative health impact assessment; Models; Environmental exposures

Introduction

Uncertainty is present in all environment health impact assessments (EHIA). It arises mostly from our lack of understanding of the associations between environmental exposures and health, but can also arise from random variations in the associations between environmental exposures and health outcomes. Uncertainty is often classified in terms of its *nature* and its *location.* [2-4] The nature of uncertainty is concerned with identifying the underlying causes of uncertainty. [2] It can be grouped in two broad types: "lack of knowledge" and "natural variability". [3] Lack of knowledge is the type of uncertainty which can be reduced with further research. Natural variability, on the other hand, describes a type of irreducible uncertainty, which is inherent in the stochastic nature of most environmental variables and health outcomes. The location of uncertainty is concerned with where the uncertainty occurs in the model, such as in the input data, parameterisation or model formulation. [2-4] According to Briggs and colleagues, [4] uncertainty can take place at different stages in the assessment, in terms of its location, such as in: the "conceptualisation" and the "analysis". The authors used the term "conceptual uncertainty" to refer to sources of uncertainty that arise at the stage of the framing of the environmental health problem, in other words, when defining the context of the assessment in the EHIA model during model formulation. They used the term "analytical uncertainty" to refer to the statistical uncertainty in the parameters and input data of the model. Another stage in which uncertainty can take place is "decision stage". According to Walker and colleagues, [3] the decision stage is concerned with how to value the outcome of the assessment model from a decision-making perspective, particular when there is uncertainty in the research evidence. Walker and colleagues synthetised numerous contributions from other fields in order

to provide an interdisciplinary framework for dealing with uncertainty. As the authors attempted to synthetise and characterise all different types of uncertainty found in the literature, they discovered the fact that many "uncertainty experts" from other disciplines agreed on one aspect: in distinguishing between the modeller's view of uncertainty and the policy /or decision-maker's view of uncertainty.

In general, EHIA models suffer from large uncertainty due to complex associations between environmental exposures and health. Quantifying the impact of uncertainty on the EHIA results is particularly important if the models are used to support health policy decisions. Quantitative information linking exposures to health outcomes might be of poor quality or not even existent. Some analysts might be reluctant to quantify the impact of uncertainty on their model results particularly when their model is based on sparse data, as this can lead to large uncertainties surrounding the central estimates of the models. The larger the uncertainty in the EHIA model outputs, the more it can undermine the credibility of the EHIA model and the comparability of results with other models. Therefore, it is necessary to review and critically appraise the methods used to quantify the uncertainty in EHIA models.

Uncertainty quantification methods

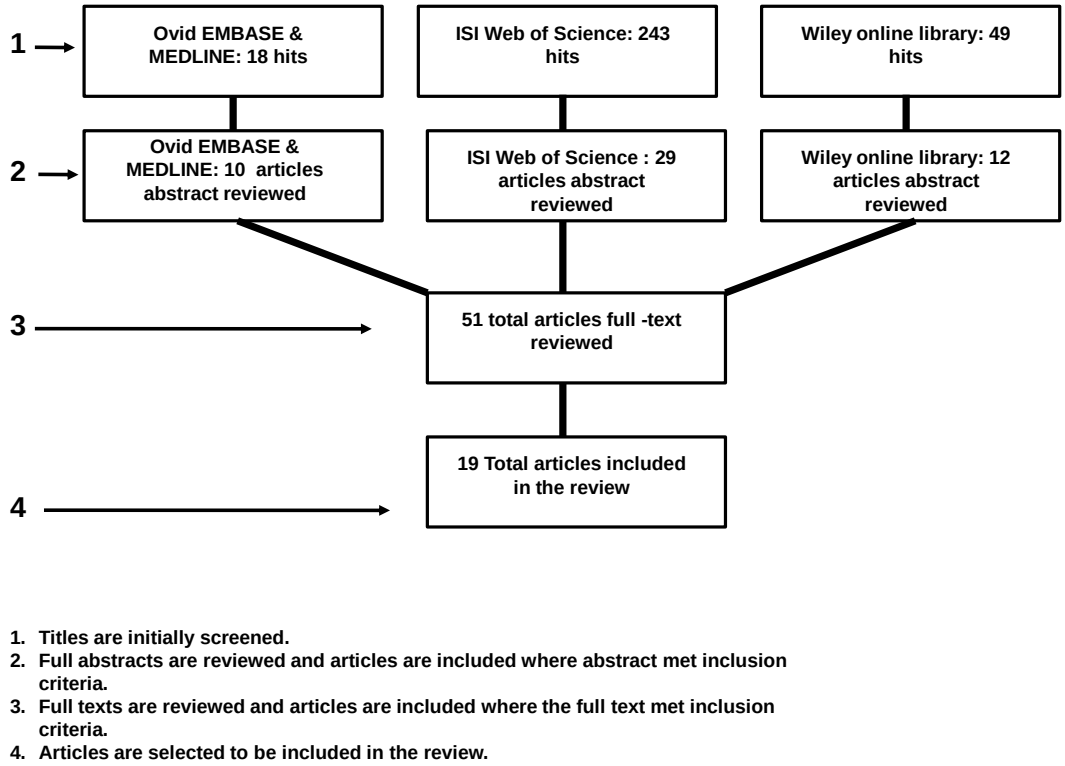
Uncertainty quantification (UQ) methods are employed to assess the impact of uncertainty on the EHIA model estimates or projections. Quantifying model uncertainty often involves "uncertainty analysis" and "uncertainty propagation". Uncertainty analysis evaluates the extent to which the uncertainty can influence the output of the assessment model. [22] The evaluation of uncertainty is conducted according to "its location" such as the input parameters or model structure.

Uncertainty analysis methods can also be used as a form of global sensitivity analysis which aims to identify the input parameters that most influence the output. Uncertainty analysis should not only deal with the uncertainty in the input parameters and their influence on the output (as sensitivity analysis) but also with the uncertainty associated with the formulation of the model. The term "uncertainty propagation", on the other hand, describes methods which can communicate (propagate) the uncertainty from its various locations to define the uncertainty of the overall output of the model. [23] The way the uncertainty is propagated depends on "its nature", whether it is due to unavoidable random variations, lack of knowledge, or both. In general, methods used to quantify uncertainty can be commonly placed within the broad categories of random sampling methods and non-probabilistic methods. In the following section, we review UQ methods used in EHIA.

Systematic review

UQ approaches for handling uncertainty in EHIA were reviewed. A literature search was conducted using: Ovid MEDLINE, EMBASE, ISI Web of Science and the Wiley online library databases. Free-text terms, combined using Boolean operators, were used in the search. Free-text was used rather than MeSH terms in order to identify non-indexed and incorrectly indexed records which would have been missed if MeSH terms had been used instead. The search was conducted using keywords: "uncertainty", "health impact assessment", "quantitative", "quantification", "model", "models", "modelling", "modeling", "modelled" , "modeled", "prediction", "simulation", "projection" and "software". Details of the search strategy are presented in Supplementary file 1. Results of the search are shown in Figure 3

Figure 3 :Results of the literature search for methods to deal with uncertainty in environmental health impact assessment



Full-texts of studies were retrieved if they seemed to be of potential interest following a screen review of their titles and abstracts. These full-texts were then screened using the inclusion and exclusion criteria described in Table 1. Limits were placed on the search so that it was confined to English language articles and dates of publication from January 2000 up to January 2011 - with the exception of the ISIS Web of Science database - where the search was confined over the last five years of publication (January 2005 to January 2011). This later strategy was used to keep the search in a manageable number of references, due to resource constraints. We selected articles in peer-reviewed academic journals focusing on quantitative modelling in relation to environmental and human-health related impact assessment studies using uncertainty quantification methods. We excluded articles in other

sources of literatures (e.g. non-peer reviewed reports, chapters in books, editorials) and also excluded those focusing on qualitative studies and non-human health related modelling studies, and environmental health studies without explicit uncertainty quantification methods.

Table1: Search inclusion and exclusion criteria for uncertainty quantification methods in environmental health impact assessment.

	Inclusion criteria	**Exclusion criteria**
Sources	Peer-reviewed journal article.	Non Academic Articles.
Article type	Original comparative research. Review article.	Reports, chapters, news article. Editorial.
Study type	Environmental human health related study within health impact assessment and quantitative risk assessment or related assessments as part of an integrated health impact assessment (quantitative or modelling-based study). Quantitative characterisation of uncertainty.	Environmental non-human health related. Other non-quantitative assessment (qualitative based study). Qualitative uncertainty characterisation.
Language	English.	Other languages.

Results from systematic review

Of the 51 articles identified by the search strategy outlined above, 19 met the inclusion criteria. [24-42] Most were excluded since they did not have an explicit quantitative uncertainty characterisation component. Other articles were excluded due to duplication of the same published studies or for not having a direct environmental health application. All papers dealt with the uncertainty in the parameters of the models (as shown in Table 2), while only 5 dealt also with the uncertainty in the model structure. [27, 35, 37, 40, 42] Individual studies identified in the

literature which share a common methodological ground were combined and summarised into the following categories: random sampling methods, second-order probability methods, Bayesian methods, fuzzy sets and deterministic sensitivity analysis methods.

Random sampling methods

Random sampling methods involve assigning distributions to parameters and repeatedly taking random samples from the assumed distributions of uncertain input parameters. In these methods, an EHIA model is run many times, using the sampled values and a distribution of the outputs is constructed. In general, random sampling methods, such as Monte Carlo (MC) techniques, [43] are used for uncertainty analysis and uncertainty propagation. Standard MC methods perform a large number of simulations using different sets of input parameters at each iteration step to generate the model outputs at those sample points. [43] The results can also be used to analyse the contribution of uncertainty in an input parameter to the uncertainty in the total output of the model. [38]

Table 2: Studies identified and included in the review

Title and First author	Type (s) and location (s) of uncertainty	Methods to characterise the uncertainty
A model for probabilistic health impact assessment of exposure to food chemicals [24]	Describes the uncertainty with regards to a full range of input parameters values to investigate the total uncertainty in the model outputs. It distinguishes between variability and uncertainty.	Random sampling-based methods by assuming distributions in the input parameters using an algorithm that provides and combines MC distributions. Probabilistic estimates in the parameters where considered in the conclusion of the model.
Health impact assessment of particulate pollution in Tallinn using fine spatial resolution and modeling techniques [25]	Uncertainty about the correctness of the model in exposure assessment and does not distinguish between the types of uncertainty.	Deterministic sensitivity analysis was performed in some input parameters. Model validation was performed with the PM2.5 and PM10 air pollution modelled levels and compared with air quality monitoring data. Uncertainty in the health impact estimates were considered in the conclusion of the model.
Decision support system for the evaluation of urban air pollution control options: Application for particulate pollution in Thessaloniki, Greece [26]	Uncertainty in the input data and parameters. Mainly incorporating both variability and uncertainty in a non-probabilistic approach.	Deterministic sensitivity analysis method through linear programming formulation (optimization) to perform sensitivity analysis. Result of parametric sensitivity analysis was not particular important for the conclusion of the model.
Parameter and model uncertainty in a life-table model for fine particles (PM2.5): a statistical modeling study [27]	Explores the uncertainty in all the input parameters and their effects on total model output. Assumes distributions and treats variability and uncertainty in all input parameters without making an explicit methodological distinction.	Both parameter and model uncertainty were propagated using MC simulation, and uncertainty analysis was conducted between all inputs and model results. The results of parametric uncertainty and the potential plausibility of the model were considered in the conclusion.
Menu Labeling as a Potential Strategy for Combating the Obesity Epidemic: A Health Impact Assessment [28]	Describes the uncertainty in the variability of the input parameter data used in a simulation model.	Sensitivity analysis was conducted in some parameters used in the simulation model. The result of the sensitivity analysis was not particular important for the conclusion of the model.

Towards health impact assessment of drinking-water privatization —the example of waterborne carcinogens in North Rhine-Westphalia (Germany) [29]	Input parameters are modelled as variability in the exposure assessment. In the dose-response assessment, the uncertainty was treated as lack of knowledge rather than variability.	Random sampling methods were used using probabilistic estimates to assume distributions in the exposure modelling. For dose response assessment, assumptions were made using potency factors with no threshold and with no uncertainty quantification method used. Probabilistic estimates in the parameters were considered in the conclusion of the model.
Quantitative health impact assessment of transport policies – two simulations related to speed limit reduction and traffic re-allocation in the Netherlands [30]	Variability in the input parameters in exposure levels and uncertainty in the outcome. Other sources of uncertainty were described but not analysed, such as the uncertainty in the exposure-effects relationships.	MC uncertainty analysis in the input parameters. The assumptions of the model and the results of the MC uncertainty analysis were particular important in the conclusion of the model.
Quantitative risk assessment of CO_2 transport by pipelines—A review of uncertainties and their impacts [31]	Uncertainty in the input parameters in the exposure-assessment simulation model.	MC sensitivity analysis is conducted on parameters of simulation models: release, dispersion and impact models. The results of parametric sensitivity analysis was particular important in the conclusion of the model.
Analysis of coupled model uncertainties in source-to-dose modeling of human exposures to ambient air pollution: A PM2.5 case study [32]	Uncertainty in parameters and structure in the model. Separates uncertainty and variability in exposure-assessment and dose-response modelling.	Evaluates parameters uncertainty using random sampling methods by assigning probability distributions and structural model uncertainty is evaluated by comparing different models with measurement air quality monitoring data. Overall uncertainties in both parameters and model structure were important in the conclusion.

Second-Order Modeling of Variability and Uncertainty in Microbial Hazard Characterization [33]	Uncertainty in the parameters by separating variability and uncertainty in the inputs of the model.	Second order probability methods using MC simulation to separate variability and uncertainty in the parameters. Bootstrap simulation was used to estimate sampling errors due to uncertainty in the limited amount of input data. Parametric uncertainty was particular important in the conclusion of the model.
Impact and uncertainty of a traffic management intervention: population exposure to polycyclic aromatic hydrocarbons [34]	Explores the uncertainty in the parameters of the exposure assessment model.	Random sampling method based on MC analysis to characterise the uncertainty, including uncertainty propagation in the output estimates. The result of probabilistic exposure estimates were important in the conclusion of the model
A Bayesian hierarchical model for urban air quality prediction under uncertainty [35]	Describes the uncertainty in the input data, parameters, and model structure and model outputs. Deals with both variability and uncertainty though a complete Bayesian Hierarchical model and using multivariate statistical methods to obtain priors.	Bayesian methods were conducted for uncertainty characterisation. In particular MCMC sampling method was conducted for uncertainty propagation to obtain posterior probability distributions. Both parametric and structural uncertainties were considered but not particular important in the conclusion of the model
An integrated fuzzy-stochastic modeling approach for assessing health-impact risk from air pollution [36]	Evaluates uncertainty in the parameters and model output using an integrated approach distinguishing variability as parameters that can be expressed as probability distributions and uncertainty as non-probabilistic.	Fuzzy set theory to model the uncertainty that could not be expressed as probability distributions and MC uncertainty propagation techniques for probabilistic parameters. The result of the uncertainty in the input parameters was important in the conclusion of the model.

Model and input uncertainty in multi-media fate modeling: Benzo[a]pyrene concentrations in Europe [37]	Uncertainty in the input parameters and uncertainty in the structure of the model.	Probabilistic modelling assuming distributions for the uncertainty in the input parameters and model uncertainty is dealt through a validation procedure comparison across models. Both sources of uncertainty were important in the conclusion of the model.
Uncertainty in health risks due to anthropogenic primary fine particulate matter from different source types in Finland [38]	Evaluates parameters uncertainty reflected on the overall model result.	MC simulation for the propagation of uncertainty in parameters and sensitivity analysis conducted between input parameters and model result. Probabilistic estimates in parametric dose-response relationships were important in the conclusion of the model.
Separation of uncertainty and interindividual variability in human exposure modeling [39]	Distinguishes between variability and uncertainty in the exposure assessment model. Uncertainty is evaluated in the parameters and model result.	Second-order MC uncertainty propagation and uncertainty analysis to assess the output parameter distribution by ranking between 100 simulated populations.
An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination [40]	Explores parameters uncertainty in the input parameter and structural model uncertainty. It distinguishes between variability and uncertainty.	Variability in the input parameters is modelled using MC simulation and uncertainty in the inputs is modelled using fuzzy-set theoretic approaches. Results of the uncertainty in input parameters were important in the conclusion of the model.
Probabilistic Framework for the Estimation of the Adult and Child Toxicokinetic Intraspecies Uncertainty Factors [41]	Describes parameters uncertainty accounting for variability and uncertainty in the input parameters.	MC sampling for variability in the input parameters and toxicokinetics uncertainty factors (index measure) it is used for uncertainty in the parameters. Probabilistic estimates in the input parameters were considered in the conclusion of the model.
Bootstrap-after-Bootstrap Model Averaging for Reducing Model Uncertainty in Model Selection for Air Pollution Mortality Studies [42]	Describes the uncertainty in the correctness and choice of models. Does not distinguish between uncertainty and variability.	BMA is conducted initially to yield a set of plausible weighted models, and then bootstrapping resampling is applied to weight each of the original plausible models. Results of both parametric and structural uncertainty were considered in the conclusion of the models.

MC, Monte Carlo; BMA, Bayesian model averaging; MCMC, Markov chain Monte Carlo; UQ, uncertainty quantification

One limitation of the standard MC technique is that it assumes that the distributions of the input parameters are known. In reality, the specific distributions of the input parameters are hardly known and the output distributions are sensitive to the selected input (prior) distributions. [44]Examples of MC methods applied to environmental health exposure and impact assessments include: modelling air pollution for fine particles, [27] transport intervention studies [30, 34] and exposure to food chemicals. [24]

Second-order probability

Second-order probability methods are facilitated by MC techniques (called second-order Monte Carlo methods). Second-order probability methods [45] attempt to distinguish between two types of uncertainty in the model parameters: "lack of knowledge" and "variability". For ease of illustration and without loss of generality, consider an EHIA model with one parameter only. In these methods, both types of uncertainty in the model parameter are propagated. This is done computationally using two loops: the outer loop propagates the "lack of knowledge" and the inner loop propagates "variability". "Lack of knowledge" defines the uncertainty in the parameters of the distribution of the model parameter (such as its mean or variance in the case of a normal distribution). The uncertainty in each of the parameters of the distribution can be expressed as a bounded interval. The inner loop propagates the "variability" in the model parameter, conditional on the distribution defined in the outer loop. Second-order MC simulation starts by selecting values of the parameters of the distribution uniformly between the lower and upper bounds in the outer loop and then fixes the distribution for the inner loop calculations. The variation in the inner loop sampling distribution represents the "variability", and the variation in the outer loop sampling distribution represents the "lack of knowledge". One limitation

of the second-order MC method is interpreting what constitutes "lack of knowledge" or "variability" in the input parameters of the model. Before the simulation is performed, the interpretation is subject to the judgement of the modeller. Potential difficulties can arise for the modeller in characterising the "lack of knowledge and "variability" in the model parameters. Examples of the application of second-order probability methods in environmental health include microbial hazard characterisation [33] and human exposure modelling of contaminants through different environmental media (air, food, soil and water). [39]

Bayesian methods

Bayesian model averaging (BMA) techniques [46] are used to handle model uncertainty by formulating alternative competing models supported by some statistical model averaging techniques. The predictions of multiple model results are combined and weighted using "information criteria". BMA applies a mixture of Bayesian computation, statistical model averaging approaches and likelihood measures. It typically involves using some information criterion-based techniques such as Bayesian Information Criterion (BIC) [47] and Akaike Information Criterion (AIC). [48] Posterior weights are assigned to the competing models reflecting their plausibility given the data, and model selection is used to reduce the uncertainty in the different model structures. BMA have been used in conjunction with bootstrapping methods. In bootstrapping methods, a single sample is taken from the parameter values of the model and used as the reference distribution from where to subsequently resample to estimate the original sampling distribution. BMA was applied to aid in the model selection in an air pollution mortality modelling study. [42]

The implementations of some Bayesian methods in the propagation and computation of uncertainty are usually supported by Markov Chain Monte Carlo (MCMC) techniques using algorithms for computing high dimensional joint distributions. [49] MCMC sampling constructs a Markov chain (discrete stochastic process) of correlated random samples with the main objective of finding the Bayesian posterior distribution of the input parameters. The rate of convergence of the MCMC algorithm is calculated using probability theory. MCMC techniques can handle the uncertainty in a large number of parameters. [21] The MCMC method has been applied in air quality modelling. [35] One potential limitation of applying Bayesian methods is the tendency to use subjectively assigned probabilities based on prior beliefs. Posterior distributions are dependent on prior beliefs of experts on the choice of prior distributions.

Fuzzy set theoretic methods

Fuzzy-based methods express the uncertainty in a non-probabilistic way via a fuzzy set. [20] A fuzzy set is defined by its elements and a membership function. A membership function measures the *degree* (between zero and unity) to which an element belongs to a set. In fuzzy set methods, membership functions are used to characterise the uncertainty in the input parameters of a model, particularly when there is insufficient information to estimate probability distributions or insufficient knowledge to define clear individual states or events. Fuzzy membership functions differ from probability distributions in one fundamental aspect. A probability function measures the "probability that an event takes place" by using a numerical probability distribution. On the other hand, a fuzzy membership function measures the "degree to which an event occurs", in other words, it measures the imprecise

nature of the definition of the event, not the probability that the event occurs. The uncertainty in a parameter that cannot be modelled using a probability distribution can be characterised instead as a "vague" parameter, representing some imprecise qualitative information on the parameter that cannot be expressed accurately. The use of fuzzy set methods can be limited (and should not replace probabilistic approaches) in circumstances when there is sufficient information and data availability to derive probability density functions. [18, 50] Examples of applications using fuzzy methods in EHIA include the human health risk assessment of groundwater contamination [40] and air pollution modelling. [36] In both examples, membership functions were constructed to map qualitative data (collected from questionnaires or guidelines) on the level of pollutant concentrations into fuzzy sets (e.g. "strict", "medium" and "loose").

Deterministic sensitivity analysis

Deterministic sensitivity analysis can be thought as a subset of uncertainty analysis. Deterministic sensitivity analysis consists in varying the values of model parameters systematically in order to explore the sensitivity of model result to changes in the parameters. In general, sensitivity analysis can be categorised into local sensitivity and global sensitivity. [51] Local sensitivity analysis evaluates the uncertainty on the model result in the vicinity of a fixed set of values of the parameters. Global sensitivity analysis evaluates the overall uncertainty with respect to the full range of values of model parameters. The use of sensitivity analysis is hampered when dealing with large number of input parameters because it would be difficult to summarise the results of the analysis in an informative manner. Deterministic sensitivity analysis does not allow modellers to evaluate the sensitivity of the model

parameters by taking into account their statistical likelihood. Compared to probabilistic-based methods, [52]deterministic sensitivity analysis can therefore be less favourable when dealing with large number of parameters. Examples of deterministic sensitivity analysis in EHIA include the health impact of air pollution, [25] risk management control strategies on air quality and health [26] and the health impacts of menu labelling on obesity. [28] The uncertainty was explored by varying the input model parameters and investigating their effect on health impacts.

Discussion

A couple of reviews have been conducted previously on health impact assessment modelling, [53, 54] but these reviews did not address uncertainty quantification (UQ) methods explicitly in the models. In this paper, we presented a systematic review of quantitative methods used to handle uncertainty in EHIA models.

Limitation and guidance on current methods

Appropriate methods for handling uncertainty from other disciplines may not have been identified since these were beyond the scope of this review. The emphasis of this review is on the relatively new established field of health impact assessment (HIA) focussing on its applications in environmental health. We limited our review to the academic literature and potential studies in the "grey" literature may have not been identified. Evidence from the methods identified in this review can be interpreted as illustrative of the existing methods used to deal with the uncertainty in EHIA.

Most random sampling methods, currently used in EHIA, use probability distributions (e.g. triangular, normal, lognormal distributions) to characterise the uncertainty in input parameters. Most sampling-based MC techniques are relatively simple in terms of formulating the computational steps and in handling of "high-dimensional" problems (i.e. models with many parameters). This is in contrast with other analytically-based methods used in engineering applications, such as differential analysis [55] which can be computationally demanding and difficult to implement when handling large numbers of input parameters. The differential analysis method requires calculating partial derivatives to estimate the uncertainty in the model outputs that results from the assigned input distributions in the model. Determining the partial derivatives can be difficult to implement. This could explain why MC techniques are more widely implemented over other analytically-based method (such as differential analysis) in EHIA models.

Most sampling-based MC techniques assume that there is sufficient data to help in defining the shapes of the distribution. Bayesian methods can address the issue of lack of data availability through the choice of prior, specified as non informative or uniform prior. The prior distributions can be elicited from experts. [56, 57] Most probabilistic techniques, with the exception of second-order MC methods, cannot fully distinguish between the two natures of uncertainty: the uncertainty due to lack of knowledge, and the uncertainty due to random variation (variability).

Probability theory is often regarded as insufficient to model the uncertainty associated with lack of knowledge. One argument against the axiomatic nature of probability is that probability theory uses some form of "equitable" probability as a

model for general uncertainty that cannot fully distinguish between a random process, imprecision or lack of knowledge. [58] In the context of health impacts or risk assessment, additional care should be taken in the treatment of ignorance as natural variability. For instance, when the modeller is not sure about the initial "shape of the distribution", that uncertainty should be simply regarded as "lack of knowledge" and therefore should not be treated implicitly as natural variability or a random process. [58] In circumstances characterised by lack of knowledge, non-probabilistic approaches such as fuzzy sets can be useful alternatives in handling the uncertainty. Fuzzy set methods can help quantify the uncertainty associated with lack of knowledge, particularly in linguistic variables that cannot be expressed precisely using classical sets or numbers.

Moreover, UQ methods currently in use are less amenable to handling the uncertainty at a more conceptual level, such as the uncertainty associated with the formulation and definition of the boundaries of the system of the model. A broader concept of uncertainty, in terms of "its location", is needed in the assessment associated with the formulation of the model, particularly when the sources of uncertainty extend far beyond the issues of parameters, input data and model structure.

The appraisal of uncertainty often excludes the selection of the framing assumptions made in many assessments. [6] In addition, when many outcomes and complex interactions are reduced or simplified into a single framing assumption, many factors are typically ignored, resulting in an oversimplified assessment. An interesting example is found in the argument for bio-fuels in EHIA. [59] Initially, the assessment

for bio-fuels presented an ideal opportunity for sustainability and providing a solution for elevated fossil fuel-cost. However, the assessment failed to consider wider potential impacts beyond the narrow focus of the intervention, such as in the area of food security.

Recommendations for future directions and further research

It is important that we continue to investigate new methods to handle uncertainty in EHIA and that we compare the impact of different methods on EHIA results. A broader perspective of uncertainty is required to understand the wider context of the issues surrounding EHIA. This is necessary to define the boundary of the system and to quantify a structure of the context of the assessment. By quantifying a causal structure in the specified context, all sources of uncertainty can be traced backward and forward, from the conceptual sources to the analytical ones. A decision-maker might not only be interested in the analytical sources of uncertainties but more fundamentally in the conceptual ones. Framing assumptions can be inevitable when attempting to quantify health effects of interventions. Decision-makers might prefer a single estimate rather than an uncertainty range or distribution, and this might dissuade the analyst from quantifying sources of uncertainty. The diversity of the modelling approaches in quantifying uncertainty is also great. Some of the above mentioned reasons could partially explain why there is no unified approach in the EHIA literature to quantify the sources of uncertainty at a more conceptual level. There are current approaches which deal with the conceptual sources of uncertainty but they rely on qualitative assessment and expert judgment. One example is the RIVM/MNP guidance for uncertainty. [7, 60] The RIVM/MNP guidance, developed by the Netherland Institute for Public Health and the Environment, provides a systematic approach to documenting and communicating the uncertainty at different

stages in the assessment. They identify sources of uncertainty, including that associated with problem framing, by means of using checklists. However, the limitation of this and similar approaches is that they rely heavily on qualitative assessment and expert judgement to deal with conceptual sources of uncertainty.

Further research and debate are needed to standardise the way uncertainty is taken into account in EHIA modelling practice. Qualitative and quantitative approaches would be best integrated into a single framework. A systematic, integrated and comprehensive framework should be provided to represent the different sources of uncertainty. Researchers conducting quantitative EHIA can benefit from an integrated framework to handle uncertainty that extends beyond the standard methods of dealing with uncertainties by incorporating different sources of uncertainties in an explicit and systematic way.

Research paper 1:- APPENDIX A - Details of literature search strategy

Search 1 – online search of Ovid MEDLINE & EMBASE

http://ovidsp.tx.ovid.com/

(18 hits – 10 initially retrieved).

Using the "Keyword" search field and Boolean search string:

uncertainty AND health impact assessment AND (quantitative OR quantification OR model OR models OR modeling OR modelling OR modeled OR modelled OR prediction OR simulation OR projection OR software)
Database: Ovid MEDLINE & EMBASE:

1. **(uncertainty and health impact assessment and (quantitative or quantification or model or models or modeling or modelling or modeled or modelled or prediction or simulation or projection or software)).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, ps, rs, nm, an, ui]**

Search 2 – online search of ISI Web of Science
http://apps.isiknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=Gen eralSearch&SID=2DJpHfL@n@HPfhlpoPo&preferencesSaved=
(243 hits, 29 initially retrieved).

Using the "topic" search field and Boolean search string:

uncertainty AND health impact assessment AND (quantitative OR quantification OR model OR models OR modeling OR modelling OR modeled OR modelled OR prediction OR simulation OR projection OR software)
"Limits" advanced search field included: latest 5 years and English language.
Lematization: "off"

Search 3 – online search Wiley
online library
http://onlinelibrary.wiley.com/advanced/search (49 hits, 12 initially
retrieved).
Using the "abstract" search field and using Boolean search string:

uncertainty AND health impact assessment AND (quantitative OR quantification OR model OR models OR modeling OR modelling OR modeled OR modelled OR prediction OR simulation OR projection OR software)

3.3. Supplementary material to chapter 3 - Scope of the methods and their relation to uncertainty

To clarify the extent to which the methods proposed in this thesis deal with uncertainty, it is important to revisit some of the concepts highlighted in the earlier chapters. The thesis classifies uncertainty in two dimensions: the *nature* and the *location*. As shown in the earlier chapters, the *nature* of uncertainty relates to the underlying causes of uncertainty: lack of knowledge or random variatibility, and the *location* of uncertainty relates to where the uncertainty occurs in the assessment. The *location* of uncertainty is dealt with using two perspectives (conceptual and analytical), and the *nature* of uncertainty is only dealt with using lack of knowledge via a deterministic domain. These two above aspects of uncertainty define the scope of how the central issues identified in the thesis are formulated in the methods.

The nature of uncertainty

There are many interpretations of what constitute lack of knowledge or random variability. Lack of knowledge is simply defined as a type of *reducible uncertainty* that can be reduced with further research. On the other hand, random variability is defined as a type of *irreducible uncertainty* that cannot be reduced with further research. It is worth noting that as more research is conducted, uncertainty relating to lack of knowledge could be found to be of random (or stochastic) nature. However, care should be taken not to treat lack of knowledge as random variability initially in the assessment. Some probabilistic approaches such as the Bayesian method can deal with both random variability and lack of knowledge. However, it is important to distinguish explicitly between the two types when conducting probabilistic uncertainty propagation. Random variability can be interpreted as a type of

"objective probability" and lack of knowledge can be defined as "subjective probability". When a modeller is not sure about the shape of the distribution initially in the assessment, such uncertainty constitutes “lack of knowledge” and should not be treated implicitly as “random variability” (or as a stochastic phenomenon) in some instances. For example, assume a modeller dealing with parametric uncertainty has observed data with only two values for a variable representing “age” of individuals (e.g. age 12 and age 45). If the modeller assumes a uniform distribution for age, the assertion is made that “age” is a random variable , and given the distributional choice, each value for age is “equally likely to occur”. In this case, the modeller is implicitly treating” lack of knowledge” as random variability (since the modeller is not sure about the shape of the distribution and has limited data: only two observed values for “age”). If the objective of the assessment is to propagate the uncertainty when dealing with limited information or incomplete data in an explicit way, treating lack of knowledge as random variability might not be the most sensible thing to do. The key assumption in the thesis for dealing with the *nature* of uncertainty (particularly with “lack of knowledge”) is explained in the following statement.

The way the uncertainty is propagated in the assessment depends on how the uncertainty is defined in terms of its nature, that is whether the uncertainty is due to unavoidable random variations or lack of knowledge.

The thesis makes this distiction explicit, in the way the uncertainty is defined in the assessment, so that the use of methods are clear when propagating uncertainty and identifying its underlying causes. One alternative to the prior example is to define

lack of knowledge in the variable "age", and treat the uncertainty as bounds or interval in the propagation of uncertainty (research paper 3 and supplementary material chapter 5).

The location of uncertainty

As shown in the earlier sections, the location of uncertainty is broadly classified in the thesis by two types (i.e. analytical and conceptual). Much work has been done on analytical uncertainty and less work has been done on conceptual uncertainty in various quantitative EHIA (shown in research paper 1). The objective of addressing conceptual uncertainty in the thesis is to focus on a wider concept of assessment of impacts. The proposed method in the thesis only addresses the framing assumptions or mapping of the causal pathways to population health, the second aspect of conceptual uncertainty. This can be an important limitation as there are other aspects of conceptual uncertainty identified in the thesis that are not explicitly addressed in the proposed method. These include aspects relating to the "context" of the assessment which concern issues of time and space, such as defining the place and time where people are considered exposed to environmental stressors, defining the specific target population or specifying more specific health outcomes in the assessment.

Whilst contextual issues of time and space are an integral part of every EHIA, a more systems-based approach to uncertainty with the emphasis on a wider concept of assessment of impacts is proposed in the method. The rationale is to encourange researchers to think as broadly as possible about the potential range of impacts before the implementation of any EHIA. As highlighted in the literature review, a

narrow focus of assessment of impacts is exemplified in the case for biodiesel, where due to its narrow focus and definition of the assessment of impacts, other wider potential health implications in the area of food security were easily overlooked. Another example of a narrow focus and definition of the assessment of impacts is found in the study of sun exposure and skin cancer. Sun exposure is associated with a reduced risk of some types of cancer.[61] Consequently, any potential intervention that effectively reduce sun exposures may overlook the risk of increasing other types of cancers. These two examples show how a narrow definition of the assessment of impacts can lead to potential health implications to be overlooked. In this thesis, therefore, the key assumption for dealing with *conceptual* uncertainty is explained in the following statement.

A more systems-based (higher-level) approach to conceptual uncertainty is necessary to shift the focus away from a narrow definition of uncertainty in the assessment of impacts.

It is important to refocus the handling of uncertainty from a narrow definition of assessment of impacts in an EHIA. The emphasis of the proposed methods is on the framing assumptions as they can help identify potential health implications. It is worth noting that "framing assumption" describes a set of concepts relating to how causal interpretations are *assumed* in the assessment. The term "framing" refers to the construction and interpretation of causal assumptions in a HIA model. The term is used in the thesis to define the mapping of the causal pathways as they relate to human health. A potential causal interpretation, as defined in chapter 4 of the thesis does not represent "causality" in the sense of Bradford Hill criteria for causation

commonly used in epidemiology. The term “causal” is used to describe a potential interpretation of causality in any HIA model. It does not represent a real causality in epidemiology. In the formulation of an HIA, such representation of causality is necessary as all models require some form of simplification of how reality “works”. The objective of a system-based (high-level) approach and wider focus on conceptual uncertainty is to be explicit about the potential causal assumptions that are made in the mapping of the causal pathways when formulating a HIA model. (research paper 2 and supplementary material chapter 4).

4. Conceptual perspective - the mapping of the causal pathways as part of conceptual uncertainty

4.1. Preamble to research paper 2 – HIA specific question to Conceptual uncertainty

Research paper 2 proposes a method to deal with one aspect of conceptual uncertainty associated with framing assumptions. The potential pathways linking the effect of a potential intervention to health outcomes are defined in the case study of the paper. The paper attempts to quantify the sensitivity of the HIA based on how the framing assumptions are defined. It argues that prior to conducting a detailed quantitative HIA of an environmental intervention it is necessary to assess the sensitivity of the assessment to the framing assumptions. In the following paper, a plausible formulation of assessment of impacts is explored as part of a more system-based approach to conceptual uncertainty. The method is applied to a case study of housing insulation where a wider concept of assessment of impacts is explored.

Housing has been chosen as an example to explore a wider concept of assessment of impact, given that housing conditions can have a significant effect on population health.[62] The objective of the case study is to assess as broadly as possible the uncertainty about the potential pathways in which a potential housing intervention can affect health and to identify the sensitivity of the HIA to those pathways (framing assumptions). The case study in the paper rather defines the HIA question more broadly.

The case study addresses the following HIA question.

What would be the public health impacts of improving housing insulation (energy efficiency)in England?

Based on the proposed intervention, the above question relates to the potential consequences the variable "*housing insulation*" can have on the outcome *"population health"*. The HIA seeks to identify all potential pathways and their associated uncertainty relating to the effects of the intervention on the outcome. Research paper 2 focus explicitly on some aspects of conceptual uncertainty in the mapping of the causal pathways whilts ignoring other parametric aspect of uncertainty (analytical uncertainty). Research paper 2 derives effect sizes from a literature search. The actual values of the effect sizes are normalised and assessed in relative terms. Input and output variables of the model are assumed to be discrete numbers in the range -1 to +1. Such normalisation can be a source of analytical uncertainty as aspects of statistical significance are not fully addressed in the selection of the effect sizes (this can be a limitation of the proposed method).

The difficulty in quantifying the uncertainty of the health impacts of housing insulation broadly arises due to the complex mechanisms between the different pathways of exposures and population health. This broader aspect of uncertainty is prioritised in the case-study example. For example, housing insulation can improve energy efficiency but it also can decrease home ventilation. A decrease in home ventilation can increase in the growth of mould. Mould in the household can increase indoor air pollutants and play a part in the increase of cardio-respiratory conditions

such as asthma. If sources of natural ventilation are assumed (e.g. window opening, building air permeability) rather than mechanically driven, ventilation in the home can decrease energy efficiency and indoor air quality. Additionally, indoor air quality can be affected by outdoor sources of air pollution assuming natural sources of ventilation (e.g. window opening). Due the complexity of these mechanisms between the different pathways of exposures, the housing intervention is modelled as a complex system (this is illustrated in the method section of the paper). Moreover, in terms of the nature of uncertainty, the intervention is modelled in a non-probabilistic space assuming no random variables or degree of randomness in the input variables.

4.2. Research paper 2

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Marco Mesa Frias
Principal Supervisor	Zaid Chalabi
Thesis Title	Modelling Uncertainty in Environmental Health Impact Assessment

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Environment International		
When was the work published?	June 2013		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	See attached cover sheet

Student Signature: _____ **Date:** 5/10/2015

Supervisor Signature: _____ **Date:** 5/10/2015

Improving health worldwide **www.lshtm.ac.uk**

Research paper 2

Assessing framing assumptions in quantitative health impact assessments: A housing intervention example

Marco MESA-FRIAS[1]; Zaid CHALABI[1]; Anna M. FOSS[2]

[1]Department of Social and Environmental Health Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.
[2]Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.

Status: Published in Environment International, doi: 10.1016/j.envint.2013.06.002

Contributions: The candidate led in the conception of the research question in collaboration with Zaid Chalabi. The candidate conducted the literature review, developed the model, discussed the result and drafted the manuscript. Anna M. Foss contributed to the review providing comments and suggestions. Zaid Chalabi contributed and provided comments in the interpretation of the results. The candidate wrote the first draft of the manuscript. He managed each round of comments and suggestions from co-authors. All the authors read and approved the final draft prior to journal submission and inclusion in the dissertation.

The candidate_____ The supervisor_____

Abstract

Health impact assessment (HIA) is often used to determine *ex ante* the health impact of an environmental policy or an environmental intervention. Underpinning any HIA are the framing assumptions, which define the causal pathways mapping environmental exposures to health outcomes. The sensitivity of the HIA to the framing assumptions is often ignored. A novel method based on fuzzy cognitive map (FCM) is developed to quantify the framing assumptions in the assessment stage of a HIA, and is then applied to a housing intervention (tightening insulation) as a case-study. Framing assumptions of the case-study were identified through a literature search of Ovid Medline (1948-2011). The FCM approach was used to identify the key variables that have the most influence in a HIA. Changes in air-tightness, ventilation, indoor air quality and mould/humidity have been identified as having the most influence on health. The FCM approach is widely applicable and can be used to inform the formulation of the framing assumptions in any quantitative HIA of environmental interventions. We argue that it is necessary to explore and quantify framing assumptions prior to conducting a detailed quantitative HIA during the assessment stage.

Keywords: Environmental Health, Risk Assessment, Modelling, Health Impact Assessment, Housing

Introduction

The extent to which an environmental policy intervention causes health-related changes is a key question in research. Health Impact Assessment (HIA) identifies possible health consequences of new policy interventions. [63-65] HIA is an area of increasing interest to policymakers in environmental health, [66-68] and there is considerable scope for innovation in the application of quantitative methodologies. [69, 70] Underpinning any HIA are the framing assumptions, which define the causal pathways mapping environmental exposures to health outcomes. However, the sensitivity of the HIA to the framing assumptions is often ignored in many assessments. Framing assumptions are inevitable when quantifying the health effect of an environmental intervention.

Housing interventions such as improving housing insulation to reduce heat loss are examples of environmental policy interventions. Improving housing insulation, as an energy efficiency measure, is encouraged as part of the UK housing regulations to reduce carbon emission and energy cost. [71] Insulating homes is not only justified on energy efficiency grounds alone, but can also be justified on health grounds. Energy efficiency measures can benefit health through increasing indoor temperature in winter. [72, 73] However, changes in the indoor environment as a result of reducing permeability can also affect health adversely. If improving insulation is not accompanied by adequate ventilation, there is the risk of increasing indoor pollutant concentrations. [74]

Housing interventions are examples of complex (environmental) interventions.[75]There is no unique definition of a complex intervention. In general, a

complex intervention has multiple direct and indirect pathways in which it can affect health. [76] The pathways associating a complex environmental intervention with health can also be ill-defined and there are often multiple health outcomes.

HIA has been used to determine the health impacts of housing policy and interventions. [77] However, large uncertainties can arise in HIA models from the lack of understanding of the complex mechanisms between the indoor environment and health. Sources of uncertainty can include the framing assumptions associated with the formulation of the HIA, in addition to the more known sources of analytical uncertainty associated with the parameters and the structure of the models. [78]

Framing assumptions arise at the "conceptualisation" of the HIA model formulation, [4] and define the causal assumptions underpinning the assessment. The framing assumptions are typically ignored when appraising the uncertainty in many assessments by discarding factors that one considers unimportant. [4, 6] Since the outcome of a HIA can be highly sensitive to the choice of the framing assumptions made initially in the assessment stage, it is important to characterise and quantify these framing assumptions.

Mathematical methods can be used to quantify the framing assumptions when defining the context of the assessment in evaluating the health impact of environmental interventions, *ex ante*. The use of complex system mathematical models has been proposed in public health. [79-81] This paper demonstrates the use of another type of complex system modelling approach, known as fuzzy cognitive mapping (FCM). In this study, we use FCM to quantify the framing assumptions in

the assessment stage of a HIA model of housing insulation, as a case-study example. The approach however is widely applicable to others examples of complex environmental interventions.

Overview of FCM method

A cognitive map is a conceptual graphical model used to represent causal assumptions. [82, 83] Cognitive maps have been used for conceptual modelling in many areas in the social sciences, such as in assessing the social implications of nanotechnologies and in describing social knowledge in the political sciences. [84, 85] Cognitive maps can be extended to incorporate imprecise qualitative knowledge into quantitative variables, known as fuzzy cognitive maps. Fuzzy cognitive maps (FCM) have been used as a modelling tool to represent conventional and Aboriginal perspectives on the determinants of diabetes. [86]

In this study, FCM is used to model framing assumptions quantitatively. Framing assumptions can be first explored with the use of causal diagrams. A causal FCM diagram shows the connections between variables in the "system of interest" and can be used to define the context of the assessment in which the environmental intervention is applied. The main emphasis of using causal FCM diagrams is on identifying causal pathways as they relate to health outcomes.

In general, FCM diagrams are directed graphs, which indicate directional links in the causal pathways. Fuzzy cognitive maps diagrams are described by a set of nodes and their causal links. In the context of this study, each node represents a key indoor factor, a health or a non-health outcome. The relationships between the nodes are

described through directional links or connections. Positive (+) and negative (–) signs imply positive and negative causality, respectively. A positive causal link between a pair of nodes means that when the amplitude (level) of one node increases, the amplitude of the other increases. A negative causal link, on the other hand, means that when the amplitude of one node increases, the amplitude of the other node decreases. A value zero (0) between a pair of nodes implies there is no causal link between the nodes.

A FCM was developed here to model the framing assumptions in the assessment stage of a HIA model of housing insulation. Fuzzy cognitive maps were then used to investigate the causal meachnisms and explain semi-quantitatively how intervention-related changes in the indoor environmental exposures can potentially affect health. Our methodological approach developed in this study is described in five main steps below.

Five steps in assessing framing assumptions

The five main steps in assessing framing assumptions are: (1) synthesising the evidence on causal pathways from the literature; (2) constructing the causal diagrams from individual studies identified from the literature; (3) representing mathematically the combined causal diagram as a system matrix; (4) measuring the structural properties of the system matrix; and (5) perturbing the system to identify causal processes. Details of the steps are described below. Refer to Appendix A for detailed mathematical description of the steps and Appendix B for a walk-through example.

Synthesising the evidence on causal pathways from the literature

Health-relevant factors and outcomes were identified in the literature to construct causal diagrams that define nodes and their links. A literature search of Ovid Medline (1948-2011) was conducted using the search terms: "housing" combined with "insulation" and "health" to identify studies investigating factors and outcomes (nodes) influencing the links between housing insulation and health. Causal pathways linking housing insulation and health were identified qualitatively. An additional hand search of the literature was conducted in Ovid Medline using the identified key factors and outcomes as search terms to determine quantitative information on the links between the nodes.

Constructing the causal diagrams from individual studies identified from the literature

Based on each published study retrieved from the literature - nodes were identified. An individual casual diagram was constructed and positive or negative associations between the nodes of the diagram were determined. Measures of effects, such as odds ratio, were subsequently used to quantify the strength of the causal links between the nodes. The measures of effects ("causal weights") were noted with each connection between a pair of nodes to represent the strength of the effects, using either the natural logarithm of an odds ratio for a health outcome, or the percentage change in indoor factors or outcomes obtained from retrieved studies in the literature (Appendix A.1).

Representing mathematically the combined causal diagram into a system matrix

Each causal diagram was then mathematically translated into a "connection matrix." The elements of each connection matrix correspond to the measure of effects

between each pair of nodes (causal weights). Each element is an algebraic number, which can be positive or negative. The value zero (0) means that there is no causal link between the nodes. The matrices from each published study were combined through summation and their values were then normalised (by dividing each element by the absolute maximum across all elements) to create a "system matrix" in which each element was in the range -1 to +1 (Appendix A.2).

Measuring the structural properties of the system matrix

The structural properties of the system matrix represent the causal structure mapping the causal pathways in the diagram. Indices are numerical measures, calculated using graph theory, [87] which characterise quantitatively the structural properties of the system. A "centrality index" shows how well connected a node (indoor factor or an outcome) is in relation to other nodes, i.e. how many links join with this specific node. The centrality index measures the centrality of the framing assumptions defined in the assessment. A high centrality index indicates high importance, whereas a low centrality index means less relevance in the system. Nodes are classified according to their input and output values (which are signed causal weights entering or leaving a node, respectively). Those nodes with only input values (i.e. arrows directed to them) can be viewed as the "outcomes" while nodes with only outputs values (i.e. arrows directed from them) may be viewed as the "drivers" or "stressors". Nodes with both input and output values can be viewed as "mediating factors" playing both roles. The centrality index is calculated by summing the magnitude of the total input and output values in the system (Appendix A.3).

Perturbing the system to identify causal processes

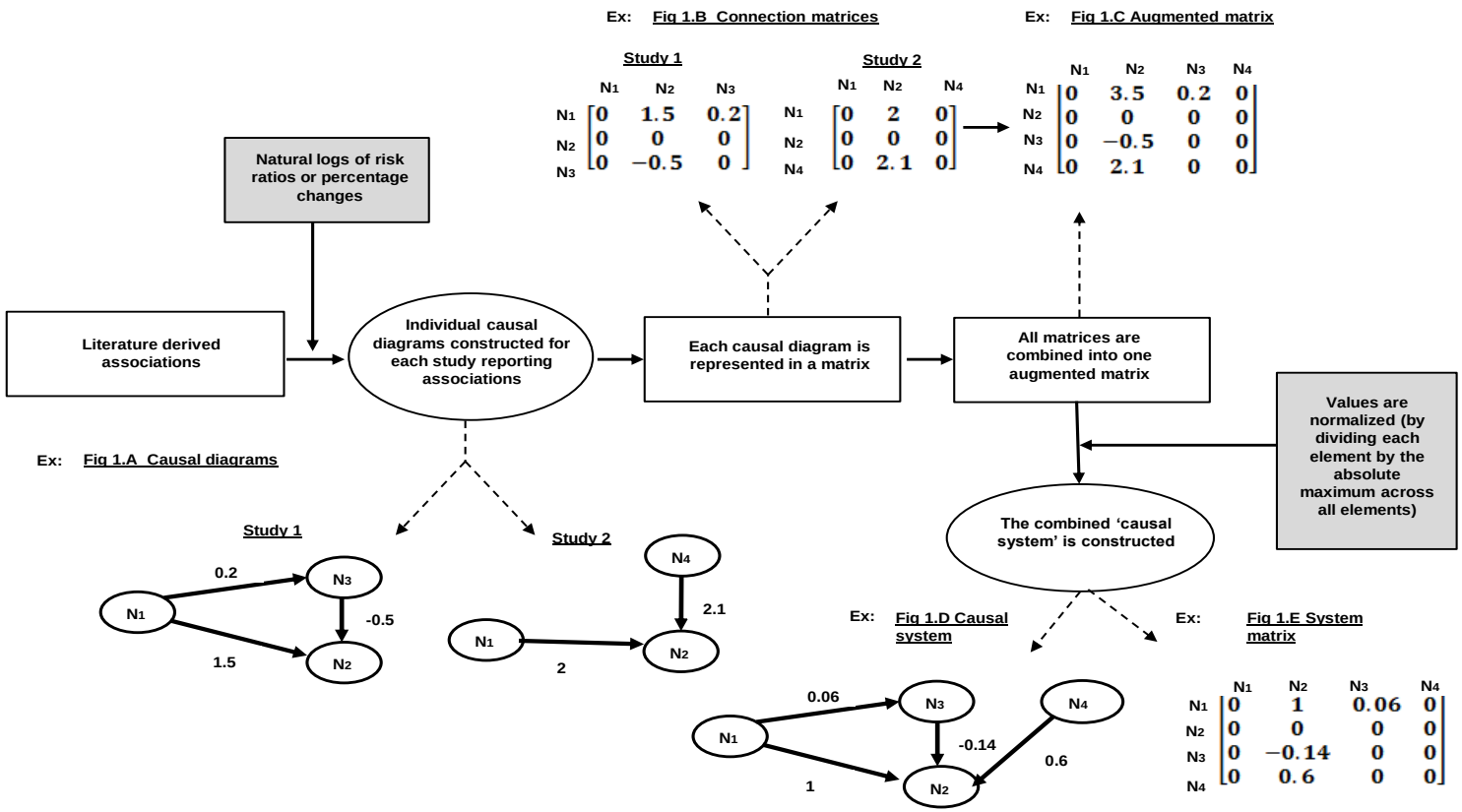
This step is concerned with assessing the sensitivity of the assessment to the framing assumptions. It explores how the intervention "works" based on the framing assumptions made initially in the assessment. Causal processes are evaluated in the system matrix by means of a deterministic perturbation between the nodes in the diagram. A "causal process" describes the mechanisms of the causal pathways in the nodes. Each node can have a "causal activity level" which measures how each node ranks in relation to each other in the causal pathway. This causal activity is represented by values between 0 and 1 in the nodes. A node with value 0 denotes the node is fully "inactive" while a node showing a value 1 means that the node is fully "active" in terms in the causal pathway. The nodes are propagated through the causal pathways in a deterministic perturbation analysis until the system reaches equilibrium. The state of the system at equilibrium depicts the key causal processes (or sources of variations) in the nodes (Appendix A.4).

Summary of procedures

For easy of illustration, Figure 4 shows the methodological approach and procedures in diagrammatic form and in mathematical matrix representation. The data are hypothetical. The initial phase of the FCM development consisted of developing individual causal diagrams for each study based on causal links derived from the literature review (Fig 4.A). The natural logs of risk ratios (or percentage changes) were calculated to define the causal weights in each of the causal diagrams. Each causal diagram was then deterministically represented in a matrix (Fig 4.B). Matrices were combined into one augmented matrix (Fig4.C).

The elements of the augmented matrix were then normalised between -1 and 1 to give the system matrix (Fig 4.E).The combined causal system is represented graphically (Fig 4.D) and in matrix form (Fig 4.E).

Figure 4: Summary of Procedures



Shaded boxes are mid-point calculations; ovals are diagrams; Ni= nodes; doted arrows are examples, straight arrows are procedures

Results

The literature search generated 40 articles from which 12 articles had sufficient qualitative information to establish the links between indoor environmental factors and health outcomes. [73, 88-98] Indoor factors linked to housing insulation that have been shown to have an effect on health, were grouped into two broad themes: indoor environmental exposures and built indoor environment.

Based on the retrieved literature, Table 3 gives a list of potential health-relevant factors linked to the housing insulation for inclusion in the causal diagrams. Factors identified in connection with the indoor environmental exposures were indoor temperature (cold), air-tightness, indoor particles, dampness and mould. Factors identified in relation to the physical aspects of the built indoor environment were insulation fabric material, and mechanical ventilation systems. Among the health outcomes identified were winter mortality, mental health, depression, and respiratory conditions such as asthma and wheezing.

In general, the identified studies had different epidemiological designs and each study focused on various associations between different indoor factors and health. This required the assignment of a more generic classification of the indoor factors and health outcomes in the causal diagram. For example, health outcomes such as wheezing, throat irritation, bronchopneumonia, winter mortality and asthma were broadly classified as: *Cardio-respiratory morbidity/mortality*.

Table 3: List of potential health-relevant factors and outcomes associated with housing insulation

Theme	Factors
Indoor Environmental Exposures	Indoor air quality Relative humidity Dampness, mould Particles generated from indoor sources (PM2.5 or PM10) Environmental tobacco smoke (ETS) Combustion (carbon monoxide, nitrogen oxides) Particles (PM2.5 or PM10) generated from outdoor sources Radon Volatile organic compounds (VOCs) Indoor temperature
Built Indoor Environment	Thermal insulation/ fabric material Mechanical ventilation systems Housing design and construction factors Air-permeability (air-tightness) Ventilation
Outcomes	Mortality Carbon dioxide/ energy savings Eye, nose, throat irritation Bronchopneumonia and pulmonary oedema Cough Asthma Wheezing Cardio-respiratory conditions Depression/ mental health Thermal comfort/ psychosocial wellbeing Fuel poverty

Indoor factors representing several pollutants affecting indoor air quality such as PM2.5, nitrogen dioxide, carbon monoxide, volatile organic compounds (VOCs), radon and environmental tobacco smoke (ETS) were classified as: *Indoor air quality*. In addition, two indoor factors corresponding to the built indoor environment were considered: *Thermal insulation and mechanical ventilation* because they are important energy efficiency measures. [77]

A total of 9 studies were identified to have quantitative information that could be used to assign measures of effects for the causal links between indoor factors and outcomes. [77, 91, 95, 99-104] Table 4 gives the key health-relevant factors and their reported quantitative associations. Studies judged to represent the same (or equivalent) link between an indoor factor and an outcome, were combined by summing the measures of effects. For example, effect sizes from factors that represented different types of pollutants such as: carbon monoxide, formaldehyde (VOCs), radon and environmental tobacco smoke (ETS) where combined by summing their effect sizes and the total effect assigned to the node *Indoor air quality*. This level of resolution was deemed appropriate to test the plausibility of the causal structure (framing assumptions) and its mechanism. The overall measures of effects were determined as described in the procedure above and inputs were assumed to deterministic (Appendix A.2).

Table 4: Key indoor factors and their reported associations as part of insulation improvements

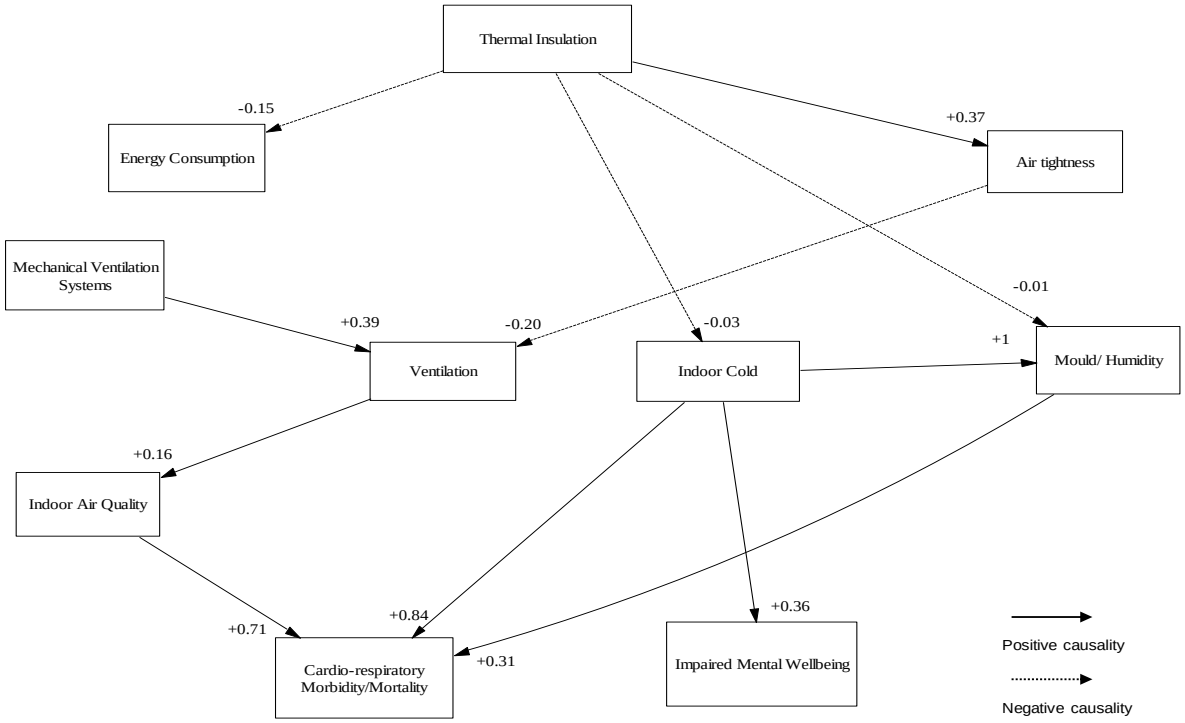
Factors	Affects Factors	Affected by Factors	Reported Association and health impacts	Strength of evidence
1. Thermal Insulation	2,3,6,8	none	Howden Chapman et al., (2007): Energy use OR 0.81(0.72 to 0.91)▼, Indoor temperature during winter increase of 0.5^{o} C (0.03 to 0.95) in bedrooms, ▲, decreased relative humidity 2.3%▼; Wilkinson et al., (2009): air permeability (air tightness) average stock improvement from 13 (m^3 /m^2 /h) to 6 (m^3 /m^2 /h)▲ ~ % 53 improvement in air tightness through insulation	++ +
2. Energy Consumption	none	1	Howden Chapman et al.(2007): Energy use OR 0.81(0.72 to 0.91) ▼	++
3. Air Tightness	4	1	Hirsh et al., (2000): ventilation decreased from geometric mean 0.73 to 0.52 per hour ~ percentage change 29%▼	+
4. Ventilation	7	3,5	Fisk et al., (2009): If ventilation rate decreases from 10 to 5 l/s-person indoor air quality reduces 23% approximately	++
5. Mechanical Ventilation Systems	4	none	Engvall et al., (2003): through an improvement of ventilation OR 0.57 (0.29 to 0.85) ▲, mechanical ventilation system decreased ocular and nasal symptoms	+
6. Indoor Temperature (Cold)	8,O	1	Wilkinson et al., (2009): Warm Front study indicated a 2% increase in risk of Cardio-vascular disease winter death for 1° C decrease in standardized indoor temperature (SIT)*; Braubach., (2007): Depression and Mental Health OR 1.404 ▼ through improving insulation; Howden Chapman et al., (2007): a decrease in self perceived cold, through insulation, improves social wellbeing = percentage change +6.2% ▲ ,emotional wellbeing + 10.9% percentage change ▲; Self-reported symptoms of cold or flu OR 0.54 (0.43 to 0.66)▼; wheezing in last 3 months OR 0.57 (0.47 to 0.70)▼ , mould OR 0.24▼	+ + ++
7. Indoor Air Quality	O	4	Mendell (2007): formaldehyde concentration OR 1.4▲ 0.34 per 20 $ug.m^{-3}$; 0.0167 ~ 2 % excess risk of allergy/asthma per $ug.m^{-3}$; Wilkinson et al., (2009): Radon %0.15 excess risk in lung cancer▲, ETS increased RR 1.30 heart disease and RR1.25 for cardio-vascular disease▲; Smith et al., (2011): Carbon Monoxide indoor pollution on pneumonia RR 0.82 (0.70-0.98) ▲	+ +
8.Mould / Humidity	O	3	Fisk et al., (2007): mould/dampness associated with an increase in asthma, cough, wheeze and upper respiratory symptoms OR 1.545 (1.34-1.75) ▲	+++

. O = health outcomes (cardio-respiratory morbidity/mortality and Impaired mental health/psychosocial wellbeing); direction of association: ▼ = reduction; ▲ = increase; + evidence from one uncontrolled study; ++ evidence from at least one prospective controlled study; +++= evidence from some prospective controlled studies

Representation of the causal system

The review of the literature identified 10 key indoor factors or outcomes and 12 associations. Figure 5 shows the causal system displaying the causal pathways linking housing insulation and health, based on the evidence available from the literature review conducted. Table 5 gives a representation of the system matrix used to calculate the centrality index and to simulate causal processes.

Figure 5: Framing assumptions in the system: modelling the process of change among indoor factors and outcomes



Positive (+) or negative (-) signs describe a positive causality or a negative causality respectively. A positive causality or a causal increase indicates when node *i* increases, node *j* increases. A negative causality or a causal decrease indicates when node *i* increases, node *j* decreases.

Table 5: System matrix linked with “Causal system”

	Thermal Insulation	Energy Consumption	Air tightness	Mechanical Ventilation System	Ventilation	Indoor Cold	Indoor Air Quality	Mould / Humidity	Cardio-Respiratory Morbidity /Mortality	Impaired Mental Wellbeing
Thermal Insulation	0	-0.15	0.37	0	0	-0.03	0	-0.01	0	0
Energy Consumption	0	0	0	0	0	0	0	0	0	0
Air tightness	0	0	0	0	-0.20	0	0	0	0	0
Mechanical Ventilation System	0	0	0	0	0.39	0	0	0	0	0
Ventilation	0	0	0	0	0	0	0.16	0	0	0
Indoor Cold	0	0	0	0	0	0	0	1.00	0.84	0.36
Indoor Air Quality	0	0	0	0	0	0	0	0	0.71	0
Mould / Humidity	0	0	0	0	0	0	0	0	0.31	0
Cardio-Respiratory Morbidity /Mortality	0	0	0	0	0	0	0	0	0	0
Impaired Mental Wellbeing	0	0	0	0	0	0	0	0	0	0

Structural assumptions

The main indoor factors and health outcomes as identified by the centrality index were indoor *cold, cardio-respiratory morbidity / mortality* and *mould / humidity*, as shown in Table 6. High centrality values reflect high connectivity of the nodes in the system. A high centrality index can be interpreted as key structural assumptions made in the assessment. Centrality overall was low among most nodes, with 7 nodes having centrality index less than unity 1.0. Figure 6 shows graphically the centrality values.

Table 6: Indoor factors and outcomes included in the system diagram ranked by their centrality indices

Factors	**Centrality**	**Type**
Indoor Cold	2.23	M
Cardio-respiratory Morbidity / Mortality	1.86	O
Mould / Humidity	1.32	M
Indoor Air Quality	0.87	M
Ventilation	0.75	M
Air Tightness	0.57	M
Thermal Insulation	0.56	D
Mechanical Ventilation System	0.39	D
Impaired Mental Wellbeing	0.36	O
Energy Consumption	0.15	O

D = drivers, O = outcomes, M = mediating factors

Main causal processes and sources of variations

As described above, the purpose of the perturbation analysis is to determine the steady-state (equilibrium) level of the causal activity of the nodes (indoor factors and outcomes). The level of causal activity of the nodes denotes the sensitivity of the assessment to the framing assumptions. Main causal processes and sources of variations can be identified via the level of causal activity in the nodes at equilibrium. Based on the causal diagram shown in Figure 5, a perturbation analysis was carried out (Appendix A.4). Figure 7 shows the level of causal activity at equilibrium for each node.

Figure 6: Distribution of centrality values in causal system

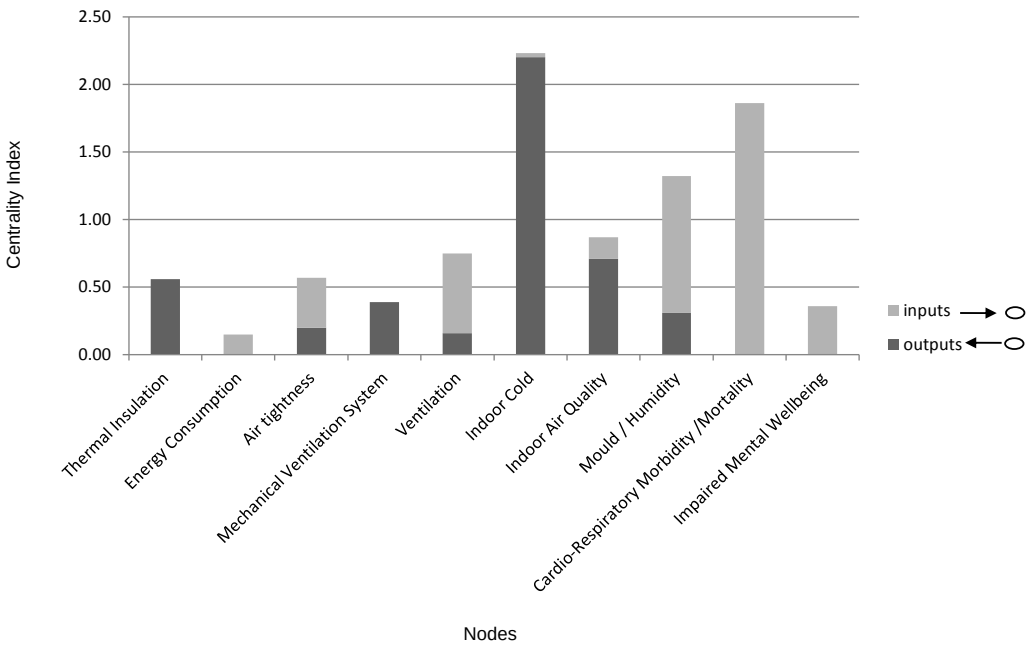
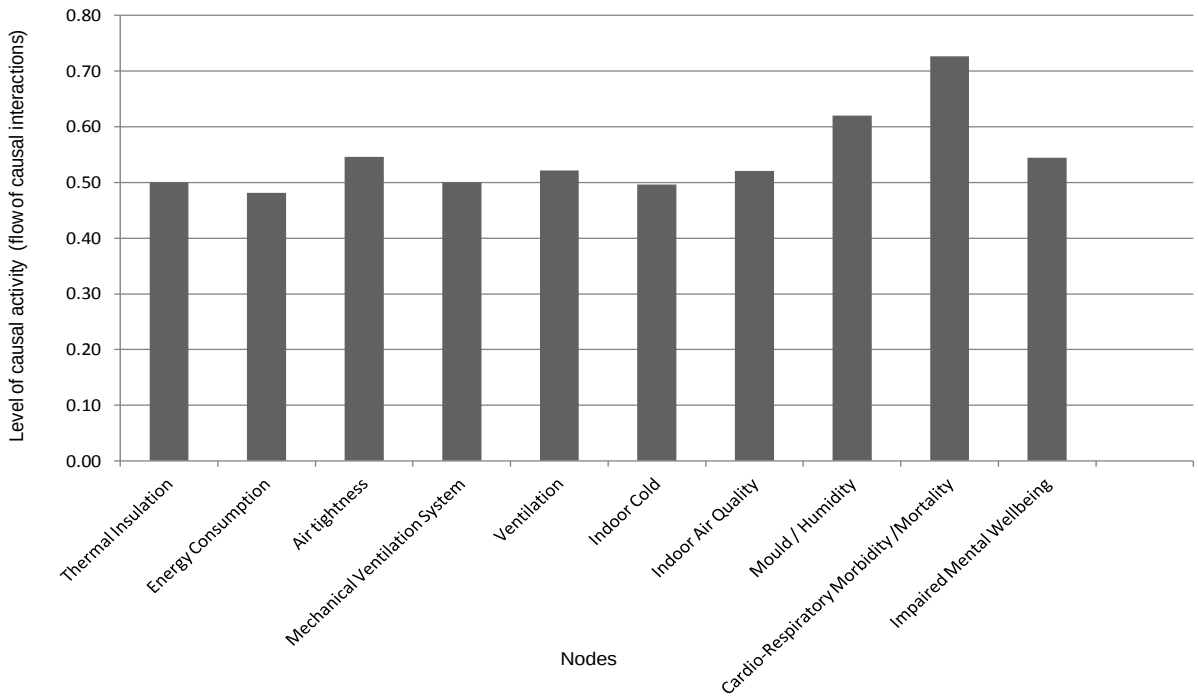


Figure 7: Causal processes after perturbation based on the structural framing assumptions of the system



Discussion

In this study, we presented a novel methodology to quantify the framing assumptions in a HIA conceptual model example of housing insulation. Framing assumptions represent a set of causal interpretations made about the system based on the evidence available in the literature. This study focused on the causal pathways linking housing insulation and health.

Indoor cold, mould, humidity and cardio-respiratory morbidity/mortality were found to be central to the framing assumptions. In addition by taking a threshold value of 0.5 (midpoint between the lowest and highest value of "causal activity"), the simulation recorded "high level of causal activity" (i.e. higher than 0.5) in the following nodes: cardio-respiratory morbidity / mortality, impaired mental wellbeing, mould / humidity, indoor air quality, ventilation and air-tightness. The threshold value of 0.5 was considered appropriate to test the sensitivity of the framing assumptions on the basis of how each factor ranked in relation to each other. Changes in the health outcome nodes (e.g. respiratory morbidity / mortality, impaired mental wellbeing) are naturally expected to be high because most pathways lead to them. What is more relevant, however, is the finding of the high level of causal activity in the nodes air-tightness, ventilation, indoor air quality, and mould/humidity. Given their high level of causal activity, these indoor factors were identified as being highly sensitive to the framing assumptions. This means that changes in these factors are particularly important because they influence health outcomes and, therefore, can cause health-related changes in relation to the intervention.

Any framing assumptions are likely to be incomplete because they are based on factors or outcomes obtained from a relatively restricted search of the published literature. In the case study example, social factors such as housing composition, socio-economic status, the behaviour of residents were not considered due to lack of quantitative information to assume causal relationships. A more comprehensive representation of the framing assumptions would require a broader range of studies to incorporate housing and social factors, health outcomes and their links. In addition, we assumed that the included studies provided the same level of evidence and where comparable in terms of population intervention, study type and study quality since our emphasis was at the system level. [79, 105] For an extensive analysis on housing insulation and health, a systematic literature review will be required with quality assessment criteria prior to selecting the studies to be included in the FCM. Weights can be assigned based on the strength of evidence obtained from a systematic review. Causal weights can be specified in the FCM without affecting the mechanics of the method. Once quality criteria of each study are assessed, and weights are assigned, the result of a FCM can be used to inform the selection of the framing assumptions prior to conduction a comprehensive quantitative HIA.

It is worth noting that most HIAs seek to assess the health impacts of an intervention before a particular policy proposal is implemented. HIA comprises various stages such as: “screening”, “scoping” , “impact assessment” “policy modification and evaluation”. [106] Of particular interest is the “impact assessment” stage, where the health impacts of a proposal are identified, and causal pathways are constructed. Assessing the sensitivity of the framing assumptions in this stage of the assessment

is important. The FCM approach can be applied to supplement this stage of the assessment.

We argue that it is necessary to quantify framing assumptions prior to conducting a comprehensive HIA. This study has highlighted the use of appropriate methods using FCM to with the framing assumptions. Decision makers should be aware that framing assumptions can have a significant impact in the outcome of the assessment. Our methodology depicts an objective method for quantifying causal assumptions at the system level. We believe that this method can handle many more complex causal pathways than that shown here.

Conclusion

This paper proposed a new method to quantify the framing assumptions in the initial stage of a health impact assessment of an environmental intervention. The method was illustrated using a housing intervention (insulation), as a case-study. The substantive findings of the approach hold promise in terms of applying it to other examples of environmental interventions. We argue that it is necessary to deal explicitly with the framing assumptions prior to conducting a full assessment of the health impacts of an environmental intervention.

Appendix

.

APPENDIX A

Constructing the causal diagrams from individual studies identified from the literature

A causal diagram is constructed based on each published study (k) identified in the literature search. For each identified study k, causal weights ($a_{ij}^{(k)}$) are derived from the reported measures of effects to describe the "strength of the causal link" between the variables (nodes). The causal weight is defined as the natural logarithm of the odds ratio (OR), if it is reported in a study between a pair of nodes i.e.

$$a_{ij}{}^{(k)} = ln(OR_{ij}{}^{(k)}) \qquad [1]$$

$a_{ij}^{(k)}$ represents the measure of effect (or causal weight) between two nodes i and j. Effect sizes expressed in other measures of effect, such as correlation coefficient or standardized mean difference, can be converted into odds ratio $(OR_{ij}{}^{(k)})$ [107] to obtain the corresponding causal weight $(a_{ij}{}^{(k)})$. If a measure of effect is reported as a percentage changes it is turned into a causal weights by expressing it in a decimal form $(e.g.\ 50\% = 0.50)$. If the measure of effects are not provided in a study, the causal weights are set to 0 for that pair of variables and for that study.

Representing mathematically the combined causal diagram into a system matrix

Each study k can be used to construct a causal diagram. A causal diagram can be represented mathematically by a $N \times N$ "connection matrix" (A) with N nodes such that the elements of the matrix A are given by

$$a_{ij}^{(k)} = \begin{Bmatrix} \pm \log\ (OR_{ij}^{(k)}) \\ 0 \end{Bmatrix} \qquad [2]$$

Each $(i\ j)^{th}$ element of the connection matrix $A_k = \left\{a_{ij}^{(k)}, i = 1..N, j = 1..N\right\}$ represents the measure of effect (or causal weight) between two nodes i and j. The causal weights are algebraic numbers which can be positive or negative. If $a_{ij}^{(k)} = 0$, it means that the nodes i and j are not connected. The connection matrices from all the identified studies are combined into a single matrix, known as the "system matrix" (S) representing all the matrices $\left\{A^{(k)}, k = 1..m\right\}$. Each element of the system matrix S (s_{ij}) is defined as

$$S = \left\{s_{ij} = \sum_{k=1}^{m} a_{ij}{}^{(k)}\ ;\ i = 1..N, j = 1 \ldots N\right\} \qquad [3]$$

where m is the total number of identified studies. Each element of the matrix S is normalised between $-1\ to\ 1$ by dividing each element by the absolute maximum across all elements [86]

$$S' = s'_{ij}$$

$$= \frac{s_{ij}}{max(|s_{ij}|)} \qquad [4]$$

where $|s_{ij}|$ denotes the absolute value of s_{ij}. s'_{ij} gives the relative "weight of evidence" for the links between any nodes i and j. For simplicity, we well refer to S'as also the system matrix.

Measuring the structural properties of the system matrix

The structural properties of the system matrix (S') can be analysed quantitatively. Indices are calculated using graph theory. A centrality index shows how "well connected" a node i is in relation to other nodes. The centrality index (c_{ij}) is simply calculated by the sum of the total input connection values ($K_i^{(in)}$) to node j and the total output connection values *(* $K_j^{(out)}$ *)* from node i

$$c_{ij} = K_j^{(out)} + K_i^{(in)} \qquad [5]$$

where $K_i^{(in)}$ and $K_j^{(out)}$are given respectively by

$$K_i^{(in)} = \sum_{j=1}^{N} |\ s'_{ji}\ | \qquad [6]$$

$$K_j^{(out)} = \sum_{i=1}^{N} |s'_{ij}| \qquad [7]$$

Pertubing the system to identify causal processes

To explore causal processes and mechanisms (including feedbacks), a deterministic perturbation analysis is conducted as follows. Denote by $\left(V^{(t)}\right)$ the n-dimensional state vector of the system at iteration t. Each nth element of the state vector represents the state of "causal activity" of the nth node. An input state vector (unit vector as initial condition) $(V^{(t)})$ at iteration t is multiplied by the system matrix S' (s'_{ij}) to generate a new vector $(V^{(t+1)})$ *at iteration* $t+1$. The resulting vector $(V_j^{(t+1)})$ is repetitively multiplied by matrix S' until the state vector reaches a stable equilibrium level [14, 86]:

$$V_j^{(t+1)} = f \quad \langle \sum_{i=1, j\neq 1}^{N} V_j^{(t)} \times (s'_{ij}) \rangle \qquad [8]$$

At each iteration each element $(u)$of the vector $V_j^{(t+1)}$ is normalised to be within the interval of $[0, 1]$ by applying pointwise a threshold function $f(u)$ (i.e. to each of its elements) [14, 108]. The threshold function is a logistic continuous function which determines the degree of activation level of the nodes after every iteration until equilibrium is reached:

$$f(u) = \frac{1}{1 + e^{-u}} \qquad [9]$$

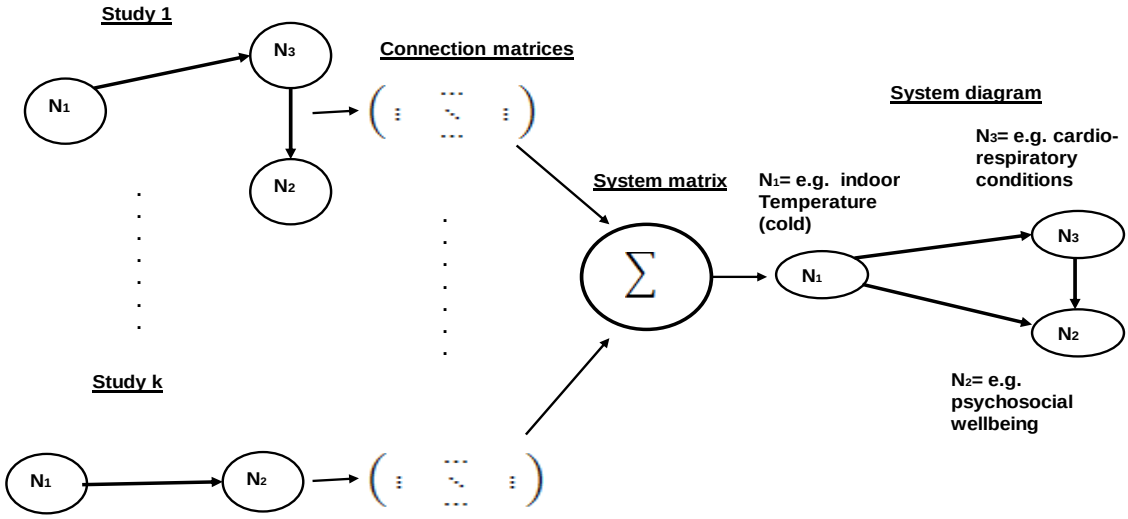
The equilibrium state describes the steady-state stable causal configuration of the system. Each nth value of the state vector represents the level of activation ("causal activity") in the nth node. The level of activation reflects how each node influences each other over a number of iterations. The level of activation is a value between 0 to 1, where 1 is the highest level of causal activity and 0 is the lowest level of causal activity. The purpose of the perturbation analysis is to measure the steady-state activity in each node in terms of feedbacks and causal mechanism.

APPENDIX B

The appendix provides a walk-through hypothetical example to guide the reader through the various steps of the approach described in the paper. The hypothetical example concerns the pathways between indoor temperature (cold), cardio-respiratory and psychosocial (wellbeing) conditions. The example is demonstrated in five pseudo-algorithmic steps.

Step #1 (Figure B.1): Combine separate casual diagrams into one system diagram. The figure shows schematically two separate studies (out of k studies) concerned with the connection of three variables (nodes), N_1 to N_3. A diagram is constructed based on each study to show the causal links between the nodes. A connection matrix is formed for each diagram and then a system matrix is constructed by combining the matrices as shown schematically below (refer also to the section "Representing mathematically the combined causal diagram into a system matrix" in Appendix A, for the mathematical details).

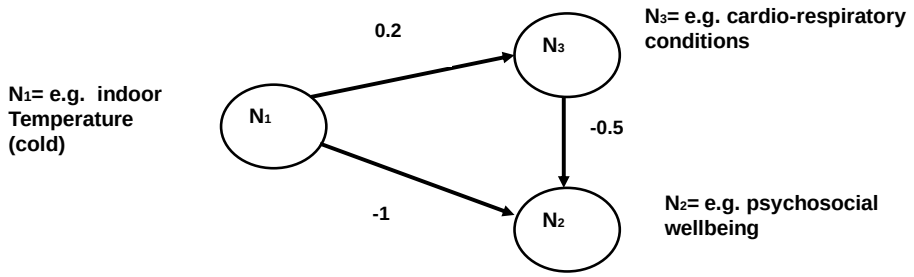
Figure B.1: Individual causal diagrams and system diagram



Causal diagrams are constructed based on each study identified in the literature review. All causal diagrams are combined to form a "system diagram

Step # 2 (Figure B.2): Parameterise the system diagram. The figure shows the system diagram obtained by combining all the studies from the literature review reporting associations between the system variables. Causal weights between the nodes of a diagram in each study are obtained from reported measures of effects (e.g. odds ratio, percentage change, etc.). The causal weights in the system diagram represent the combined "strengths" or "relative weights" of evidence across all reported links (see also section "Representing mathematically the combined causal diagram into a system matrix" and equation [4] in Appendix A).

Figure B.2: The system diagram



Causal weight in the system diagram = s'_{ij} (e.g. $s'_{1,2}$ = -1) derived and combined from all reported associations obtained in the literature review

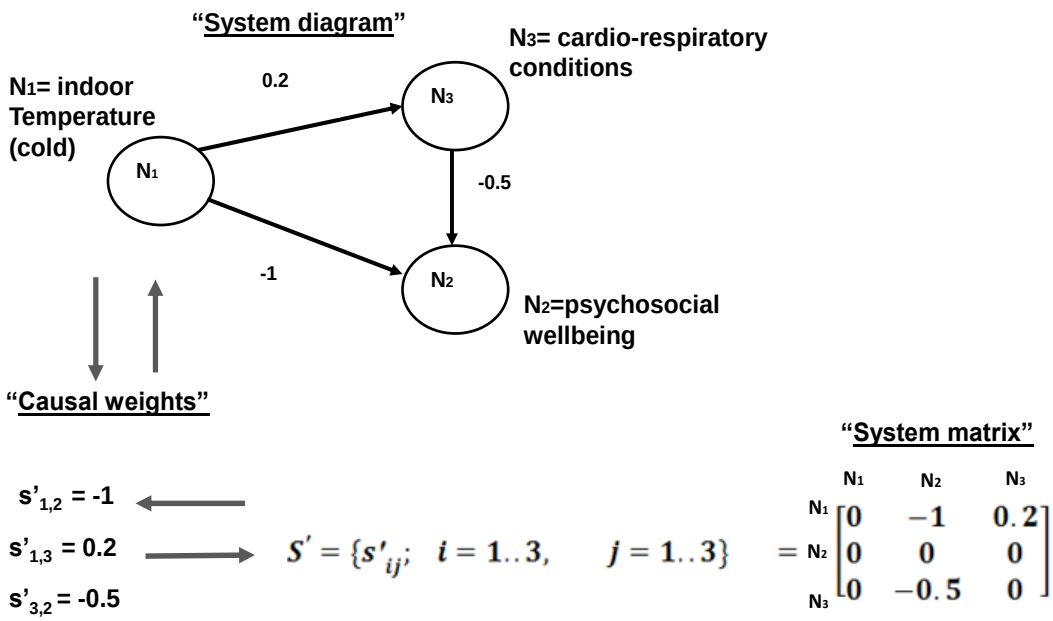
Positive causality	Negative causality
Strong = 1	Strong = -1
Medium = 0.5	Medium= -0.5
Poor = 0.2	Poor= -0.2

The System diagram represents the combination of all studies reporting associations from the literature review. Causal weights, assigned to each link in the system diagram represent the "strength" or "relative weight" of evidence. Causal weights can also be described qualitatively (as shown in the table).

It is important to note that the causal weights are algebraic quantities. A positive causal link (between nodes i and j) means when node i increases, node j increases. A negative causal link (between the same nodes i and j) on the other hand means that when node i increases, node j decreases. The amplitudes of the causal weights in a system diagram can be interpreted qualitatively as indicating the qualitative strength of the links e.g. "Strong" if the amplitude of the causal weight ≥ 0.90; "Medium" if it is ≥ 0.50, and "Poor" if it is ≤ 0.50.

Step # 3 (Figure B.3): Construct the system matrix from the system diagram .A system diagram (showing the causal links between nodes) can be represented by a matrix. For a system diagram with three nodes (N_1 to N_3), we construct a 3 × 3 system matrix (S') as follows. If there is a causal link between node N_1 and N_3, place the causal weight value as an entry in row N_1 and column N_3 in the matrix S' e.g. $s'_{1,3} = 0.2$. If a causal link is not directed between a pair of nodes, set the causal weight to 0 e.g. $s'_{2,3} = 0$ (refer to section "Constructing the causal diagrams from individual studies identified from the literature" in Appendix A for the mathematical details).

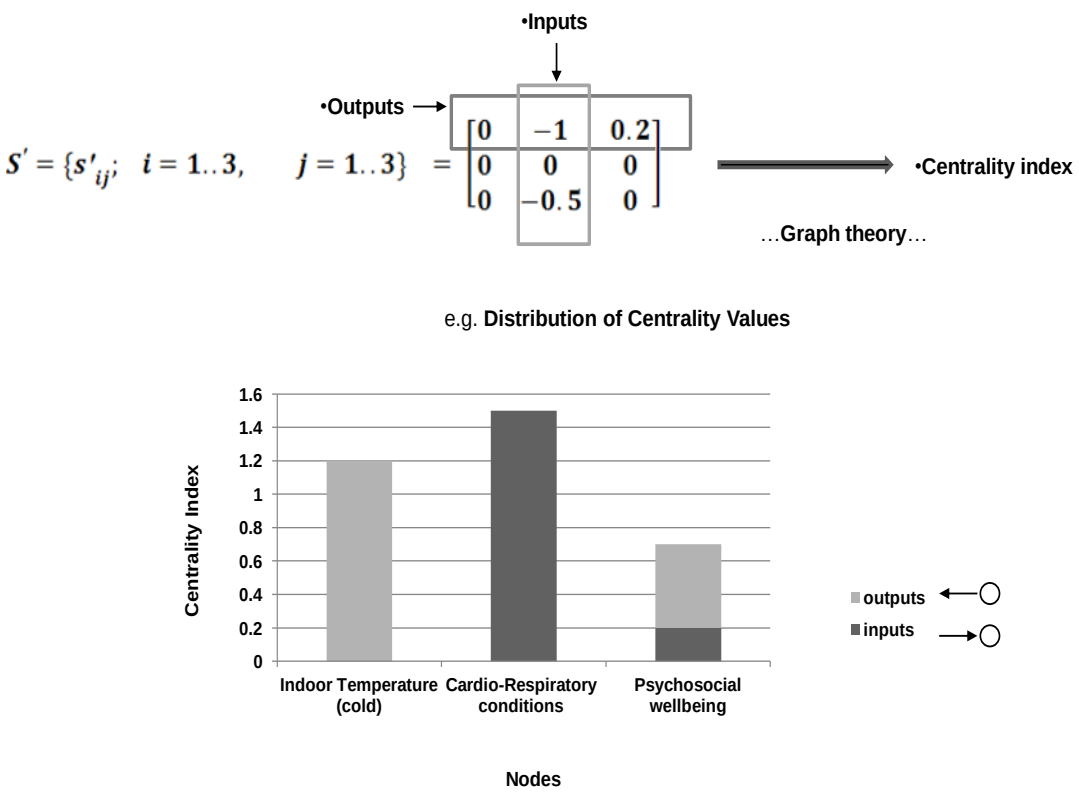
Figure B.3: Mathematical matrix representation



The graphical structure of the system diagram can be represented mathematically in the form of matrix. Each causal weight value (positive or negative) is placed as an entry in the appropriate row and column in the matrix.

Step # 4 (Figure B.4): Determine the structural properties of the system using a quantitative measure called the centrality index. The centrality index is obtained by summing the total "inputs" and the total "outputs" connection values (causal weights) of the matrix (See equations [5], [6] and [7] in Appendix A). Input connection values are the values of the causal weights corresponding to the columns of the system matrix; output connection values are the values of the causal weights corresponding to the rows of the system matrix. The values of the centrality indices of the nodes and the outputs from the nodes and inputs to the nodes are shown in the table and figure below.

Figure B.4: Centrality index



Structural properties of the system can be analyzed once the system matrix is obtained using the centrality index. The centrality index is calculated by the total sum of the absolute values of column and row entries in the system matrix

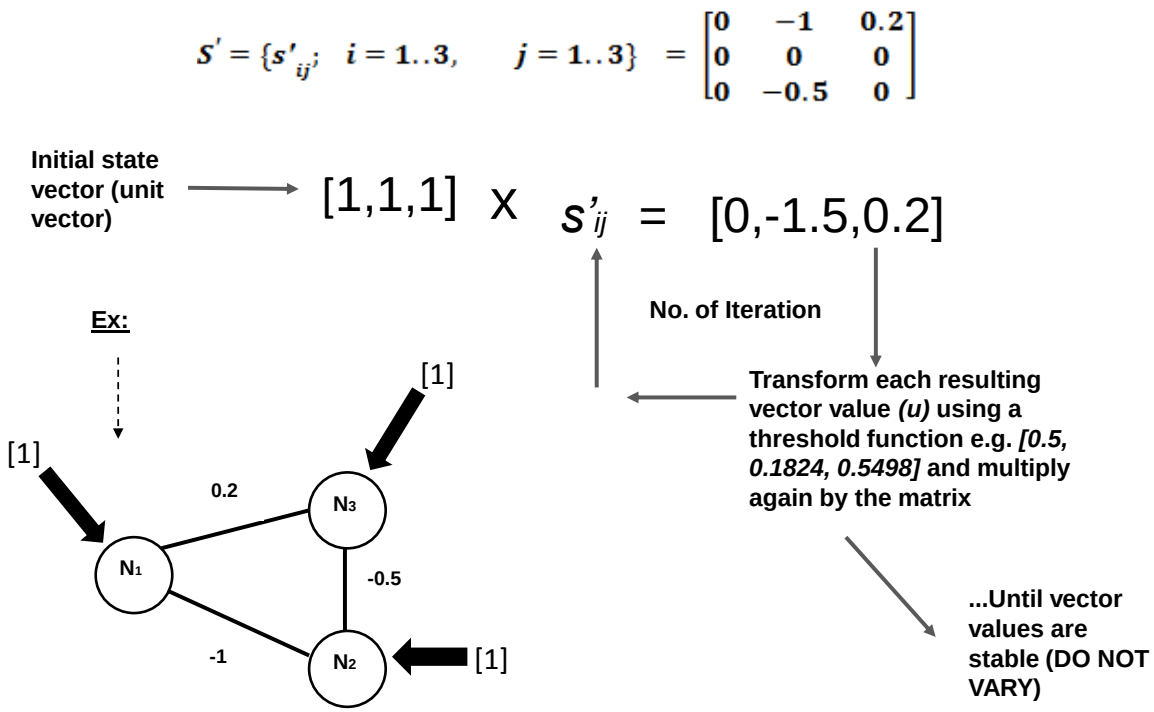
Table B.4: Summary of centrality index, inputs and outputs

Nodes	inputs	outputs	Centrality
N1= Indoor Temperature (cold)	0.00	1.20	1.20
N2= Cardio-Respiratory Conditions	1.50	0.00	1.50
N3= Psychosocial wellbeing	0.20	0.50	0.70

Step # 5 (Figure B.5): Pertub the system to identify causal processes. The perturbation analysis of the causal processes is guided through a deterministic iterative calculation procedure. The overall result of the perturbation analysis is presented in a graph and in a table as shown below.

The deterministic perturbation process is conducted in two iterative sub-steps which consist of: (1) matrix multiplications and (2) application of a threshold function. The two iterative sub-steps are explained as follows. An initial state vector (unit vector) is multiplied by a matrix $S'(s'_{ij}, i = 1..N, j = 1..)$ and then each element (u) of the resulting vector is normalised using a threshold function $f(u)$ to create a new normalized state vector $(V_j^{(t+1)})$. In subsequent iterations, the state vector $(V_j^{(t+1)})$ is repeatedly multiplied by the matrix S' until all state vector values reach an equilibrium (refer to the section "perturbing the system to identify causal processes" and equations [8] and [9] in Appendix A, for the mathematical details).

Figure B.5: Complex causal process in the perturbation analysis



Simulating causal processes consist in two iterative steps: (1) matrix multiplication by a state vector and (2) application of a threshold function.

The steps below give the details of the perturbation analysis calculations for the hypothetical example.

Matrix multiplication by a state vector:

$$V_j^{(t+1)} = f \quad \langle \sum_{i=1, j\neq 1}^{N} V_j^{(t)} \times (s'_{ij}) \rangle$$

where $V = 1 \times n$ is a state vector that contains the values of the nodes.

$$S' = \{s'_{ij};\ i = 1..3,\ j = 1..3\} = \begin{bmatrix} 0 & -1 & 0.2 \\ 0 & 0 & 0 \\ 0 & -0.5 & 0 \end{bmatrix}$$

Initial state vector (unit vector) $V = [1,1,1]$ represents weakly perturbed inputs assigned to each node.

Resulting vector = $V \times s'_{ij} = [1,1,1] \times \begin{bmatrix} 0 & -1 & 0.2 \\ 0 & 0 & 0 \\ 0 & -0.5 & 0 \end{bmatrix} = [0, -1.5, 0.2] =$

$[u1, u2, u3]$

Application of threshold function:

$$f(u) = \frac{1}{1 + e^{-u}}$$

$$N_1 = f(u1) = \frac{1}{1 + e^{-1\times 0}} = 0.5$$

$$N_2 = f(u2) = \frac{1}{1 + e^{-1\times -1.5}} = 0.1824$$

$$N_3 = f(u3) = \frac{1}{1 + e^{-1\times 0.2}} = 0.5498$$

$New\ state\ vector$ $\mathrm{V}^{(1)} = [0.5, 0.1824, 0.5498]$ = Iteration No. 1

$New\ state\ vector$ $V^{(2)} = [0.5, 0.556034, 0.524979]$ = Iteration No. 2

$New\ state\ vector$ $V^{(3)} = [0.5, 0.559100, 0.524979]$ = Iteration No. 3

.....Until stable pattern conditions (equilibrium) are reached ……

Table B.5 Summary of iterations

No. of iterations until stable conditions (equilibrium):	**1**	**3**	**2**
	Node 1	**Node 2**	**Node 3**
	1.000000	1.000000	1.000000
Iteration No.1	0.500000	0.182426	0.549834
Iteration No. 2	0.500000	0.315416	0.524979
Iteration No. 3	0.500000	0.318106	0.524979
Steady state	**0.500000**	**0.318106**	**0.524979**

Figure B.6: Causal processes between the nodes after perturbing the system

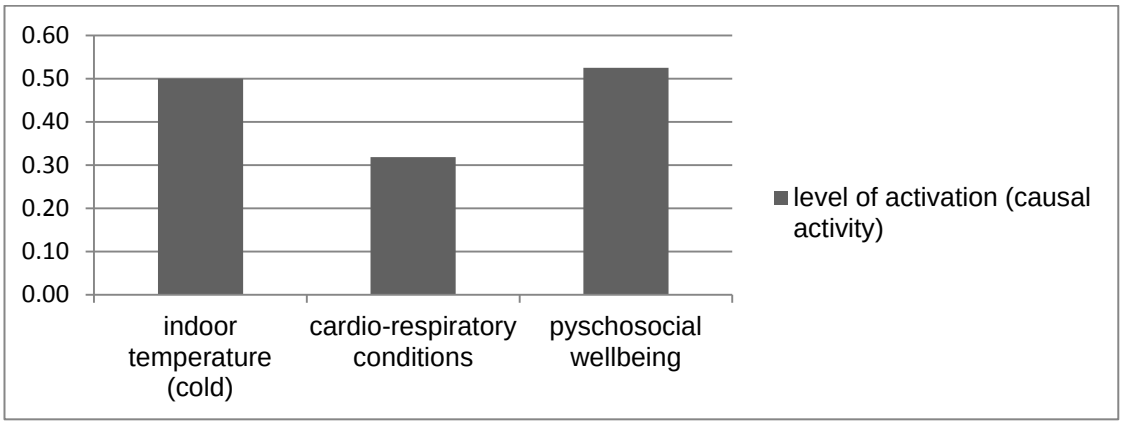


Figure B.6 shows the stable pattern reached at equilibrium. The figure describes the steady-state system behaviour taking into account the feedback processes and the pattern of causal mechanisms between the three variables: indoor temperature (cold), cardio-respiratory and psychosocial (wellbeing). Each element of the steady-state state vector represents the equilibrium level causal activity (level of activation) in the individual node (variable). The level of activation measures how each node influence one another over a number of iterations. The result of the perturbation analysis describes how the system functions based on its structural properties. To understand a causal system fully, it is important to analyse its structural properties as well as its dynamic causal behaviour.

4.3. Supplementary material to chapter 4 – Further analysis based on research paper 2

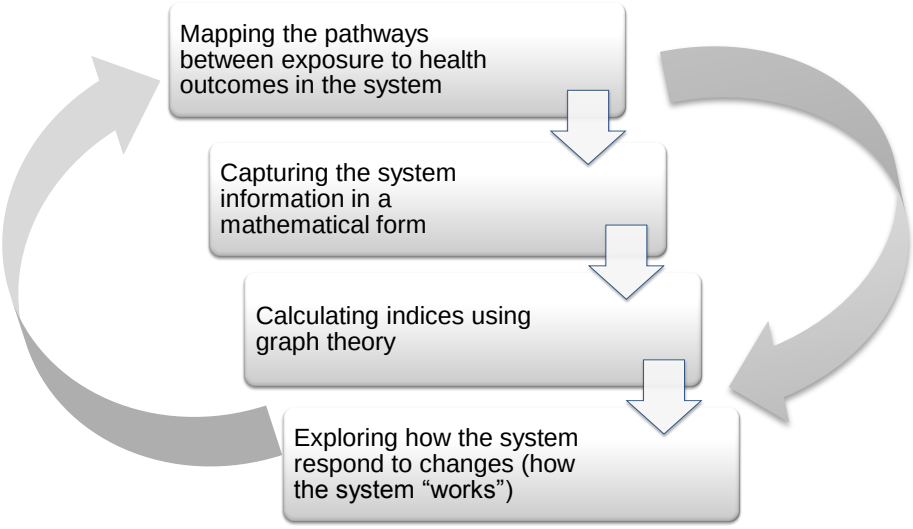
The HIA model applied to the case-study example addresses conceptual uncertainty in the mapping of the causal pathways to health outcomes based on a literature search. The emphasis of the proposed method is on the characteristics of the assessment at a system-level. In particular, the emphasis is on "how" the system works rather than "what" works. The method does not address detailed parametric observations and sources of heterogeneity (i.e. "what" works). Sources of heterogeneity that relate to seasonality or time where occupants are exposed, age, place of exposures (i.e. home, apartments, semi-detached homes), specific outcome measure (including seriousness of a disease) were not fully addressed in the method. This is a potential limitation of research paper 2.

There are many uncertainties to consider when modelling the impact of a housing intervention. The uncertainty in defining the causal pathways of exposure to population health outcomes was prioritised in the case study. The assumed casual pathways were explored through a perturbation analysis performed deterministically, rather than stochastically to capture some underlying mechanism. This is in contrast to a stochastic simulation where a random element is introduced in the input variables to demonstrate empirical association between the variables.

The emphasis of paper 2 was on the higher level conceptual sources of uncertainty which included: (i) mapping specific pathways of exposures to health outcomes, (ii) defining the direction of causality (positive or negative and potential magnitude) in

addition to (iii) determining "how" the system works, in other words how the system respond to changes based on the assumed causal pathways. The key steps to explore these uncertainties can be depicted in Figure 8, and they will be revisited in this section.

Figure 8 Key processes in exploring conceptual uncertainty in the assessment



The objective of revising these steps is to illustrate how the mathematical formulations relates to the case-study and how alternative formulations can be potentially better. These steps can be followed in an iterative manner or repeatedly (as depicted in Figure 8) . They can be described in various logical steps as follows.

Mapping the pathways between exposure to health outcomes

The review of the literature in research paper 2 identified the pathway of exposures linking housing insulation and health. Alternative pathway of exposures and determinants can be identified and formulated (in addition to research paper 2 definition of the assumed causal structure). This section defines a variety of

potential exposures in the indoor environment of relevance to public health and the potential effects on health. Most common environmental hazards affecting indoor environmental conditions and air quality can be indoor temperature, relative humidity, air particles, allergens, mould and radon. [93, 109, 110] The quality of the external physical environment can also play an important indirect role such as safety from crime, proximity to congested areas, transportation and places. [111]

In terms of the potential effect of a housing intervention, evidence seems to show association with physical and mental health improvement from insulating hourses. [73, 91, 112-115] Housing conditions and exposures associated with poor indoor air quality have been shown to have an effect on various health outcomes such as cardio-respiratory conditions, psychosocial well-being and general quality of life. [101, 109, 116, 117] In this section, the outcome measure is also defined broadly. Outcomes such as wheezing, throat irritation, bronchopneumonia, winter mortality and asthma are broadly classified as: *cardio-respiratory morbidity/mortality* ,and for mental health, *psychosocial well-being* is defined as an outcome. This level of resolution is deemed appropriate to test the plausibility of the causal structure given that the emphasis of the intervention is at the system-level.

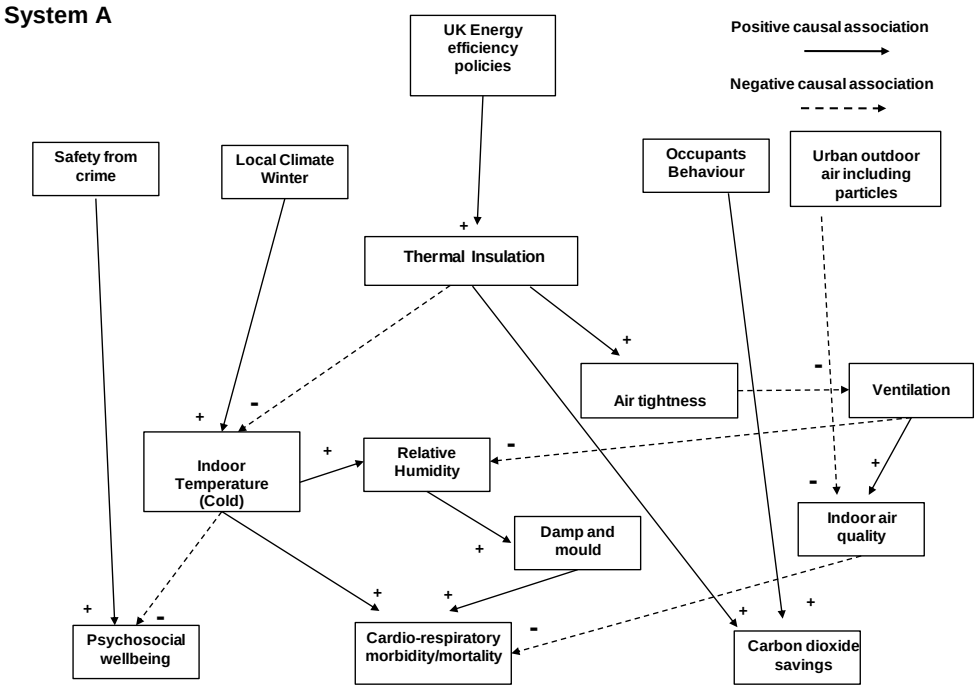
Capturing the information in a mathematical form

The system is represented mathematically by a matrix for the purpose of preserving and analysing its casual structure. In the case-study example of paper 2 each input in the matrix was re-scaled into the interval of -1 to 1 by dividing each elements of the matrix by the maximum absolute value as followed:[86]

$$a'_{ij} = \frac{a_{ij}}{\max(|a_{ij}|)}$$

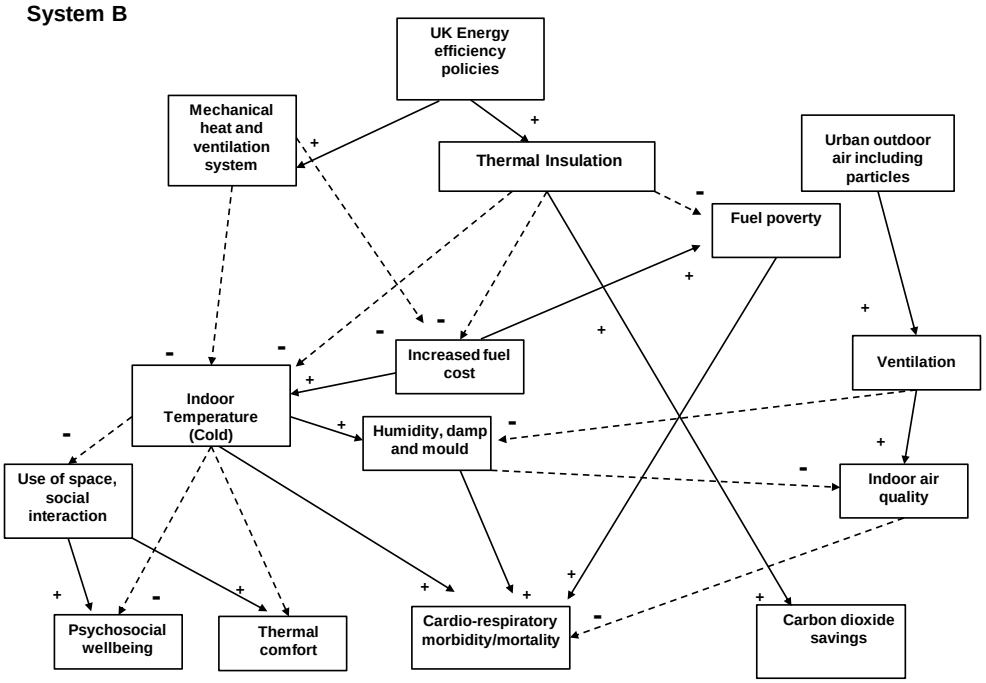
Figure 9A below illustrates an example of an assumed causal system on the basis of qualitative description taken from the literature.[118]

Figure 9A: "System A" defining and structuring the system (based on qualitative description from Bone et al 2010)[118]



A second representation of a causal system can be equally plausible by adding concepts such as "*fuel poverty,*" "*fuel cost*" and "*thermal comfort*." (Figure 9B)

Figure 9B: "System B" defining and structuring the system



Trivalent weights of -1, 0 and 1 are used in both systems to describe the assumed causal associations in order to simplify the illustration. However, different weights can be added to the "links" using numerical values between -1 and 1 to describe the different "strengths" of the assumed causal associations.

The resulting matrix for "system A" and "system B" are given below in Table 7A and Table 7B

Table 7A: Matrix from "System A"

Insulation	UK energy efficiency policies	safety from crime	Local Climate: Winter	Occupants behaviour	Urban outdoor air-including particles	Thermal insulation	Cold	Relative humidity	Air tightness	Ventilation	Indoor air quality	Damp and mould	Psychosocial wellbeing	Cardio-respiratory morbidity/mortality	Carbon dioxide savings
UK energy efficiency policies	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
safety from crime	0	0	0	0	0	0.	0	0	0	0	0	0	1	0	0
Local Climate: Winter	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Occupants behaviour	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Urban outdoor air-including particles	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
Thermal insulation	0	0	0	0	0	0	-1	0	1	0	0	0	0.	0	1
Cold	0	0	0	0	0	0	0	1	0	0	0	0	-1	1	0
Relative humidity	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Air tightness	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
Ventilation	0	0	0	0	0	0	0	-1	0	0	1	0	0	0	0
Indoor air quality	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
Damp and mould	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Psychosocial wellbeing	0	0	0	0	0	0	0	0	0	0	0	0	0	0.	0
Cardio-respiratory morbidity/mortality	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carbon dioxide savings	0.	0	0	0	0	0	0	0	0	0	0	0	0	0	0

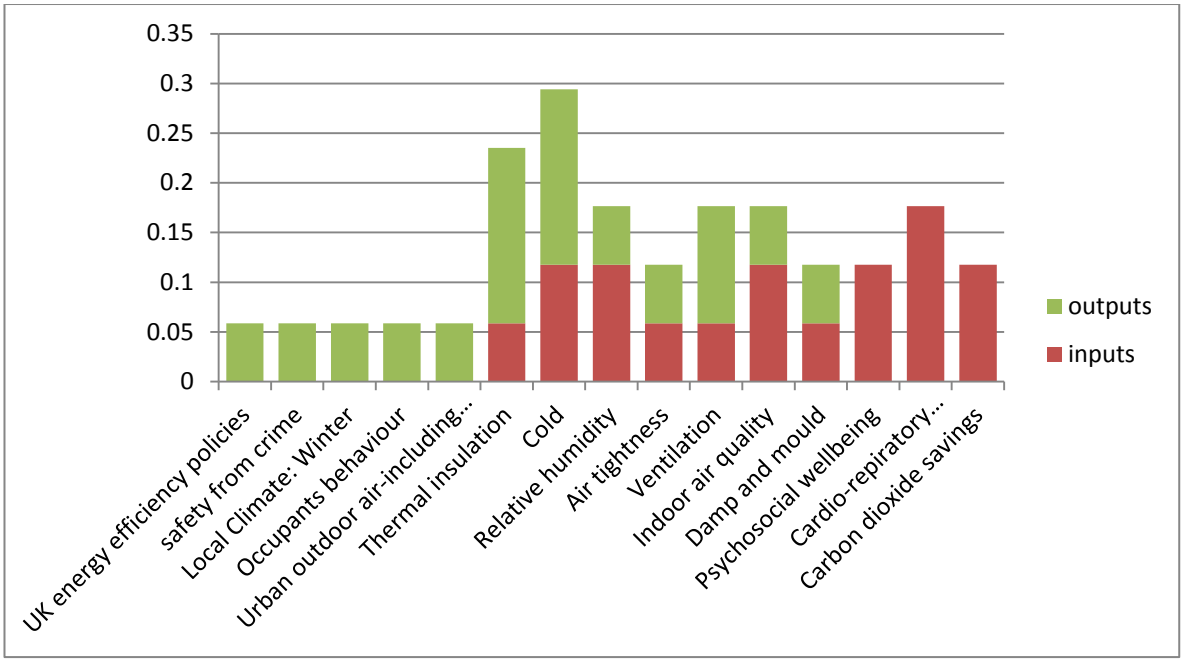
Table 7B: Matrix from "System B"

Insulation	UK energy efficiency policies	Mechanical Heat and ventilation systems	Thermal insulation	Urban outdoor air-including particles	Fuel Poverty	Cold	Increased fuel cost	Ventilation	Humidity, damp and mould	Indoor air quality	Use of space, social interaction	Psychosocial wellbeing	Thermal comfort	Cardio-respiratory morbidity/mortality	Carbon dioxide savings
UK energy efficiency policies	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Mechanical Heat and ventilation systems	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0	0
Thermal insulation	0	0	0	0	-1	-1	-1	0	0	0	0	0	0	0	1
Urban outdoor air-including particles	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Fuel Poverty	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Cold	0	0	0	0	0	0	0	0	1	0	-1	-1	-1	1	0
Increased fuel cost	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
Ventilation	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0
Humidity, damp and mould	0	0	0	0	0	0	0	0	0	-1	0	0	0	1	0
Indoor air quality	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
Use of space, social interaction	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
Psychosocial wellbeing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thermal comfort	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cardio-respiratory morbidity/mortality	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carbon dioxide savings	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Calculating indices using graph theory

Whilst the basis of the presumed causal structure for "system A" and "system B" are based on qualitative descriptions, the structural properties of each casual system can be compared quantitatively. Comparison between different systems can be made using graph theoretic indices. For example, a centrality index shows how well connected a variable (node) is in relation to other variables. A summary of the centrality index derived from “system A” are displayed in Figure 10A. Each bar in the figure at x-axis represents each variable with their corresponding value of the centrality index at the y-axis.

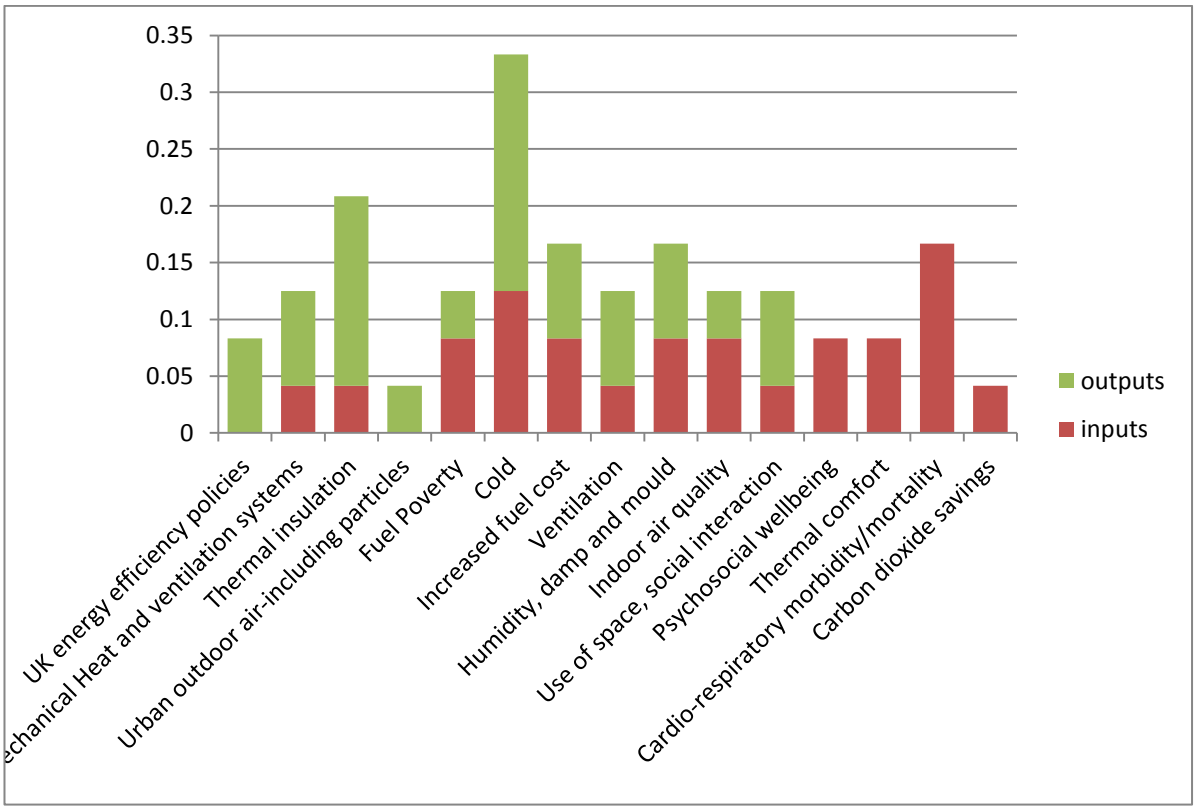
Figure 10A: Summary of most central variables for “System A*”



* Y axis normalised (divided by the total number of connection in each graph)

Centrality index in Figure 10B shows cold, thermal insulation, relative humidity, ventilation and indoor air quality as the most important variables in the system, given the assumed causal interpretations.

Figure 10B: Summary of most central variables for "System B"*



* Y axis normalised (divided by the total number of connection in each graph)

For "system B", the most important variables based on the assumed causal structure are *cold*, *thermal insulation*, *humidity,damp* and *mould*, *increased fuel cost* and *cardio-respiratory morbidity/mortality*.

Graph theoretic indices can help "include" rather than "exclude" the framing assumptions quantitatively or semi-quantitatively in the appraisal of uncertainty. The indices can provide a quantitative measure for comparison between different

systems. However, these comparisons so far are based on the *static* structural properties of the causal systems. Such comparisons are only made on the basis of their structural characteristics and not on their function. To provide guidance on which structure to use it is important to analyse how the system respond to changes.

Exploring 'how the system respond to changes

This part is concerned with how the casual pathways are propagated , in other words, 'how the system works" based on the structural assumptions. The term " structural assumptions" is used to define the pathways of exposures and causal interpretations. To explore "how the system works" or more accurately how the system responds to changes in a set of inputs, small simultaneous changes (perturbation) are given to each node. Each node is allowed to reach its maximum relative value of 1 (activation) so that once their feedbacks or underlying mechanisms are taken into consideration, the nodes can reach a stable pattern. To explore the feedbacks and underlying mechanism, a perturbation analysis is conducted as follows (it is also described in Appendix A): an input state vector (V^1) is multiplied by the connection matrix $(A_{i,j})$ to generate a new vector (V^{t+1}). The resulting vector $(V_j{}^t)$ is repetitively multiplied by $(A_{i,j})$ until the values of each vector are stable.[14, 86]

$$V_j^{(t+1)} = f \quad \langle \sum_{i=1, j\neq 1}^{N} V_j^{(t)} \times (s'_{ij}) \rangle$$

Each result value (x) of the vector is kept within the interval of [0, 1] by a applying a threshold function $f(u)$[14, 108] after every iteration as follows:

$$f(u) = \frac{1}{1 + e^{-u}}$$

The stable condition describes the system feedback and causal mechanism under baseline scenario. Each resulting value of a vector represents the level of activation (causal activity) in the individual nodes. The causal activity reflect*s how each node influence one another over a number of iterations*. It is important to emphasise that the *causal activity* measures how the system function or behaves based on its characteristics, in other words, based on its assumed centrality (i.e. the centrality index). The *centrality index* measures the static properties of the system and the causal activity measures its function based on its static properties. The centrality index and the causal activity are considered two distint measures.

The following results indicate how the structural assumptions affect how the system operates. Figure 11A-B shows the causal activity (level of activation) of each node in "system A" and "system B" once their feedbacks are taken into account.

Figure 11A: Feedbacks mechanism under baseline scenario "System A"

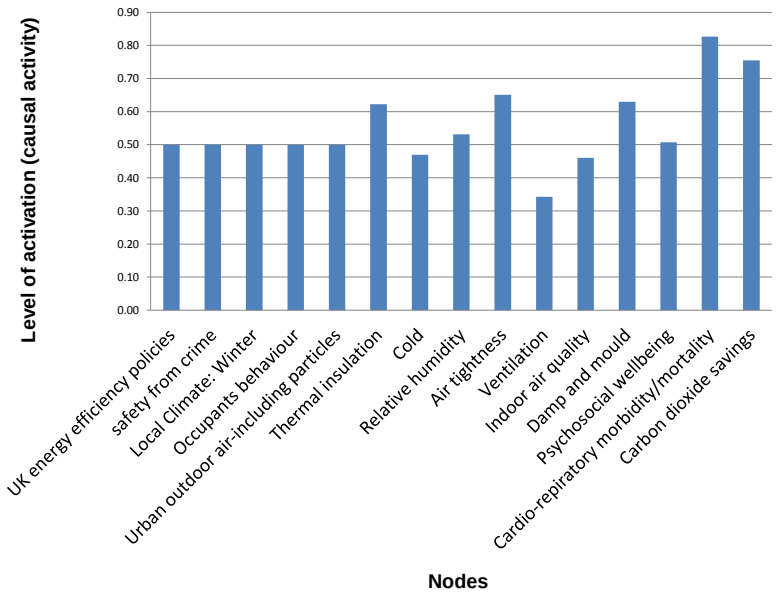
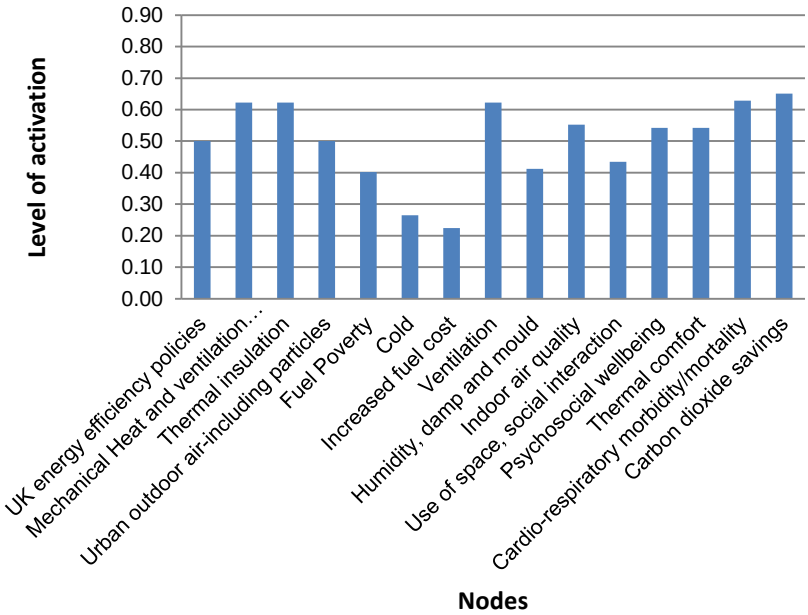


Figure 11B: Feedbacks mechanism under baseline scenario "System B"



The illustration results demonstrate a higher feedback response in *"cardio respiratory conditions"* relative to other variables within the "system A" structure. On the other hand in "system B", *carbon dioxide savings* displays a higher feedback

response relative to other variables such as *cardio respiratory conditions* and ventilation. The result of the perturbation analysis from each structural assumption can provide guidance on which structure to use depending on the policy question to be adopted. The result from "system A" demonstrates both a higher feedback response in *'cardio respiratory conditions'* and *'carbon-dioxide savings'* relative to other nodes in the system. A high level of activity in both nodes can be explained by the causal assumptions made in "system A" measured by the centrality index. Both nodes (*cardio respiratory conditions* and *carbon dioxide savings*) are assumed to be outcomes variables in the system. Outcome variables are expected to have a high level of causal activity given the direction of the causal associations pointing towards them. On the basis of the feedback mechanisms, "system A" is chosen for further for exploration in this example given its higher feedback on the outcome of interest.

Furthermore, the resulting vector of activation obtained under baseline conditions "system A" (V_j^{base}) can be used to make comparison between scenarios. Policy induced changes can be hypothesised based on how the system will respond to future changes. UK energy efficiency policies are expected to become more and more stringent in the future to meet carbon reduction targets. Supposing, the UK government decides to insulate all existing homes to achieve zero net carbon emission. An 'insulating all existing homes' scenario can represent the maximum value the node *'UK energy efficiency policies'* can obtain in "system A". Mathematically, this is denoted by keeping the node during the entire iteration process of the perturbation analysis to a constant numerical value of 1 (maximum value). The resulting vector obtained {UK energy policies" (V_j^{UK})} is compared

against the {baseline scenario"(V_j^{base})} as follows. A relative change (Δ) is used to compare both scenarios:

$$\Delta = \frac{(V_j^{UK} - V_j^{base})}{V_j^{base}}$$

The extent to which the UK energy efficiency scenario influences the system is measured by noting the value of relative change (Δ) from baseline scenario. The relative change (Δ) in this example shows how the system responds to changes in UK energy efficiency policies scenario compared to the baseline scenario (no change) as shown in Figure 12 below.

Figure12: UK energy policies change scenario

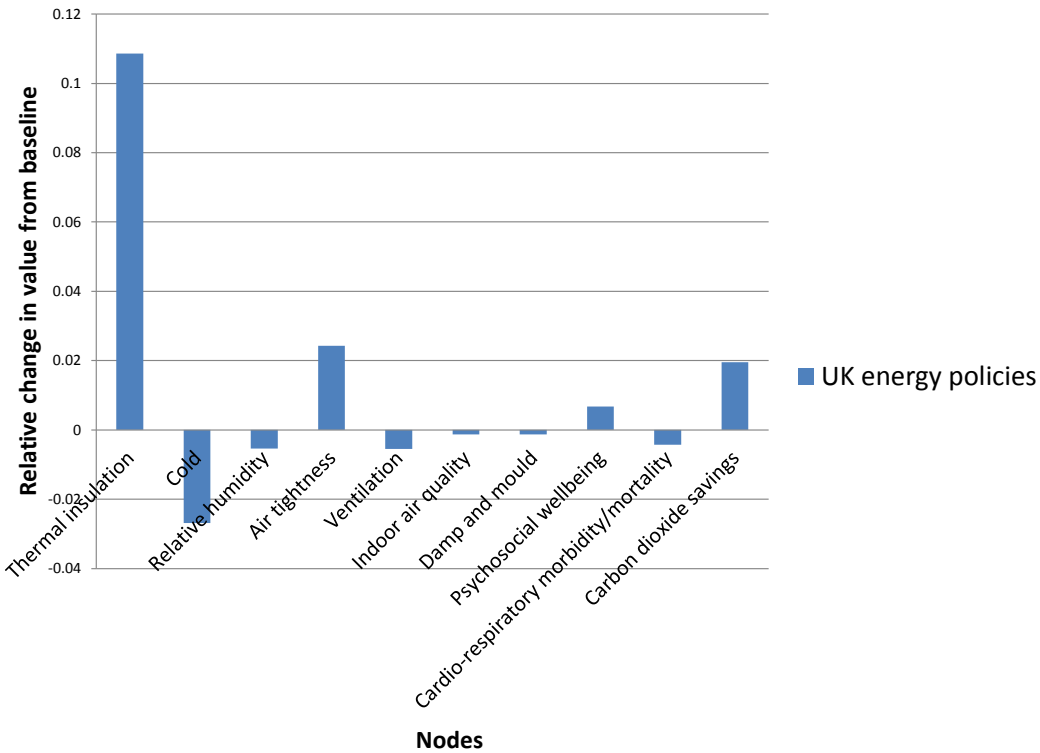


Figure 12 shows the relative change in the nodes (change above/below average in relation to each other), with an increase in thermal insulation, carbon dioxide savings, air-tightness, and a decrease in cold. Additionally, a decrease in cold, relative humidity and ventilation are shown as a result of the change-model scenario (UK energy efficiency policies). Results of such causal changes can be qualitatively validated for consistency and tested whether these changes are expected if the UK government insulate all existing homes. Such results should be interpreted in relative terms and qualitatively. These values represent qualitatively the theoretical value of what the node could measure in reality (e.g. thermal insulation = change in W/K.m in all dwellings from baseline, air tightness = change in m3/(m2·h) @50 Pa from baseline, carbon dioxide savings = change in CO2 ppm from baseline and psychosocial well-being = change in Ryff Scales from baseline). However, the FCM is not a tool suitable for parameter estimation but rather a conceptual modelling tool. The perturbation analysis can test the plausibility of the structural assumptions (measured by the centrality index) based on how the system operates. Each causal structure represents a framing assumption and a perception "how the system works". Through the perturbation analysis, the causal activity (behaviour) of a system can only be observed. It is important to observe the causal activity of a system to test its structural assumption. According to Pearl et al. (2000),[119] a system's true causal structure can only be explicitly recognised by fully changing the state (perturbation) of each node and observing the consequence. This can be explored by initially assigning that maximum value to each node in a unit vector. In reality, the maximum value represents an extreme theoretical value of what the node could measure (e.g. indoor cold 1= lowest °C; ventilation 1= highest ACH). Each causal structure can be

compared and tested for qualitative consistency based on the result of the perturbation analysis.

It is worth noting that the process of perturbation analysis might not always converge in some causal systems, particularly when modeling a casual system involving many feedback links. However, more important than knowing whether a system reaches a steady state condition, it is the guaranteed that the modeling process does not change the stability of the system being model. In other words, that the modeling process is consistent in capturing the essence of the qualitative system without oversimplifying its complexity. In addition, the pattern of the causal mechanism in the perturbation analysis can provide the basis for the selection of variables. The level of activation in each node highlights how each variable ranks in relation to each other once we allow each node in the system reach their maximum value. The perturbation analysis in practice can be used for model reduction, where the most important variables are retained based on: (i) their level of causal activity in their feedbacks and (ii) the result of the centrality index. For example, from "system A" results, 9 variables can be retained (out 15 variables) such as insulation, airtightness, damp and mould, relative humidity, ventilation, cold, indoor air quality, carbon dioxide savings and cardio-respiratory morbidity/mortality.

5. Analytical perspective – propagating lack of knowledge or limited information as part of analytical uncertainty

5.1. Preamble to research paper 3 – HIA specific question to Analytical uncertainty

Research paper 2 dealt with one aspect of conceptual uncertainty associated with the framing assumptions in a HIA of housing insulation. Changes in ventilation rates (among other factors) were identified in research paper 2 as being sensitive to the framing assumptions in the causal pathways. Another aspect of uncertainty identified in the framework of the thesis is the "analytical uncertainty" associated with the input parameters and outputs of a HIA model.

Research paper 3 focuses on analytical uncertainty and estimates the uncertainty in two input parameters: the population exposure scenarios and the exposure-response function of a HIA model. Lack of understanding or information is assumed to be the primary source of analytical uncertainty. The HIA model assumes lack of information in relation to: the definition of exposure scenarios, the distribution of the population to the exposure scenarios, the extrapolation of the exposure-response function to different subpopulation, geographical location and the assumption of a linear threshold above and below a particular value. The uncertainty is characterised via fuzzy sets defined based on evidence from a literature search and propagated using interval analysis calculations. Research paper 3 provides an analytical framework for quantifying health impacts and handling analytical uncertainty. The analytical framework is applied to a case-study example of indoor housing

ventilation exposure scenarios in England. The case study addresses the following HIA question.

What would be the population health burden if the population of England were exposed to different housing ventilation scenarios?

The above question is defined broadly to include the entire population of England. To a limited extent the complex set of issues of parametric uncertainty in the case-study example are only addressed in research paper 3. Ventilation rates is modelled as the only causal variable in the case-study, and other casual factors such as sources of indoor air pollution in homes, time spent indoors, geographical locations with sources of outdoor air pollution in rural or urban areas are not accounted in the HIA model. Most HIA tend to be qualitative and quantifying the health outcomes and their associated uncertainty in a HIA is still limited in practice. [120] Research paper 3 attempts to reduce the gap in current HIA by providing an initial framework to quantify the health impacts and associated uncertainty.

5.2. Research Paper 3

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Marco Mesa Frias
Principal Supervisor	Zaid Chalabi
Thesis Title	Modelling Uncertainty in Environmental Health Impact Assessment

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Environment International		
When was the work published?	November 2013		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	See attached cover sheet

Student Signature: _____ **Date:** 5/10/2015

Supervisor Signature: _____ **Date:** 5/10/2015

Improving health worldwide **www.lshtm.ac.uk**

Research Paper 3

Quantifying uncertainty in health impact assessment: A case study example on indoor housing ventilation

Marco MESA-FRIAS[1]; Zaid CHALABI[1]; Anna M. FOSS[2]

[1]Department of Social and Environmental Health Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.
[2]Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.

Status: Published in Environment International doi:10.1016/j.envint.2013.10.007
Corrigendum also published in Environment International
doi:10.1016/j.envint.2014.03.001

Contributions: The candidate led in the conception of the research question in collaboration with Zaid Chalabi. The candidate developed the framework, conducted the literature review, discussed the result and drafted the manuscript. Anna M. Foss contributed to the review, providing comments and suggestions. Zaid Chalabi provided comments in the interpretation of the results. The candidate wrote the first draft of the manuscript, and managed each round of comments and suggestions from co-authors. All the authors read and approved the final draft prior to journal submission and inclusion in the dissertation.

The candidate_____ The supervisor_____

Abstract

Quantitative health impact assessment (HIA) is increasingly being used to assess the health impacts attributable to an environmental policy or intervention; and/or the burden due to present conditions. As a consequence, there is a need to assess uncertainties in the assessments because of the uncertainty in the HIA models. In this paper, a framework is developed to quantify the uncertainty in the health impacts of environmental interventions or exposures scenarios and is applied to evaluate the impacts of poor housing ventilation. The paper describes the development of the framework through three steps: (i) selecting the exposure metrics and quantifying the evidence of potential health effects of the exposure; (ii) estimating the size of the population affected by the exposure and selecting the outcome measure; (iii) quantifying the health impact and its uncertainty. The framework introduces a novel application for the propagation of uncertainty, based on fuzzy set theory. Fuzzy sets are used to propagate parameter uncertainty in a non- probabilistic space and are then used to calculate the uncertainty in the morbidity burdens associated with three indoor ventilation exposure scenarios: *poor*, *fair* and *adequate*. The case-study example demonstrates how the framework can be used in practice, to quantify the uncertainty in health impact assessment where there is insufficient information to carry out a probabilistic uncertainty propagation.

Keywords: Environmental Health, Risk Assessment, Modeling, Health Impact Assessment, Housing.

Introduction Health Impact Assessment (HIA) evaluates prospectively the health impacts attributable to an environmental policy or intervention and or the burden of disease due to present conditions. HIA requires sources of evidence and a number of analytical tools are available for the estimation of health impacts that range from qualitative to quantitative methods. To date, most HIA methods have been qualitative rather than quantitative. Although some quantitative HIAs have been conducted in the past, [30, 54, 68, 69] their take-up has been slow. The quantification of health impacts in a HIA has desirable features for decision support. It provides a measure of the magnitude of health consequences of an environmental policy or intervention. Also, it can help decision-makers evaluate the significance of the potential health impacts based on the assessment before a policy or an intervention is implemented.

Although quantifying the health impacts is desirable in HIA, such quantification can be met with limitations in practice. [121] Quantifying the health impacts requires the knowledge of various measures such as exposure-response functions (or relative risks), location and size of the population affected, and the distribution of exposure over the affected population. Limitations on conducting a quantitative HIA can occur due to lack of information on the above measures or lack of evidence on the causal pathways linking changes in exposure with health outcomes. Such limitations, commonly characterised by "lack or imprecision in knowledge", can be an important source of uncertainty in the quantification of health impacts. [3]

Uncertainty is inherent in most environmental HIA, partly due to lack of understanding of the associations between environmental exposures and health

outcomes, or due to random variations in these associations. [4] Most approaches for quantifying uncertainty in environmental HIA models cannot deal with uncertainty due to lack of knowledge. [78] It is important to note that lack of knowledge yields to imprecision in parameters. Most probabilistic approaches assume that the uncertainty in model parameters is due to random variations and they characterise the uncertainty in model parameters using probability distributions. However, random variation in model parameters is only one type of uncertainty in environmental HIA. Uncertainty in model parameters might also arise from limitations in knowledge (or incomplete data), and it is important to incorporate methods that can deal with the uncertainty due to this limitation. As such, this paper provides an alternative non-probabilistic approach to incorporate parameter uncertainty due to imprecision in knowledge using an application of fuzzy set theory, which is novel in health impact assessment. Fuzzy set theory provides a method for characterising uncertainty in a non-probabilistic space. Fuzzy set theory is a method that does not require knowledge of statistical properties of parameters such as its mean, variance or correlations to propagate its uncertainty, which makes it an ideal method to handle uncertainty that might arise due to imprecision in knowledge or incomplete data. [20]

We believe that a more comprehensive examination of HIA for handling the uncertainty in the quantification of health impacts is required. As methods for the quantification of health impacts are beginning to take-up, [53, 122, 123] this paper adds to this literature by developing and applying a new HIA modelling framework to quantify the health impacts and its uncertainty. In this paper, we will focus on parametric uncertainty. Other issues associated with uncertainty such as the formulation of a model or framing assumptions are addressed elsewhere. [124] Our

approach involves the development of a case-study example and the application of the HIA framework in three sequential steps: (i) selecting the exposure metric and quantifying the evidence of potential health effects of the exposure, (ii) estimating the size of the population affected by the exposure and selecting the outcome measure, and (iii) quantifying the health impacts and associated uncertainty. The framework is demonstrated through a HIA case study which examines the health impact of housing ventilation in England.

Housing ventilation case-study

Housing energy efficiency measures, and changes in building designs are currently implemented as part of the UK government's effort to reduce carbon greenhouse emissions and energy cost from domestic sources. UK government initiatives require improvements in insulation retrofits to avoid heat loss and encourage energy savings. [71, 125] However, there are concerns regarding changes in building designs retrofits and energy efficiency measures because they can potentially reduce indoor ventilation rates due to an increase of air-tightness. [126, 127] It is worth noting that ventilation needs are not always considered when assessing the performance of energy-efficiency interventions, and some studies suggest that a majority of newer airtight energy efficiency homes are under-ventilated. [128] It is important therefore to ensure an adequate ventilation level in dwellings for better health and well-being. In the next section, we explore how indoor ventilation can affect health through the development of a quantitative framework.

Methods

Quantitative framework for HIA

In general, the key steps for quantifying health impacts in a HIA include: (1) selecting the exposure metric and quantifying the evidence of the potential health effects of the exposure; (2) estimating the size of the population affected by exposure and selecting an outcome measure; (3) quantifying the health impacts and associated uncertainty. The steps are applied to the case-study of housing ventilation as follows.

Step 1: Selecting the exposure metric and quantifying the evidence of potential health effects of the exposure

Adequate ventilation is required to remove indoor pollutants, with several studies having associated poor indoor ventilation with negative health outcomes. [95, 129-131] Common negative health outcomes reported due to poor ventilation exposure include allergies, rhinitis, asthma, wheezing, among others. Several qualitative reviews have concluded that a minimum ventilation rate of 0.5 air changes per hour (ACH) is required for health reasons. [132-136] However, these reviews have not produced quantitative *summary* estimates associating poor indoor ventilation and health.

Currently most quantitative studies rely on different experimental intervention studies, to provide estimates of an association between ventilation rate and health. Some experimental intervention studies have provided inconclusive results due to limitations in the size of the population, measurement methods of ventilation, and the diversity of geographical locations and climate. [137-140] No previous study has provided quantitative *summary* estimates based on epidemiological study design. It is important to review the evidence based on epidemiological studies, with studies

that have adjusted for key confounders, to assess limitations and provide a quantitative summary estimate. We conducted a systematic review and a meta-analysis, as an initial step towards quantifying the evidence and determining the strength of the association between poor ventilation rates and health outcomes.

Systematic search and meta-analysis

A systematic search was conducted in the Ovid Medline academic database from inception (-1948) through to August 2012, using the following free-text search string: "Ventilation" OR "Ventilation Rate" OR "Air flow*" OR "Air exchange*" AND "Health" OR "Sick Building*" OR "Allergy*" OR "illness*" OR "Asthma" AND "Housing" OR "Home" OR "Apartment" OR "Dwelling" OR "Building" OR "Residence" in the title and the abstract. Details of the search strategy are shown in Appendix A. Papers were screened according to the following inclusion criteria: (i) studies published in peer-reviewed articles in English; (ii) original studies that used primary data (e.g. not reviews, commentaries, etc.); (iii) studies that provided a measure of effects (e.g. odd ratios or relative risks, hazard ratios); (iv) only studies of cohort, cross-sectional or case-control study design were included; (v) studies which defined health outcomes and measurement of ventilation. Studies meeting the inclusion criteria were carefully examined, and their main characteristics were recorded. The following information was extracted from included studies: authors, year of publication, study design, geographical location, study population, building setting (offices, residences, schools), sample size, health outcomes assessed, ventilation exposure measurement, degree of adjustment and effect estimates for a given ventilation exposure category. Ventilation exposures were defined and classified into two categories: "low ventilation" for ventilation rates below 0.5 air

changes per hour (ACH), and "reference ventilation" for ventilation rates above or equal to 0.5 ACH.

Studies presented different effect estimates (e.g. relative risks, odds ratio, and hazard ratio) alongside several types of risk comparison groups for measures of ventilation exposures. We standardised the effect estimates and the different types of risk comparison into a log scale assuming a log-linear relationship of health symptoms with ventilation category. Risk comparisons were defined into two categories: ventilation rate greater than 0.5 ACH ("reference group") versus lower ventilation rate less than 0.5 ACH ("exposure group"). The natural log of the effect estimates and standard errors were calculated from the published studies estimates and confidence intervals (CIs). Odds ratio (ORs), using random-effects models, and 95 % CI were used to represent the final quantitative summary estimate and associated uncertainty. In addition, quality scoring or weighting of studies was not performed because quality scoring can introduce some bias. [141] We instead assessed heterogeneity using subgroup analysis to examine the sensitivity of different aspects of the studies had on final study results (Appendix A).

Step 2: Estimating the size of population affected by exposure and selecting outcome measure

For the population affected by the exposure, we identified the total population of England up to mid-2011 projections from the UK Office of National Statistics (ONS) data. In terms of outcome measures, common symptoms in relation to poor ventilation exposures were identified through the systematic review and meta-analysis from Step 1. In this case-study, we defined the outcome measure as

respiratory-related morbidity to describe the range of symptoms associated with poor ventilation exposure. Based on this outcome definition, we identified data from the Health Survey for England (HSE) - 2010 report on respiratory health to obtain estimates of the total annual number of existing respiratory-related morbidity cases in England. [142]

Step 3: Quantifying the health impacts and associated uncertainty

We calculated the health impacts of poor ventilation and associated uncertainty as part of the case-study. This step involved quantifying the percentage increase in morbidity risk (i.e. excess morbidity risk) due to poor ventilation exposure and estimating the excess annual number of cases by comparing the disease burden from three theoretical population exposure scenarios. The methods are briefly described below.

Firstly, the excess morbidity risk due to poor ventilation exposure per ACH below threshold (0.5 ACH) was quantified by calculating the natural logarithm of the odds ratio and its 95 % CI, obtained from the result of the meta-analysis in the previous step. Secondly, the health impacts of housing ventilation were estimated by comparing the disease burden (i.e. annual respiratory-related morbidity cases) attributable to the exposure under three ventilation exposure scenarios: (i) *poor* ventilation with ventilation rates less than 0.48 ACH, (ii) *fair* ventilation with ventilation rates between 0.19 ACH and 0.77 ACH, and (iii) *adequate* ventilation with ventilation rates of at least 0.48 ACH and above. These exposure scenarios were classified according to ventilation standards for indoor air quality. [143]

Fuzzy set approach to uncertainty

In the absence of sufficient information to quantify probabilistically the uncertainty in a parameter, fuzzy set theory can be used for this purpose. In general, fuzzy set theory has been used to quantify parameter uncertainty in a non-probabilistic space. [19, 20, 144] Fuzzy sets are defined by a *membership function* that measures the "degree" (between zero and unity) to which a parameter value belongs to a set. [144, 145] Fuzzy sets can also be defined by their *lower/upper α-cut* bounds, which defines the interval of a fuzzy set. In this case-study, fuzzy sets were used to characterise the imprecise nature of the definition of each ventilation exposure scenario and the uncertainty in the log odds ratio obtained from the meta-analysis in the previous step. In order to perform common arithmetic operations with fuzzy sets such as multiplication, division, subtractions and other operations, interval arithmetic was used. Interval arithmetic using fuzzy sets performs arithmetic operations with interval values by specifying a lower and upper α-cut bound to determine the minimum, and maximum values of the interval. We present the mathematical definitions of a fuzzy set, the membership function and the lower/upper α-cut bounds, followed by their illustration.

Definitions

A fuzzy set is described mathematically as: [144]

$$A(x) = \{x, \mu_A(x) |\ x \in X \ and\ \mu_A \in [0,1]\}, \qquad [1]$$

where x is an element of the set X; $A(x)$ is a fuzzy set of X; $\mu_A(x)$ is the membership function. The membership function of fuzzy set $A(x)$ can be given by:

$$\mu_A(x) = \begin{cases} 0, & if\ x < a_1 \\ \left|\frac{x - a_1}{a_2 - a_1}\right| & if \quad a_1 \leq x \leq a_2 \\ \left|\frac{a_3 - x}{a_3 - a_2}\right| & if \quad a_2 \leq x \leq a_3 \\ 0, & if \quad x > a_3 \end{cases} \qquad [2]$$

where $a_1, a_2, a_3,$ are real numbers. The values of $\mu_A(x)$ range from 0 to 1, where "1" denotes full membership of the set and "0" denotes no membership. By membership we refer to the *degree* in which a value belongs to a set. For example, "the closer $\mu_A(x)$ is to 1, the more likely is that an element of x belong to A, and the closer $\mu_A(x)$ is to 0, the less likely is that an element of x belong to A"

Additionally, a fuzzy set is defined by specifying its lower and upper α-cut bounds as follows, for $\alpha \in (0,1)$ and $a_1 \leq a_2 \leq a_3$, we obtain:

$$A_\alpha := [A_L(\alpha), A_U(\alpha)] \qquad [3]$$

$$= [(a_2 - a_1) \times \alpha + a_1, \quad -(a_3 - a_2) \times \alpha + a_3], \qquad [4]$$

where A_α is the α-cut bounds of A, which describes an interval of confidence at level α whose membership values are greater than the value at α. The lower bound of the interval is defined as $A_L(\alpha) = inf\{x \in \mathbb{R}: A(x) \geq \alpha\}$; and the upper bound of the interval is defined as $A_U(\alpha) = sup\{x \in \mathbb{R}: A(x) \geq \alpha\}$ where the terms inf and sup mean respectively the greatest lower bound and the lowest upper bound. See below the description of the Figure 13B for more detail.

Interval arithmetic operations with fuzzy sets are approximated using the α-cut bounds for each $\alpha \in (0,1)$. Arithmetic operations are given in a general form as:

$$(A \otimes B)_\alpha = A_\alpha \otimes B_\alpha, \qquad [5]$$

where $\otimes = +, -, * \; or \; /$ are basic arithmetic operations and A, B are arbitrary fuzzy sets.

For example, addition operations using fuzzy sets are given in a general form by

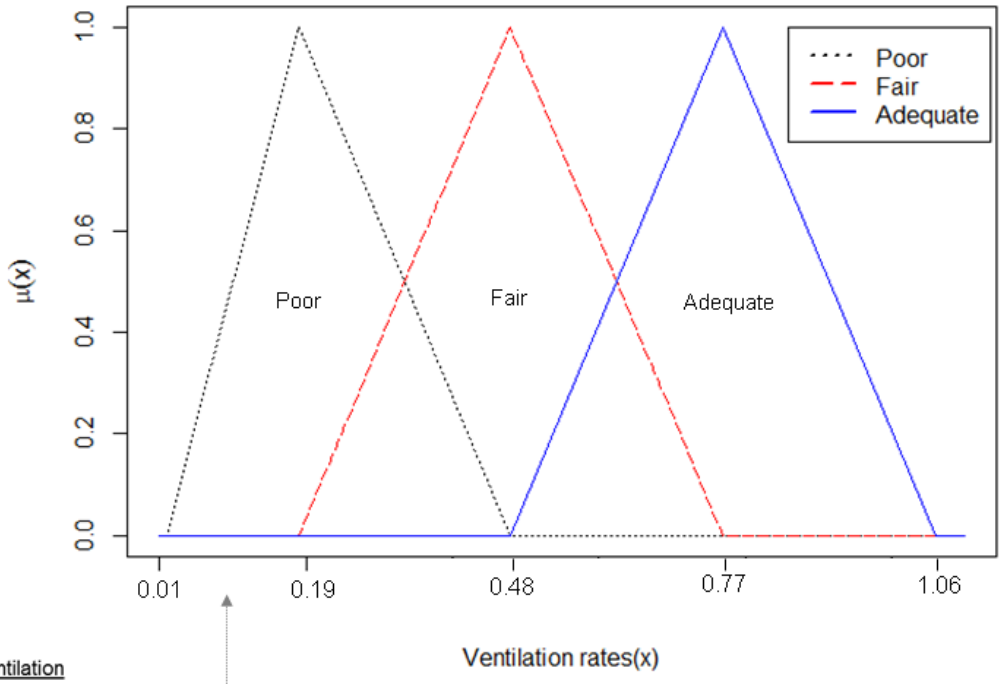
$$(A + B)_\alpha = A_\alpha + B_\alpha = [A_L(\alpha) + B_L(\alpha), \quad A_U(\alpha) + B_U(\alpha)] \qquad [6]$$

Details of interval arithmetic operations using fuzzy sets are shown in Appendix B.

For ease of understanding, we present Figure 13 to explain the mathematical definitions and operations of the fuzzy set approach. Figure 13A illustrates the concept of a fuzzy set, and its membership function. The x-axis displays the ventilation rate (air changes per hour - ACH) and the y-axis displays the degree of membership. X is the set of all feasible ventilation rates, and x is a single ventilation rate which belongs to this set. Three subsets of X are shown in this figure: "*poor* "ventilation, "*fair*" ventilation and *"adequate"* ventilation. The dotted, dashed and continuous lines define respectively the membership functions of *poor*, *fair* and *adequate* ventilation sets. To explain the concept of a membership function, consider the *poor* ventilation fuzzy set. The poor ventilation set is defined mathematically with equation [2]. In this set, ventilation rates 0.19 ACH belong unequivocally to this set. As the ventilation rate increases above 0.19 ACH, the degree of membership

of the *poor* ventilation set decreases linearly until it reaches zero at $x = 0.48$ ACH. Conversely, as the ventilation rate decreases below 0.19 ACH, the degree of membership of the same set decreases linearly until it reaches zero at $x = 0.01$ ACH. Figure 13B illustrates the concept of a fuzzy set and its interval arithmetic operations using the α-cut bounds. The x-axis in the figure displays *poor* ventilation set A and *fair* ventilation set B and their summation set $(A + B)$. The y-axis displays the α -cut of the fuzzy sets. To explain interval arithmetic using the α-cut bounds, consider the fuzzy set A. The lower and upper α-cut bound of fuzzy set A is defined analytically with equation [4] to preserve the interval form of the fuzzy set during arithmetic operations. The lower bound $A_L(\alpha)$ describes the interval values (e.g. 0.01 to 0.48) of the fuzzy set when $\alpha = 0$ and the upper bound $A_U(\alpha)$ represents the centre value (e.g. 0.19) of the fuzzy set when $\alpha = 1$. Fuzzy set A is added with Fuzzy set B using equation [6]. The α -cut bounds of the resulting fuzzy set $(A + B)$ are obtained by substituting the values "1" and "0" for α in the equation. Further details of interval arithmetic operations using fuzzy sets are shown in appendix B.

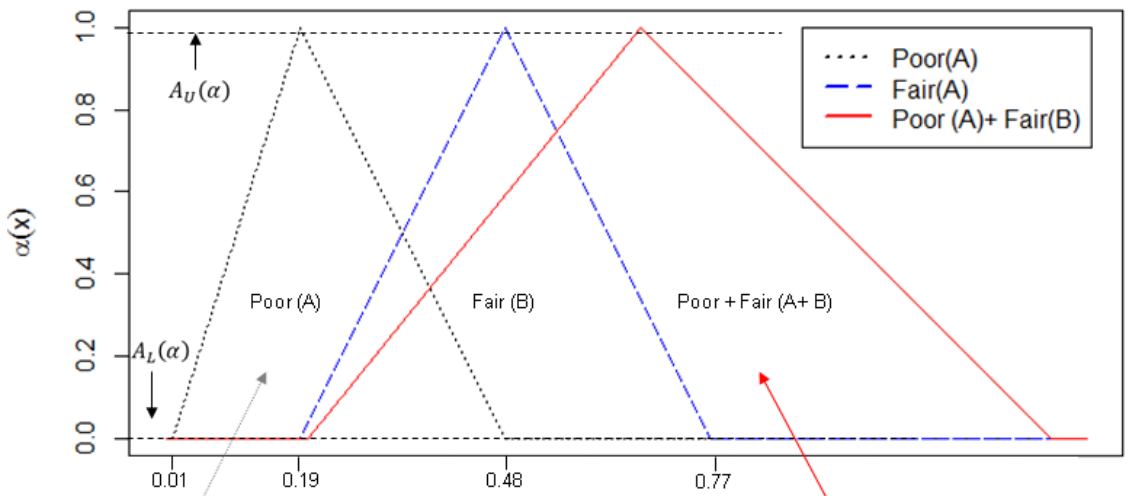
Figure 13A: Example graphical representation of fuzzy sets with ventilation exposure scenarios.



Ex: Poor ventilation

$$\mu_{poor}(x) = \begin{cases} 0, & if \;\; x < 0.01 \\ \left|\dfrac{x - 0.01}{0.19 - 0.01}\right| & if \;\; 0.01 \le x \le 0.19 \\ \left|\dfrac{0.48 - x}{0.48 - 0.19}\right| & if \;\; 0.19 \le x \le 0.48 \\ 0, & if \;\; x > 0.48 \end{cases}$$

Figure 13B: Interval arithmetic operation with fuzzy sets using the lower/upper α-cut bounds.



Fuzzy interval arithmetic operations

Ex: Lower/upper $\alpha - cut$ bounds for poor:

$A_\alpha = [A_L(\alpha), A_U(\alpha)] = [(0.19 - 0.01)\alpha + 0.01, -(0.48 - 0.19)\alpha + 0.48]$

$A_\alpha = (0.18)\alpha + 0.01, -(0.29)\alpha + 0.48$

By substituting 1 and 0 for α, we obtain the lower/upper $\alpha - cut$ bounds of the interval

Ex: The sum $A + B$ for every $\alpha \in (0, 1)$ as:

$(A + B)_\alpha = A_\alpha + B_\alpha = [A_L(\alpha) + B_L(\alpha), \quad A_U(\alpha) + B_U(\alpha)]$

$= [(a_2 - a_1)\alpha + a_1] + [(b_2 - b_1)\alpha + b_1],$

$[-(a_3 - a_2)\alpha + a_3] + [-(b_3 - b_2)\alpha + b_3]$

Calculating the burden of ventilation exposures using fuzzy sets

We estimated the annual morbidity burdens attributable to the three ventilation exposure scenarios. The process consisted of various steps. We first calculated risk ratios associated with all ventilation exposure scenarios. The risk ratios were calculated assuming a log-linear function based on the level of ventilation exposure, the odds ratio and the unit threshold associated with the odds ratio. [146] In the risk ratio, two input parameters were defined as fuzzy sets: the ventilation exposure scenario (i.e. poor, fair and adequate), and the excess risk in morbidity due to ventilation below 0.5 ACH threshold. The excess risk in morbidity was obtained by taking the natural logarithm of the odds ratio with its 95 % CI, and mapping the bounds of the 95 % CI to the bounds of the fuzzy set as shown in Appendix B. The risk ratio for each scenario is given by

$$RR_i = exp[E \times (0.5 - X_i)] \qquad [7]$$

where RR_i is a fuzzy set which describes the risk ratio adjusted to the exposure in ventilation scenario i, E is a fuzzy set describing the excess risk in morbidity due to ventilation below a threshold unit in ACH. "0.5" is the ventilation threshold unit from which a health effect is observed per degree below ACH, and X_i is a fuzzy set describing the exposure parameter for each ventilation scenario i.

Changes in morbidity burdens attributable to the three ventilation exposure scenarios were calculated using the population attributable fraction (PAF). The PAF is an

epidemiological method that calculates the health effect due to changes in exposure for the whole population (exposed and unexposed). [147-149] The PAF is given by:

$$PAF_i = \frac{p(RR_i - 1)}{p(RR_i - 1) + 1} \qquad [8]$$

where p is the proportion of the population exposed (a value "1" for p represents that everyone in the population is exposed), and RR_i is the risk ratio associated with the exposure for each ventilation scenario.

In addition, we calculated the total number of annual respiratory-related morbidity cases attributable to changes in indoor ventilation as the final outcome of the assessment. The annual morbidity burden (AMB) attributable to the three ventilation exposure scenarios is given by: [149]

$$AMB = PAF_i \times B \qquad [9]$$

where B is the total annual number of existing respiratory-related morbidity cases in England and PAF_i is the population attributable fraction corresponding to ventilation exposure scenario i.

Results

Selecting the exposure metric and quantifying the evidence of potential health effects

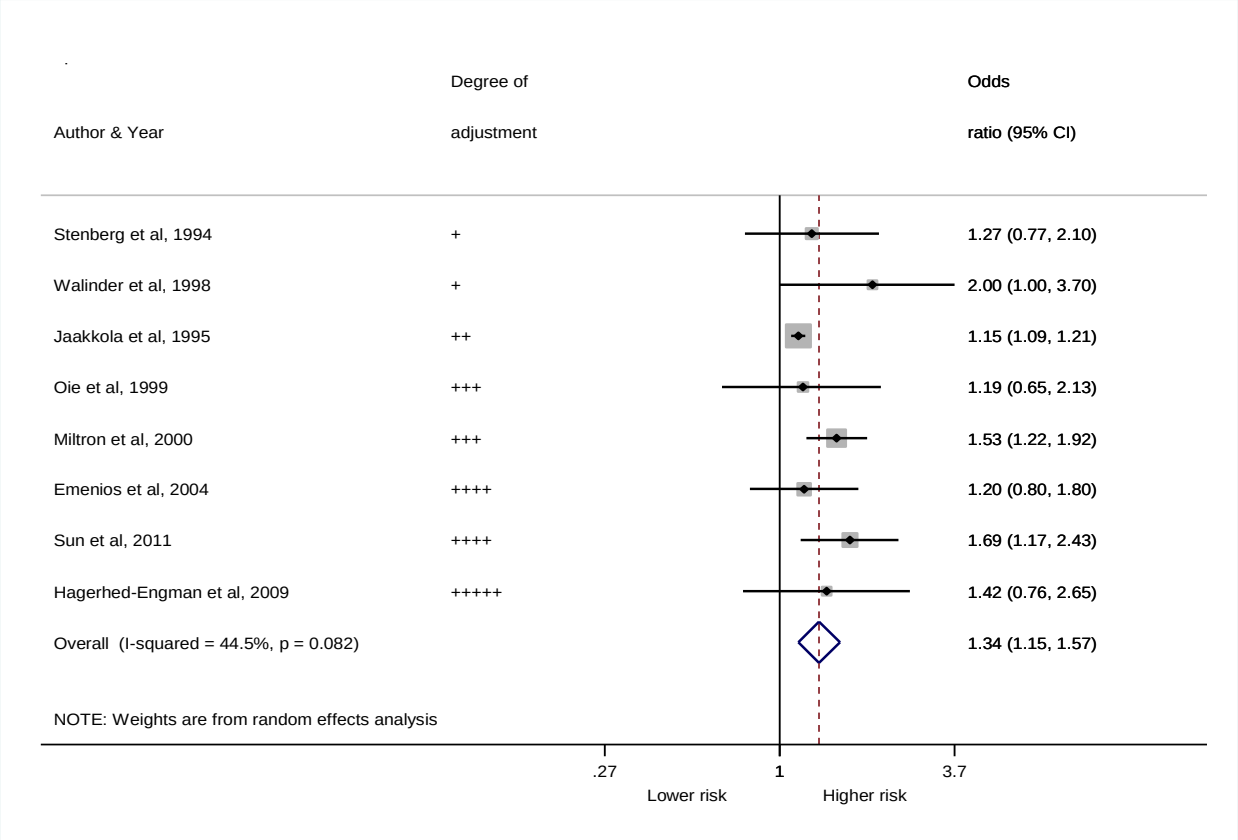
As part of the systematic search, the literature yielded a total of 621 peer reviewed

articles, of which 586 articles were deemed to be irrelevant or duplicates after reviewing titles and abstracts, leaving 35 articles to be retrieved for further evaluation. Of the 35 studies assessed, 8 articles met the inclusion criteria [150-157] and were included in the meta-analysis (9,826 participants). Studies included in the meta-analysis controlled for a number of confounders, including age, sex, crowding, building age, history of eczema, asthma, allergic rhinitis and outdoor temperature. Table 8 shows the characteristics of included studies in the meta-analysis. The result of the meta-analysis yielded an overall estimate of 1.34 OR (95 % CI 1.15 to 1.57) as shown in Figure 14, which gives a quantitative *summary* measure, with uncertainty in the 95 % CI, of the association between poor ventilation exposure (less than 0.5 ACH) and health. There was no evidence to suggest that the pooled estimate of OR and its 95% CI were affected significantly by heterogeneity (Appendix A)

Table 8: General characteristics of included studies

Author, year	Design	Geographical location	Study population	Health outcomes assessed
Stenberg et al., 1994	Survey questionnaires, case-control study, 3 months, 1989	Sweden	464 office workers stratified for geographical areas with 83 % of women in each group of the sample	Sick building Symptoms (SBS)
Jaakkola et al., 1995	Questionnaires, cross sectional study, 12 months, 1991-1992	Finland	399 office workers selected randomly from 14 mechanical ventilated office buildings	Ocular, nasal symptoms and allergic reactions
Walinder et al., 1998	Self-administered questionnaires, cross sectional study, 24 months, 1993-1995	Sweden	234 school personnel working in the main buildings of 12 randomly selected primary schools	Nasal symptoms
Oie et al., 1999	Survey questionnaires, case-control study, 24 months, 1992-1993	Norway	172 children in residence homes	Bronchial obstruction
Milton et al, 2000	Questionnaires, cross sectional study, 12 months, 1994-1995	US	3,720 employees of a large manufacturer in 40 buildings	Monthly short-term sick-leave
Emenius et al., 2004	Cohort study, 24 months, 1994-1996	Sweden	4089 children in residence homes	Wheezing
Hagerhed-Engman et al., 2009	Survey questionnaires, case-control study, 6 months, 2001-2002	Sweden	400 children in residence homes	Asthma, rhinitis, eczema
Sun et al., 2011	Survey questionnaires, case–control study, 12 months, 2006–2007	China	348 college students in college dorms at Tianjin University	Wheezing, rhinitis, dry cough

Figure 14: Result of meta-analysis: odds ratio (95% CI) for respiratory-related morbidity in high ventilation > 0.5 ACH (reference group) compared to low ventilation < 0.5 ACH (exposure group.



+ adjustment for sex and age only; ++ for these plus history of atopy (e.g. history of eczema, asthma, allergic rhinitis and others); +++ for these plus crowding and building age ; ++++ for these plus smoking); +++++ for these plus outdoor temperature.
Cochran Q= 12.61 (df= 7) P= 0.082, Tau squared = 0.0175.

Estimating population affected by exposure and selecting outcome measure

Based on the UK Office of National Statistics, the population in England is projected to be 53 million (53,107,000 people) up to mid-2011 projections. [158] In terms of outcome measures, common symptoms in relation to poor ventilation exposures were identified as: allergies, rhinitis, asthma, wheezing and others (Table 8). Some authors have grouped these conditions under the terms "building-related symptoms" or "sick building syndrome" to describe a range of outcomes associated with indoor environmental exposures. [159, 160] According to the HSE report of annual respiratory-related cases in England, a total of approximately 8% including children and adults had reported in the last 12 months symptoms of wheezing, asthma and whistling in the chest. [142] In the HSE report, this was estimated to be a total of 4.2 million (4,178,720) of the current annual respiratory-related morbidity cases in England.

Quantifying the health impacts and uncertainty

The input parameters defined as fuzzy sets in the HIA model are shown in Table 9. Table 10 shows the morbidity burdens and corresponding uncertainty under the three ventilation scenarios. The negative values refer to health gains. In relation to annual respiratory-related morbidity cases, an excess of 371,097 cases were estimated under the poor ventilation scenario. Under the fair ventilation scenario, 24,997 total annual respiratory-related morbidity cases were attributable to the exposure. In the adequate ventilation scenario, a reduction of approximately 352,562 cases in annual morbidity cases were estimated in England. The uncertainty bounds under the poor ventilation scenario ranged between 99,398 – 1,028,008 cases attributable to the exposure. Under the fair ventilation scenario, a reduction of 539,846 cases and an increase of 706,364 cases were estimated attributable to the exposure. Under the adequate

ventilation scenario, a reduction of between 1,197,605 and 48,605 cases were estimated approximately. The fuzzy sets describing the adjusted risk ratios used in the calculation for each scenario is given in Figure 15. Each fuzzy set gives the interval values which describe the propagation of uncertainty in the parameters with an interval bounds and the centre value of the interval. Additionally, the fuzzy sets describing the total annual morbidity burden are given in Figure 16.

Table 9: Input parameters and corresponding fuzzy intervals

Definition	Explanation	Intervals
Ventilation rate (X):		
Poor	Poor ventilation rates in air changes per hour (ACH) for "low" indoor air quality	$(0.01\ \text{ACH} \leq \boldsymbol{X} < 0.48\ ACH)$
Fair	Fair ventilation rates in air changes per hour (ACH) for "medium/fair" indoor air quality	$(0.19\ ACH \leq \boldsymbol{X} \leq 0.77\ ACH)$
Adequate	Adequate ventilation rates in air changes per hour (ACH) for "high" indoor air quality	$(0.48\ ACH \leq \boldsymbol{X} < 1.06\ ACH)$
Increase in risk (E):		
Excess risk in morbidity	Excess percentage (%) increase in respiratory-related morbidity risk derived from meta-analysis	$(0.14 \leq \boldsymbol{E} \leq 0.45)$

Table 10: Annual respiratory-related morbidity burdens attributable to each ventilation exposure scenario

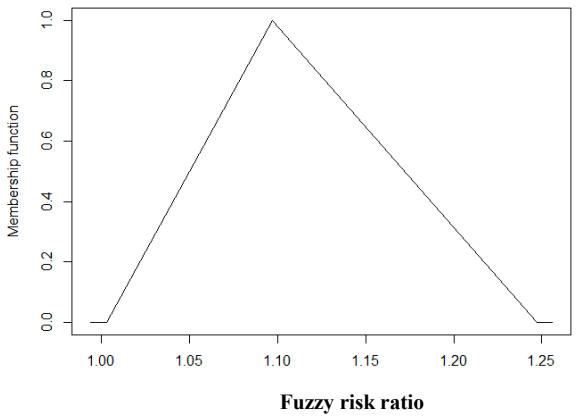
	Mid-year 2011 England population *	**Excess respiratory-related morbidity cases (n) attributable to each exposure scenario**
<u>Ventilation scenarios</u>	$\mathbf{53,107,000}$	
Poor ventilation	-	$(99{,}398 \quad \mathbf{371,097} \quad 1{,}028{,}008)$,
Fair ventilation	-	$(-539{,}846 \quad \mathbf{24,997} \quad 706{,}364)$
Adequate ventilation	-	$(-1{,}197{,}605 \quad \mathbf{-352,562} \quad -48{,}605)$

*source: ONS, 2012

Figure 15: Fuzzy sets describing uncertainty propagation of adjusted risk ratios in model parameters.

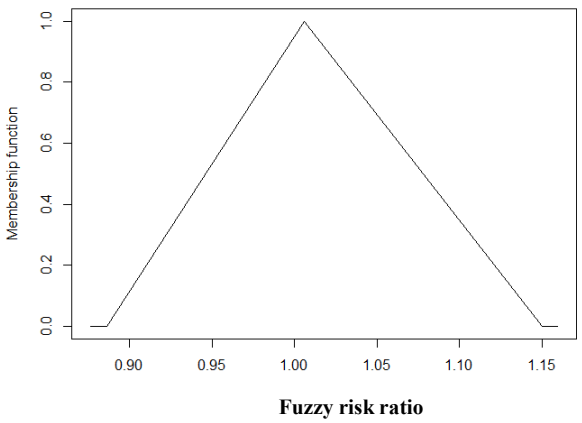
a) Poor ventilation scenario

RR_{poor} = (1.003 **1.097** 1.247)



b) Fair ventilation scenario

RR_{fair} = (0.886 **1.006** 1.150)



c) Adequate ventilation scenario

$RR_{adequate}$ = (0.777 **0.922** 1.009)

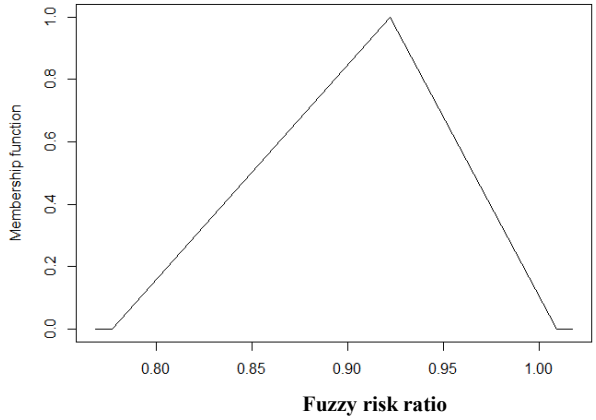
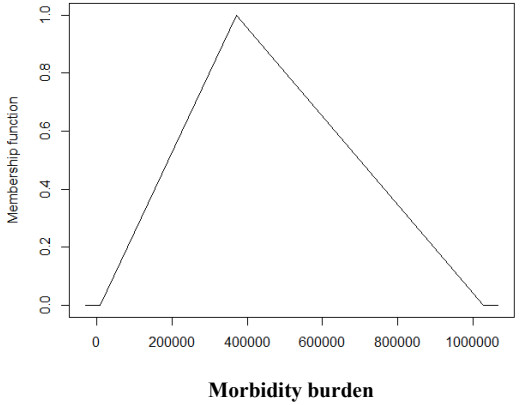


Figure 16: Annual respiratory-related morbidity burdens attributable to changes in indoor ventilation scenarios and corresponding uncertainty described in fuzzy sets

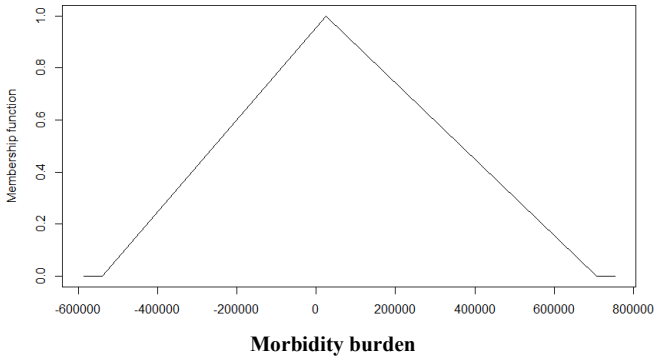
a) Poor ventilation scenario

AMB_{poor} = (99,398 **371,097** 1,028,008)



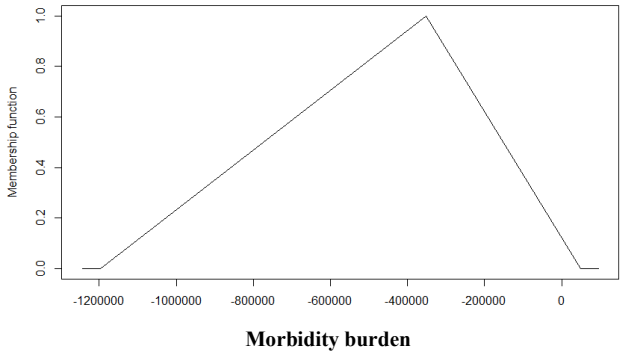
b) Fair ventilation scenario

AMB_{fair} = (-539,846 **24,997** 706,364)



c) Adequate ventilation scenario

$AMB_{adequate}$ = (-1,197,605 **-352,562** -48,605)



Discussion

In this study, we provided a framework that can be used as part of the assessment stage of a HIA. We applied the framework to a case-study example of indoor housing ventilation in England. In the case-study, we used meta-analysis to get an estimate of the odds ratio of the association between indoor ventilation and health, and a health impact model to calculate respiratory-related morbidity burdens attributable to changes in indoor ventilation exposures.

Findings from the case-study

The literature search in the case-study identified a total of 8 studies with 9,826 participants that were included in the meta-analysis from which an exposure response relationship was derived: 1.34 OR (95% CI: 1.15 to 1.57), for ventilation rates below 0.5 ACH. We believe that the finding from the meta-analysis contributes to the body of the evidence linking poor ventilation rates and health. To the best of our knowledge, this is the first meta-analysis providing summary estimates of the associations between indoor ventilation rates and health using only epidemiological study designs. Results from the meta-analysis seem consistent with other research that summarised the evidence on ventilation rates and health using other study designs, where the authors concluded that a decrease in ventilation rates (from approximately 0.5 to 0.2 ACH) increases the prevalence of respiratory-related symptoms between 12 % - 32 %. [100] In addition, based on the ventilation exposure scenarios from the case-study result, we found using a health impact model that respiratory-related morbidity due to “poor” ventilation scenario can potentially be a significant contributor to the total annual respiratory-related morbidity cases in England. We also found that imprecision in the definition of each exposure scenario

was of major significance in the outcome of the model. In this paper, the uncertainty in the outcome of the model can be described into three groups. The first consisted of morbidity burdens ranging from 99,398 to 1,028,008 cases, which is characterised by poor ventilation exposures (between 0.01 ACH to 0.48 ACH). The second group is in the range of -539,846 to 706,364 cases, which represents a compromise between a reduction and an increase of morbidity cases, and it is characterised by fair ventilation exposures (between 0.19ACH to 0.77ACH). The third group ranges from -1,197,605 to -48,605, as a reduction of cases, which is characterised by adequate ventilation exposures (between 0.48ACH to 1.06 ACH). Thus, the lowest level of ACH of ventilation exposure resulted with the greater impacts on health.

Our finding emphasises the need to ensure adequate ventilation levels to minimise the potential health effects from poor ventilation exposure as buildings become more airtight in England. There is evidence in the wider literature to suggest that low ventilation rates increases air-borne pollutants concentration. [153] For instance, one extensive review has suggested that ventilation rates lower than 0.5 ACH in cold climates can increase the risk of negative health outcomes. [133] Ventilation rates between 0.5 and 1.5 ACH in the UK are considered sufficient to stop condensation and to control indoor pollutants. [161] In addition, this research finding adds to this evidence-based suggesting that 0.5 ACH can be considered an actual threshold from which a population health effect based on the exposure can be observed.

Strengths and limitations

As part of the framework, we applied a method based on fuzzy set theory to deal with the uncertainty in the parameters of a model. Given the lack of probabilistic

information in some input parameters (e.g. statistical information regarding ventilation exposure for the English housing stock), the application of fuzzy set theory was considered appropriate for the quantification of uncertainty, as an alternative way of handling uncertainty to the probabilistic approach. The uncertainty in each exposure scenario was represented using fuzzy sets, and their spread was determined based on plausible information on ventilation rates' guidelines for indoor air quality. We also characterised the uncertainty in the 95 % CI of the odds ratio as a fuzzy set, which was used as an input parameter for the health impact model. Fuzzy sets were defined in this study with a triangular membership function with one interval and a centre value of the interval, which represents the lower and upper bounds of the fuzzy set respectively. It is important to note that there are many choices for membership functions of fuzzy sets such as trapezoids and Gaussian membership functions, which are described elsewhere. [20]

This study was only able to quantify the health impacts of housing ventilation in England based on limited information found in the literature. The meta-analysis presents some limitations. The diversity of the study populations, geographical locations from individual studies in the analysis can make the overall estimate sub-optimal for the English context. Another limitation is that the smaller number of eligible studies (8 studies) might have influenced the power of the meta-analysis, although such bias and limitation regarding the small number of studies can be reduced as more studies become available in the literature. We also defined each exposure ventilation scenario with specific ventilation rate categories. Other ventilation rates categories were not considered in this analysis. For example, there are very high ventilation categories which exceed ventilation rates greater than 1.06

ACH, which were not considered. We also incorporated the uncertainty in the 95 % CI of the odds ratio using fuzzy sets without probabilistic guarantees or distributional assumptions. A potential limitation of the fuzzy set approach is that the fuzzy set does not incorporate knowledge regarding correlation and other statistical information in parameters, and this can be a limitation in circumstances when there is sufficient information to incorporate statistical information such as mean, correlations and other.

A different point of view of uncertainty

When comparing the proposed method with other probabilistic approaches is important to note that both approaches deal with different aspects of uncertainty. Uncertainty can arise in the assessment from two underlying causes. Uncertainty can arise due to imprecision in knowledge because of limited information, or due to random variability found in the stochastic nature of most real-world variables. It could be argued that the fuzzy-set method provides a better measure for the characterisation of the uncertainty in circumstances characterised with limited information about statistical parameters or imprecision in knowledge. On the other hand, probabilistic approaches can provide a better characterisation of uncertainty if suitable assumptions can be made on the statistics of the variability in the input parameters. Monte Carlo (MC) methods rely on random sampling and simulations, to obtain probability distributions from which statistical parameters can be estimated to characterise the uncertainty. These methods assume model parameters to be random variables, using statistical inference with sampling techniques to obtain parameter distributions of the random variable. However, such probabilistic approaches to uncertainty propagation can be less suitable to deal with the

uncertainty associated with lack or imprecision in knowledge than the fuzzy set approach. Assuming random variability in model parameters when there is limited statistical information can lead estimates in epidemiological models to very different conclusions if suitable assumptions on the statistics of the variability cannot be made [162]. We instead propagated the uncertainty using interval arithmetic with the fuzzy set approach, and provided some interval estimates for the characterisation of uncertainty without assuming random variability in model parameters, but rather assuming the information in parameters to be imprecise by nature.

Conclusion

We have proposed a non-probabilistic framework using fuzzy set theory to quantify the uncertainty in HIA and applied it to housing ventilation as an example. The framework could also enable the quantification of the health impacts by following three steps: (i) selecting the exposure metric and quantifying the evidence of potential health effects of the exposure, (ii) estimating the size of the population affected by the exposure and selecting the outcome measure, and (iii) quantifying the health impacts and associated uncertainty. The framework is demonstrated through a HIA case study which examines the health impact of housing ventilation in England. We have argued that this framework can be applicable to other examples of quantitative HIA where there is insufficient information for a probabilistic analysis. This includes situations where the uncertainty in model parameters cannot be described by probability density functions, because of either of lack of statistical information or the input parameters are not precisely defined.

Supplementary Material: Appendix A and B

APPENDIX A

Selecting the exposure and quantifying the evidence of potential health effects

Details of search strategy in Ovid Medline database

#1 search (ventilation or ventilation rate or air flow or air exchange).ti,ab
#2 search (health or sick building* or allergy* or illness or asthma).ti,ab
#3 search (housing or home or apartment or dwelling or building or residence).ti,ab
#4 search #1 or #2
#5 search #3 and #4
#6 limit #5 to English language
#7 remove duplicates from #6

Ovid Database fields:

- ab: Abstract
- ti: Title

Figure A.1. Flow diagram of included studies

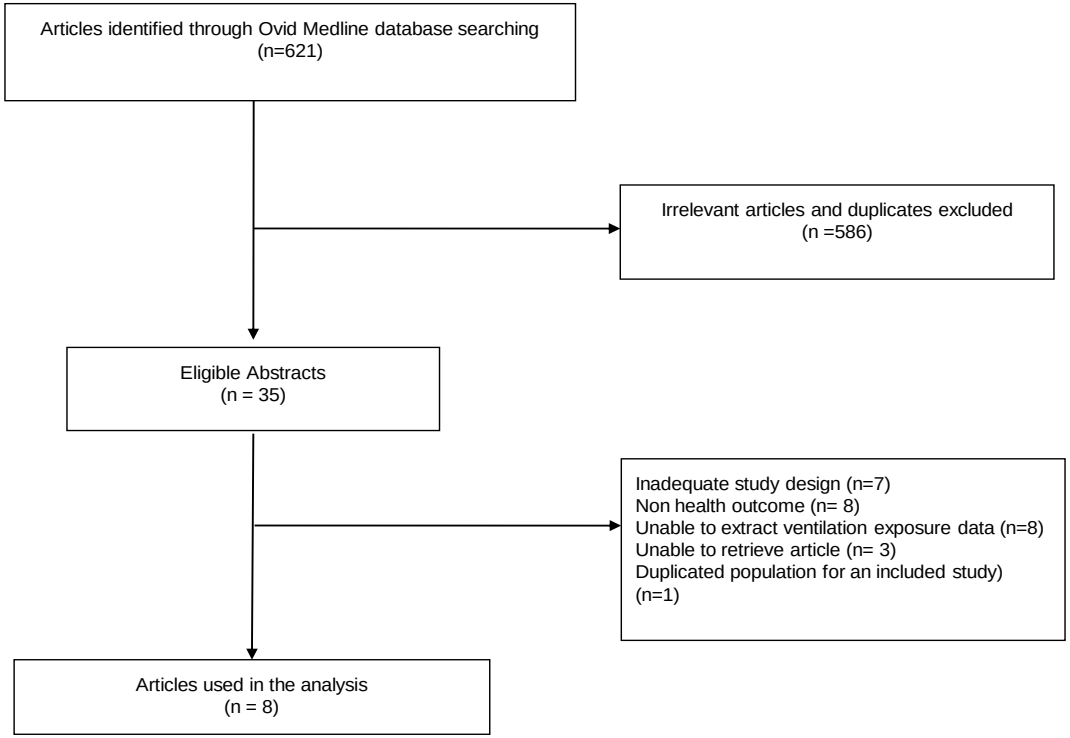


Table A.1: Subgroup analysis with summary Odds Ratios of the association between respiratory-related symptoms and poor ventilation

Subgroup	No. of Studies	Total population studied	OR (95 % CI)	Test for heterogeneity	
				τ^2	P
Study design					
Case-control	4	1,384	1.45 (1.14-1.85)	0.0000	0.7120
Cross-sectional	3	4,353	1.38 (1.04-1.82)	0.0401	0.0150
Cohort	1	4,089	1.20 (0.80- 1.80)	N/A	N/A
Building setting					
Office	3	4,583	1.28 (1.03-1.60)	0.0228	0.0530
School	2	582	1.76 (1.28-2.42)	0.0000	0.6600
Residence	3	4,661	1.24 (0.93-1.67)	0.0000	0.8940
Study size					
<1000	6	2,017	1.31 (1.09-1.57)	0.0164	0.1960
>1000	2	7,809	1.34 (1.15-1.57)	0.0014	0.3050
Geographical location					
Sweden	4	5,485	1.36 (1.05-1.76)	0.0000	0.6170
Finland	1	399	1.15 (1.09-1.21)	N/A	N/A
Norway	1	172	1.19 (0.65-2.13)	N/A	N/A
US	1	3,720	1.53 (1.22-1.92)	N/A	N/A
China	1	348	1.69 (1.17-2.43)	N/A	N/A
Study population					
Children	3	4,661	1.24 (0.93-1.67)	0.0000	0.8940
Adults	5	5,165	1.41 (1.13-1.75)	0.0344	0.0150

Sub-group analysis stratified by different aspects of the studies is shown in Table A.1. Heterogeneity was examined using the Cochran's Q test which is a classical measure that tests the statistical significant ($p < 0.1$) variation among study outcomes and it is used to estimate P values for heterogeneity. The tau-squared τ^2 statistics was also used to assess the degree of variation among study outcomes due to true

substantial heterogeneity ($\tau^2 > 1$) rather than variations due to random chance [163]. In the analysis stratified by building settings, three studies in school buildings showed an association between poor ventilation and respiratory-related morbidity of 1.76 OR (95 % CI: 1.28 to 2.42). One study, stratified by geographical location in China provided an estimate of 1.69 OR (95 % CI: 1.17 to 2.43). Additionally, studies stratified in the adult population showed a higher summary estimate of 1.41 OR (95 % CI: 1.13 to 1.75) compared to those studies conducted only in children 1.24 (95 % CI: 0.93 to 1.67). The difference in results, for example, between children and adults, might be explained by some sources of heterogeneity. In the adult population, the Cochran's Q test P value for heterogeneity was 0.0150, which indicates some heterogeneity among study outcomes. Heterogeneity is expected in the result, given the diversity of studies included and the small number of studies included. However, there is no evidence to suggest that substantial sources of heterogeneity ($\tau^2 > 1$) can extensively affect the conclusion of the meta-analysis and its 95 % CI. The limited evidence available due to the small number of studies, however, does not make it possible to evaluate and adjust for more covariates to explore other sources of heterogeneity. In Appendix B we quantify the uncertainty in the overall result of the meta-analysis (in the 95 % CI of the OR) using fuzzy sets to incorporate this limited information.

APPENDIX B

Quantification of health impacts and their uncertainty

To aid explanation, the method used to calculate the health impacts, and their uncertainty is described chronologically in the steps below. Fuzzy sets are used to quantify the uncertainty in the input parameters and the outputs of the health impact assessment model. In this section, we use the fuzzy sets to: (a) quantify the percentage increase in morbidity risk; (b) characterise the uncertainty in ventilation exposure; (c) describe interval arithmetic operations using fuzzy sets.

a) Quantify the percentage increase in morbidity risk

The meta-analysis carried out in Appendix A gave the mean value of the odds ratio (OR) of the association of respiratory symptoms with poor ventilation as 1.34, and its 95% confidence interval as [1.15 to 1.57]. These values can be transformed into percentages by taking the natural logarithm of the odds ratio (OR):

$$E = \ln(OR) \qquad [B.1]$$

where E is the percentage excess risk in respiratory-related morbidity due to ventilation rates below 0.5 air changes per hour (ACH).

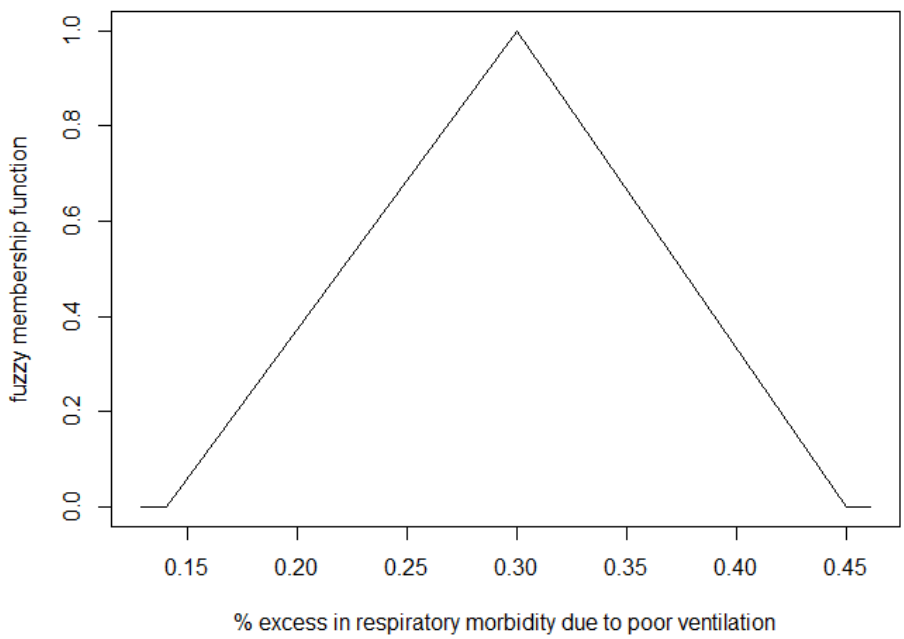
Using equation [B.1], we obtain the central estimate to be 30% and the lower- and the upper bounds of the 95% confidence interval as 14% and 45%.

For example, if we define $E(x)$ as a fuzzy set on x (See also Figure B.1) and we apply equation [1] from the main manuscript, the membership function for all parameters can be defined as:

$$\mu_{\tilde{E}}(x) = \begin{cases} 0, & x < 0.14 \\ \left|\dfrac{x - 0.14}{0.30 - 0.14}\right| & 0.14 \leq x \leq 0.30 \\ \left|\dfrac{0.45 - x}{0.45 - 0.30}\right| & 0.30 \leq x \leq 0.45 \\ 0, & x > 0.45 \end{cases} \qquad [B.2]$$

where 0.14, 0.45 and 0.30 are the lower, upper and centre values describing the excess in respiratory-related morbidity due to ventilation rates below 0.5ACH.

Figure B.1. Fuzzy set $E(x)$ with membership function describing the excess risk in respiratory-related morbidity below 0.5 ACH.



b) Characterise ventilation exposure

Information on ventilation rates were obtained from a study by the Building Research Establishment (BRE) which estimated ventilation rates in a sample of 33 UK dwellings (built after 1995). [164] Ventilation rates were found in the range of 0.19-1.06 air changes per hour (ACH) during the winter and summer months. [132, 164] Three possible ventilation exposures scenarios are defined based on the above information and classified according to ventilation rates standards for indoor air quality: [143] (i) *adequate* ventilation rates for "high" indoor air quality; *(ii) fair* ventilation rates for "medium" indoor air quality; and (iii) *poor* ventilation rates for "low" indoor air quality. If we apply equation [2] from the main manuscript section, the ventilation exposure values from each category are defined as follows, where the unit of ventilation rate is ACH:

$$\mu_{poor}(x) = \begin{cases} 0, & if \;\; x < 0.01 \\ \left|\dfrac{x - 0.01}{0.19 - 0.01}\right| & if \;\; 0.01 \leq x \leq 0.19 \\ \left|\dfrac{0.48 - x}{0.48 - 0.19}\right| & if \;\; 0.19 \leq x \leq 0.48 \\ 0, & if \;\; x > 0.48 \end{cases} \qquad [B.3]$$

$$\mu_{fair}(x) = \begin{cases} 0, & if \;\; x < 0.19 \\ \left|\dfrac{x - 0.19}{0.48 - 0.19}\right| & if \;\; 0.19 \leq x \leq 0.48 \\ \left|\dfrac{0.77 - x}{0.77 - 0.48}\right| & if \;\; 0.48 \leq x \leq 0.77 \\ 0, & if \;\; x > 0.77 \end{cases} \qquad [B.4]$$

$$\mu_{adequate}(x) = \begin{cases} 0, & if\ \ x < 0.48 \\ \left|\dfrac{x - 0.48}{0.77 - 0.48}\right| & if\ \ 0.48 \leq x \leq 0.77 \\ \left|\dfrac{1.06 - x}{1.06 - 0.77}\right| & if\ \ 0.77 \leq x \leq 1.06 \\ 0, & if\ \ x > 1.06 \end{cases} \qquad [B.5]$$

where μ_X is the membership function of the fuzzy set X_i for each exposure scenario i, where i is either "*poor*", "*fair*" or *adequate*. The three ventilation exposure scenarios are specified in the following intervals: *Poor* ventilation $(0.01\ ACH\ \leq x < 0.48\ ACH)$; *Fair* ventilation $(0.19 \leq x < 0.77\ ACH)$; and *Adequate* ventilation $(0.48 \leq x < 1.06\ ACH)$.

c) Describe interval arithmetic operation using fuzzy sets

Fuzzy interval arithmetic operations are approximated by finding the lower and upper α-cut bounds of a fuzzy set. See example 1 below.

<u>Example 1</u>

Assume one fuzzy set A with three parameter values $A = (2,4,8)$. If we apply equation [2-4] from the main manuscript section, we obtain the fuzzy sets and its corresponding α-cut bounds as follows.

$$\mu_A(x) = \begin{cases} 0, & if\ \ x < 2 \\ \left|\dfrac{x - 2}{4 - 2}\right| & if\ \ 2 \leq x \leq 4 \\ \left|\dfrac{8 - x}{8 - 4}\right| & if\ \ 4 \leq x \leq 8 \\ 0, & if\ \ x > 8 \end{cases} \qquad [B.6]$$

The lower/upper α-cut bounds of A_α (where $\in [0,1]$) are:

$$A_\alpha := [A_L(\alpha), \quad A_U(\alpha)]$$
$$= [(4-2)\alpha + 2, \quad -(8-4)\alpha + 8]$$
$$= [2\alpha + 2, \ 8 - 4\alpha] \qquad [B.7]$$

Common interval arithmetic operations such as addition, subtraction, multiplication and others are approximated using equation [B.8-B.12] (as shown below).[165-167]

(Addition)

$$(A+B)_\alpha := A_\alpha + B_\alpha$$
$$= [a_1 + b_1 + (a_2 - a_1 + b_2 - b_1)\alpha, \ a_3 + b_3$$
$$- (a_3 - a_2 + b_3$$
$$- b_2)\alpha] \qquad [B.8]$$

(Subtraction)

$$(A-B)_\alpha := A_\alpha - B_\alpha$$
$$= [(a_1 - b_3) + (a_2 - a_1 + b_3 - b_2)\alpha, \ (a_3 - b_1)$$
$$- (a_3 - a_2 + b_2 - b_1)\alpha] \qquad [B.9]$$

(Multiplication)

$$A_\alpha \times B_\alpha = [((a_2 - a_1)\alpha + a_1) \times ((b_2 - b_1)\alpha + b_1),$$
$$(a_3 - (a_3 - a_2)\alpha)$$
$$\times (b_3 - (b_3 - b_2)\alpha] \qquad [B.10]$$

<u>Substituting the lower/upper α-cut bounds</u>

During interval arithmetic calculations, the interval values of a fuzzy set (i.e. lower, upper and a centre value of the interval) are found by substituting 1 and 0 for α in a given equation, as shown in the examples below.

<u>Example 2</u>

Assuming the following values e.g., $A = (2, 4, 8)$, $B = (3,6, 9)$, the α-cut of fuzzy set A and B can be described as follows by using equation [2-4] from main section of the manuscript as follows:.

$$A_\alpha = [(4-2)\alpha + 2, -(8-4)\alpha + 8]$$

$$= [2\alpha + 2, -4\alpha + 8]$$

$$B_\alpha = [(6-3)\alpha + 3, -(9-6)\alpha + 9]$$

$$= [3\alpha + 3, -3\alpha + 9]$$

For all $\alpha \in [0,1]$, multiply A_α and B_α as two regular intervals using equation [B.10]

$$A_\alpha \times B_\alpha = [2\alpha + 2, -4\alpha + 8] \times [3\alpha + 3, -3\alpha + 9]$$

$$= [(2\alpha + 2) \times (3\alpha + 3), \quad (-4\alpha + 8) \times (-3\alpha + 9)]$$

We obtain:

$$= [6\alpha^2 + 12\alpha + 6, \quad 12\alpha^2 - 60\alpha + 72]$$

Substituting $\alpha = 0$:

$$A_0 \times B_0 = [6, 72]$$

Substituting $\alpha = 1$:

$A_1 \times B_1 = [6 + 12 + 6, \quad 12 - 60 + 72] = [24, 24] = 24$

Therefore a fuzzy set is obtained, which is an approximation of $A \times B$:

$$A \times B \cong (6, \quad 24, \quad 72)$$

(Division)

$A_\alpha \div B_\alpha = [(a_2 - a_1)\alpha + a_1]/([b_3, -(b_3 - b_2)\alpha]), [a_3 - (a_3 - a_2)\alpha]/([(b_2 - b_1)\alpha + b_1])$ [B.11]

<u>Example 3</u>

Similarly, approximated values of $A \div B$ can be expressed as fuzzy sets by dividing the interval $A_\alpha \div B_\alpha$ for all $\alpha \in [0,1]$, using equation [B.11]:

$A_\alpha \div B_\alpha = [(2\alpha + 2)/(-3\alpha + 9)], [(-4\alpha + 8)/(3\alpha + 3)]$

<u>When $\alpha = 0$:</u>

$A_0 \div B_0 = [2/9, \quad 8/3]$

$= [0.2, \ 2.6]$

<u>When $\alpha = 1$:</u>

$A_0 \div B_0 = [4/6, \quad 4/6]$

$= 0.6$

Therefore, the approximation of fuzzy sets A and B is:

$$A \div B \cong (0.2, \quad 0.6, \quad 2.6)$$

(Exponential)

Exponential operations with values restricted by a fuzzy set can be performed and approximated using the α-cut also (as shown in Example 2 and 3).[19, 168] If we consider a fuzzy set e.g. $A_\alpha = (a_1, a_2, a_3)$, the exponential of a fuzzy set A is:

$$A_\alpha = [exp(a_2 - a_1)\alpha + a_1),\ \ exp(\ a_3 - (\ a_3 - \ a_3)\alpha\ )] \qquad [B.12]$$

(Logarithm)

Logarithmic operations with fuzzy sets can be performed using equation [B.13], given by:

$$A_\alpha = [\ ln(a_2 - a_1)\alpha + a_1),\ \ ln(\ a_3 - (\ a_3 - \ a_3)\alpha\ )\ ] \qquad [B.13]$$

5.3. Supplementary material to chapter 5 – Further analysis based on research paper 3

The main emphasis of research paper 3 was on characterising the uncertainty in two input parameters of a HIA model of ventilation exposures in England. Research paper 3 attempted to quantify the impact on parametric sources of uncertainty, in particular the definition of exposure scenarios and the exposure response function (excess risk). This section provides further explanation of the process of quantifying the uncertainty in the two parameters and it explores how the use of fuzzy shapes other than triangular fuzzy sets in the propagation of uncertainty might affect the conclusion.

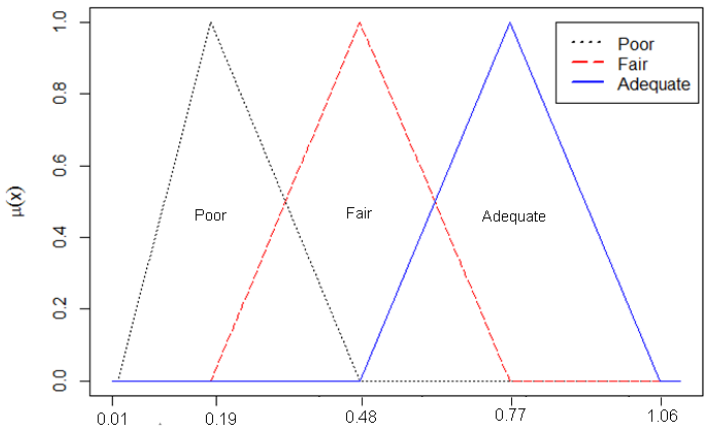
Uncertainty in the definition of exposure scenarios

A key aspect in defining the exposure scenarios is the choice of interval around the fuzzy sets used to characterise the uncertainty in the three ventilation exposure scenarios. According to a literature search, evidence on ventilation exposures by the Building Research Establishment (BRE) was found in the range of 0.19-1.06 ACH.[132, 161] At the time of the PhD study, the above range is the only evidence found on ventilation exposure involving energy efficient homes in the UK. The initial source of uncertainty in this example can be described by lack of information or understanding about the potential exposures in England, and not on some random variations at this stage. Therefore, the uncertainty in the range is characterised in a bounded interval using fuzzy sets without assuming random or stochastic variation. It is worth noting that fuzzy sets are not mutually exclusive, in other words, a fuzzy value can belong to two or more sets. This is in contrast to probability where all positive values in probability can only represent an unique value of a variable. A

fuzzy membership function can measures *the degree* to which a value can belong to two or more sets, and it should not be interpreted as the probability of a value belonging to a set. For example, consider the degree of membership of a ventilation value from the three ventilation exposure scenarios defined in paper 3 (Table 11 below).

Table 11: Fuzzy membership (Example of ventilation rate values ACH)

Ventilation rates ACH	Degree of membership to poor μ(x)	Degree of membership to fair μ(x)	Degree of membership to adequate μ(x)
0.3	0.62	0.38	0
0.75	0	0.07	0.93
0.44	0.14	0.86	0
0.65	0	0.41	0.59



The above example illustrates that each potential ventilation rates ACH value could belong to one or more fuzzy sets defined by their corresponding membership function (i.e. the degree to which a value belong to a set).

In paper 3, an universal space of 0.01-1.06 ACH is divided using the three fuzzy set categories defined (Table 11) to include the empirical evidence on ventilation exposures found by the BRE (i.e. 0.19 to 1.06 ACH). The central values in each fuzzy sets are based on expert opinion from the BRE depicting the relationship between ventilation level needed to maintain air quality and energy efficiency: ~0.19 ACH poor, ~0.48 ACH fair and ~ 0.77 ACH adequate ventilation. Such categorisation is rather arbitrary given the lack of evidence or consensus; hence the importance of defining fuzzy sets to account for imprecision in knowledge and allowing the possibility of the true value to belong to one or more ventilation scenarios (i.e. fuzzy sets). A very low ventilation value of 0.01 ACH is assumed in the interval to allow values less than 0.19 ACH to be included in the fuzzy set relating to poor ventilation. However, the smaller the value, the less likely the value is assumed to belong to poor ventilation. In reality ventilation values less than 0.19 ACH would belong to an extremely poor ventilation category. Whole-house ventilation less than 0.19 ACH are highly unlikely, and this is the reason for such initial categorisation in the fuzzy set. Moreover, ventilation values less than 0.01 ACH are not defined in the space and are not assumed in the exposure scenarios as they would not considered a plasusible representation (i.e. ventilation rates less than 0.01 ACH do not exist in reality).

An upper range of 1.06 ACH is given to the fuzzy set corresponding to adequate ventilation, as values more than 1.06 ACH would not be considerered adequate from an energy efficient standpoint. In general , ventilation strategies from an energy-efficient standpoint, should achieve a total (uncontrolled plus controlled) whole

house ventilation rate of around 0.5ACH - 1 ACH to maintain indoor air quality and energy efficiency based on the evidence from the literature.[143]

Uncertainty in the exposure-response function

There are multiple sources of uncertainty that arises in the definition of the exposure-response function. The exposure of interest is ventilation rate and its potential impact on population health in England. One source of uncertainty is found in the specifications of the health outcome measure. There are not specific outcomes associated with some conditions where building occupants experience negative health outcomes. These conditions are linked to the occupants housing or building exposure. The set of conditions are known as "sick building symptoms" or "sick building syndrome" in the literature.

No previous evidence based on epidemiological studies was found to provide a quantitative *summary* measure of the effect of ventilation exposures on a *specific* health outcome measure at the time of this study. In research paper 3, a systematic review is conducted to identify potential negative outcomes which included allergies, asthma, wheezing, bronchial obstruction and less specific outcomes known as sick building symptoms which were suggested to be caused by inadequate ventilation.[151] In the pooling of individual studies in research paper 3, some co-morbidities changed from confounder to mediator (e.g. asthma, allergic rhinitis) depending on the specific outcome studied (e.g. sick building syndrome) (Figure 14, research paper 3).

A more general outcome measure is defined as *respiratory-related morbidity* to describe the population health effects. Such level of resolution was considered

appropriate in the case study (the objective of which is to deal with the uncertainty when there is a lack of information). In this case, the source of uncertainty arises due to insufficient evidence to quantify with a degree of accuracy the effect of the exposure on a specific outcome measure.

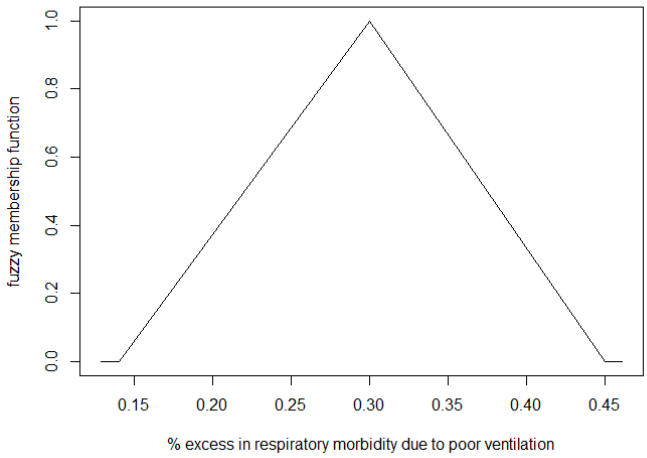
Ventilation rates are measured as air changes per hour (ACH) in some studies whilst other studies measure ventilation rates is measured in litre per second (l/s). The relationship between ventilation rate in ACH and ventilation rate in l/s is described in the following equation: $ACH = [\ ventilation\ rate\ \ l/s) \times 300\ (s/hr) \times 0.001\ (m^3/s)/$ $[room\ volume\ (m^3)]$. The volume parameter was generally taken from each study reporting ventilation rates in l/s. An additional source of uncertainty arises in the assumption of an exposure-response function below or above the threshold of 0.5 ACH. In the meta-analysis, risk comparisons are defined into two categories: ventilation rate greater than 0.5 ACH ("reference group") versus lower ventilation rate less than 0.5 ACH ("exposure group"). The effect sizes from each study that reported different types of risk comparison were standardized into a log scale by assuming a log-linear relationship of health symptoms with ventilation category in ACH.

Another potential source of uncertainty arises in the specification of the population considered exposed. There are significant variations in the definition of the target population considered to be exposed. The target population can be further refined by *who*, *when*, *where* and for *how long* people are considered exposed. However, due to lack of information about the distribution of the population of England exposed to

specific ventilation scenarios, the target population is defined as the total population of England without taking into account sub-populations.

Given the above sources of uncertainty, the exposure-response is characterised as a fuzzy set (described in Appendix B from research paper 3) to allow imprecision in knowledge. In addition, a subgroup analysis is conducted (described in Appendix A from research paper 3) to assess whether potential sources of heterogeneity could affect the result of the pooling of studies in research paper 3. Based on the result from the subgroup analysis, no significant sources of heterogeneity is assumed in the 95% CI contained within the OR of the meta-analysis. There are, however, other potential sources of uncertainty as the underlying studies in the meta-analysis did not include other sub-populations such as older people. The values from the meta-analysis are transformed into percentages via a log transformation. The central estimate in the fuzzy set is assumed to be 30% and the lower- and upper-bounds 14% and 45% respectively for the exposure response function. Figure 17 below illustrate the fuzzy set characterising the log-transformed exposure-response function assumed in the analysis (also shown in Appendix B).

Figure 17. Fuzzy set $E(x)$ with membership function describing the exposure-response function (or excess risk) in respiratory-related morbidity below 0.5 ACH.

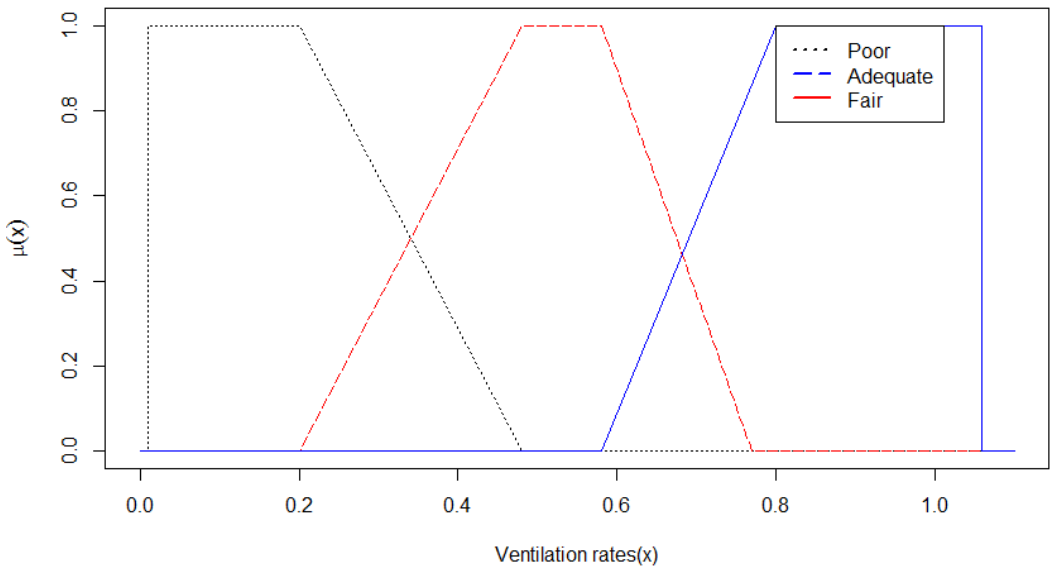


Other shapes other than triangular fuzzy sets in the propagation of uncertainty

The underlying assumptions in the three ventilation exposure scenarios described in research paper 3 warrants further investigation in exploring shapes other than triangular fuzzy sets. The central values assumed in the triangular fuzzy sets were based on expert opinion from the BRE. A universal space of 0.01-1.06 ACH of ventilation rate was also assumed and categorised in various fuzzy intervals for each ventilation scenario. Alternative assumptions were made as follows. Additional assumptions are made regarding where the input values are assumed to lie within the fuzzy intervals in the three fuzzy sets. Instead of assuming a a single value to lie within the intervals (as a central value), a range of values are assumed to lie within each fuzzy intervals: ~0.01 to 0.20 ACH *p*oor, ~0.48 to 0.58 ACH *fair* and ~ 0.80 to 1.06 ACH for *adequate* ventilation. The three ventilation exposure scenarios are specified using a trapezoidal fuzzy set as follows: *Poor* ventilation $(0.01\ ACH\ \leq x < 0.48\ ACH)$ with assumed central values between $(0.01\ ACH\ to\ 0.20ACH)$;

Fair ventilation $(0.20\ ACH\ \leq x < 0.77ACH)$ with assumed central values between $(0.48\ ACH\ to\ 0.58\ ACH)$; *Adequate* ventilation $(0.58ACH\ \leq x < 1.06\ ACH)$ with assumed central values between $(0.80ACH\ to\ 1.06\ ACH)$. Figure 18 below depict the three ventilation scenarios represented by trapezoidal fuzzy sets, in contrast to research paper 3 where triangular fuzzy sets were used originally.

Figure 18: Trapezoidal fuzzy set with three ventilation scenario



$$\mu(x)_{scenario} = \begin{cases} 0 & if \quad x < a_1, \\ \left|\frac{x-a_1}{a_2-a_1}\right| & if \quad a_1 \leq x < a_2, \\ 1 & if \quad a_2 \leq x \leq a_3, \\ \left|\frac{x-a_3}{a_4-a_3}\right| & if \quad a_3 < x \leq a_4, \\ 0 & if \quad a_4 < x, \end{cases}$$

where $a_1 \leq a_2 \leq a_3 \leq a_4 \in \mathbb{R}$ are all real numbers depicting potential ventilation rates ACH; and where $\mu(x)$ is the membership function for each ventilation scenario.

The interval of trapezoid fuzzy set is defined by specifying its lower$A_L(\alpha)$ and upper $A_U(\alpha)\alpha$-cut bounds as follows, for $\alpha \in (0,1)$ and $a_1 \leq a_2 \leq a_3 \leq a_4$:

$$A_\alpha \coloneqq [A_L(\alpha), A_U(\alpha)]$$
$$= [(a_2 - a_1) \times \alpha + a_1, \quad -(a_4 - a_2) \times \alpha + a_3],$$

Interval arithmetic operations are given in the general form as described in research paper 3.

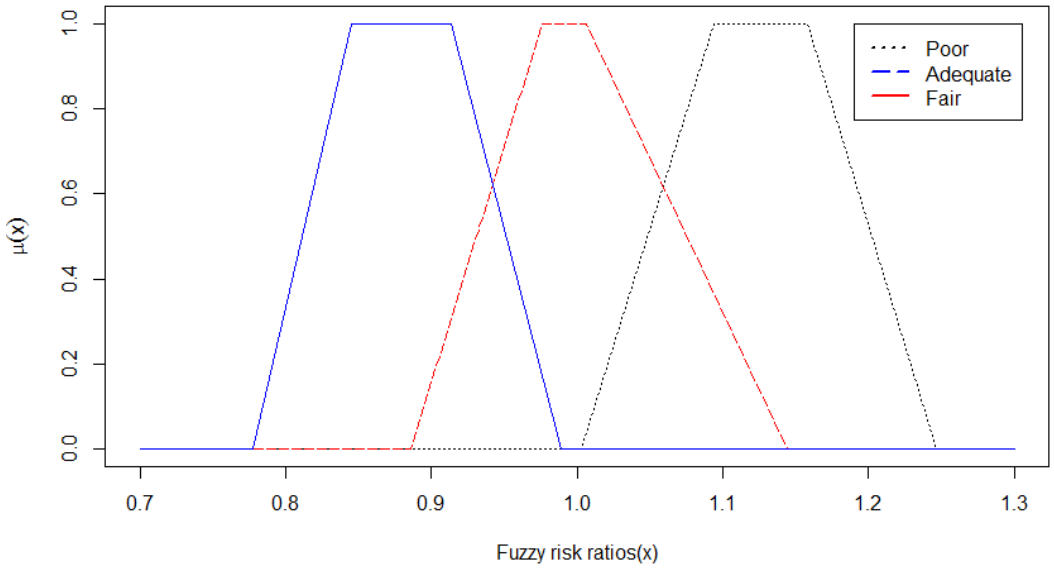
$$(A \otimes B)_\alpha = A_\alpha \otimes B_\alpha,$$

where $\otimes = +, -, * \; or \; /$ are basic arithmetic operations and A, B are arbitrary fuzzy sets.

The uncertainty contained within the interval in the fuzzy set is propagated using formulae [7-9] from research paper 3.

The result from the uncertainty propagation are given in the following two figures (Figure 19 and Figure 20). The two figures provides the resulting intervals for the trapezoidal fuzzy sets describing the uncertainty propagation of the adjusted risk ratios, and the total annual respiratory-related morbidity burdens attributable to three scenarios.

Figure 19: Trapezoidal fuzzy sets describing uncertainty propagation of the adjusted risk ratios



The resulting intervals of the adjusted risk ratio are given in Table 12 below

Table 12: Resulting fuzzy interval from adjusted risk ratio uncertainty propagation

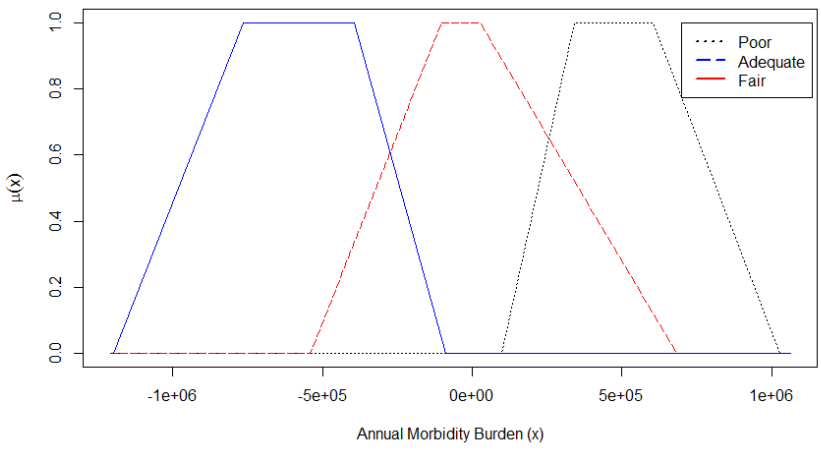
Ventilation scenarios	**Fuzzy risk ratio (intervals)**
Poor ventilation	(1.002 **1.094** **1.158** 1.246)
Fair ventilation	(0.885 **0.976** **1.006** 1.144)
Adequate ventilation	(0.777 **0.845** **0.913** 0.988)

The interval size of the fuzzy sets is preserved through the propagation of uncertainty. Different sizes of intervals are noticed for each resulting fuzzy sets in Figure 19. Due to a narrow characterisation of the central values initially in the set *Fair* (~0.48 to 0.58 ACH), the resulting interval is narrower compared to other fuzzy sets.

The resulting lower-and-upper bound of the risk ratio do not meaningfully change in the *poor,* and *fair* ventilation scenario (Table 12) compared to other triangular fuzzy sets described in research paper 3 (Figure 15). For the *adequate* ventilation scenario, a meaningful change is noticed in the risk ratio. The resulting risk ratio does not contain 1; this result is due to the initial re-categorisation of the lower bound in the *adequate* fuzzy set (i.e. 0.58 ACH minimum bound) which is reflected in the output of the upper bound of the interval.

The fuzzy sets describing the total annual respiratory-related morbidity burden are given below in Figure 20.

Figure 20: Trapezoidal fuzzy sets of the total annual respiratory-related morbidity burdens attributable to the three ventilation scenarios.



In addition, the resulting intervals of the total annual morbidity burdens are given in Table 13 below

Table 13: Resulting fuzzy interval from total annual respiratory-related morbidity burdens uncertainty propagation

Ventilation scenarios	**Total annual respiratory-related morbidity cases (n) attributable to each exposure scenario**
Poor ventilation	$(99{,}398 \quad \mathbf{339{,}730} \quad \mathbf{604{,}763} \quad 1{,}028{,}008)$
Fair ventilation	$(-539{,}846 \quad \mathbf{-101{,}502} \quad \mathbf{25{,}758} \quad 682{,}006)$
Adequate ventilation	$(-1{,}197{,}605 \quad \mathbf{-764{,}440} \quad \mathbf{-393{,}527} \quad -87{,}064)$

The corresponding (+) or (–) signs correspond to an excess of morbity cases or a reduction of morbidity cases respectively (Table 13).

Comparing the results using different shapes

Compared to results obtained in research paper 3, the resulting morbidity cases attributable to poor ventilation scenario do not change in the lower-and-upper bound of the interval. Under *poor* ventilation scenario, the uncertainty ranged from 99,398 to 1,028,008 morbidity cases, provided the population of England as whole is exposed to that scenario. The central values (bold) reflects the assumption made regarding *where* the initial range of values (ventilation rate exposures) are assumed to lie within the interval in the fuzzy sets. Provided the assumption is correct about the initial set of ventilation exposures (0.01 ACH to 0.20ACH), the total number of cases would be between 339,730 to 604,763. Under *fair* ventilation scenario, the resulting lower-and-upper bound of the interval also do not significantly change compared to results obtained in research paper 3 (i.e. the uncertainty ranged between 539,846 cases prevented to 682,006 extra cases in both results). Given how the central values are assumed (and provided the assumption is correct), the results would range from a reduction of -101,502 cases and an excess of 25,758 cases. Under an *adequate* ventilation scenario, the uncertainty ranged between -1,197,605 to -87,064 morbidity cases prevented. The upper bound of the interval changed from -48,605 cases prevented (research paper 3) to -87,064 cases prevented. This difference reflects the initial re-categorisation of the lower-bound in the adequate fuzzy set (i.e. 0.58 ACH minimum bound). Provided the assumption is correct about the initial set of ventilation exposures (0.01 ACH to 0.20ACH), the total number of cases prevented would be between -393,527 to -764,440.

Discussion

The role of the membership function plays an important role in interpreting the results obtained from the fuzzy set. Given that the membership function of the fuzzy set in the central values is $\mu(x) > 0.90$, the morbidity cases can be interpreted qualitatively as "highly probable" of being associated with the corresponding fuzzy set scenario. An important caveat is the assumption that the initial input values are correct (i.e. central values of ventilation rates in ACH are assumed to lie within the fuzzy intervals). If the membership function of the fuzzy set corresponding to morbidity cases is $\mu(x) > 0.80$, such values of morbidity cases can be interpreted qualitatively as "probable". In summary, the degree to which a value could belong to a set can be interpreted qualitatively based on its membership function. Results can be communicated qualitatively as follows: ["very highly probable" $\mu(x) > 0.90$], ["highly probable" $\mu(x) > 0.80$], ["highly" $\mu(x) > 0.65$] , ["probable" $\mu(x) > 0.50$]. Such interpretation deviates(based on this particular perspective) from traditional approaches to classical statistical inference, where the interval could have varying associated probabilities of 95% or 90 %. The fuzzy set approach does not take into account random variations due to "chance", and therefore, analogies with confidence level and significance in statistics should be avoided in the interpretation of the fuzzy set output.

6. Discussion

6.1. Introduction

Environmental health impact assessment is an area of increasing interest to public health practitioners; [66-68] there is considerable scope for innovation, particularly in the application of quantitative methodologies. [69, 70] Models are often used to quantify *ex ante* the health effects attributable to an environmental exposure or intervention. [30, 169-172] However, these models suffer from great uncertainty due to the complex associations between environmental exposures and health outcomes. [1, 173]

Current quantitative approaches are less amenable to handling the conceptual uncertainty associated with the framing assumptions. [78] This thesis attempts to address this difficulty by broadening the way uncertainty is taken into account in environmental health impact assessments, and by including framing assumptions as part of an analytical framework.

The overall aim of this study was to determine how best to quantify uncertainty in a more explicit and systematic way than in the context of environmental health impact assessment by focusing on five specific objectives which were to:

1. Describe the complexity of how uncertainty arises in environmental health impact assessment and classify the uncertainty to be amenable for quantitative modelling

2. Critically appraise the strengths and limitations of current methods used to handle the uncertainty in environmental health impact assessment.
3. Develop a novel quantitative framework for quantifying the uncertainty in environmental health impact assessment
4. Formulate two detailed case-study examples on health impact assessment of indoor housing interventions.
5. Apply the framework to the two case-studies

The next section discusses and summarises the overall findings from the thesis. **Section 6.3** addresses the contributions of thesis to the literature. **Section 6.4** discusses its limitations. **Section 6.5** identifies areas of further research and **Section 6.6** discusses the implications for applied modelling researchers and policy making. The last section provides the conclusion.

6.2. Overall finding of the thesis

The conceptual review identified different sets of concepts and tools. It adapted existing theoretical frameworks to define a quantitative framework that included a different "perspective" of uncertainty. Novel quantitative tools were applied in the field of environmental health impact assessment, building on previous approaches from other fields. This conceptual review, additionally, provided a rationale for the quantitative tools chosen in the framework. The proposed quantitative tools addressed some aspects of uncertainty that were identified in the framework. These aspects of uncertainty identified in the framework were prioritised in the application of the case study examples.

As a general summary of the findings, this study highlights the theoretical underpinnings of current uncertainty quantification methods used to deal with uncertainty in environmental health impact assessment (research paper 1). It also identifies the strengths, methods and limitations. Research paper 1 highlights the need for alternative approaches for characterising uncertainty. Because there is often limited quantitative information to use in environmental health impact assessments, [1, 174] novel approaches were identified as alternatives to probabilistic methods when dealing with limited data or information. The review in paper 1 demonstrates the need for a broader definition and perspective of uncertainty. Framing assumptions are considered to be an important part of that broader definition of uncertainty in the thesis. In addition, the review promotes better practices to incorporate sources of uncertainty associated with the framing assumptions beyond standard methods used for dealing with uncertainty. The purpose of raising issues associated with framing assumptions is to stimulate better reflections or debates about the potential intended or unintended consequences an environmental intervention can have on population health than usually recognised in environmental health impact assessment (EHIA). One example was provided in the case of biofuels in the area of EHIA. Due to the narrow definition of the assessment of impacts, other wider impacts in the area of land use, food security and consequent impacts on public health were not considered. Another example associated with issues of framing assumptions is in the study of sun exposure and skin cancer with interventions to limit sun exposures to avoid skin cancer. Sun exposure has been associated with a decrease of certain cancer and, therefore, any intervention that effectively reduce sun exposure might inadvertently cause harms.

Research paper 2 is the first analysis of the thesis. It develops a methodology to deal with “conceptual uncertainty” associated with the framing assumptions. This method quantifies the sensitivity of an HIA to the framing assumptions and determines the key pathways, which have most influence on health, by analysing the causal map linking housing insulation to health outcomes in the case study example. The case study demonstrates the potential of the method and its wider applicability. It concludes by arguing the necessity for exploring and quantifying framing assumptions prior to conducting a detailed quantitative HIA during the assessment stage. The analytical framework in research paper 2 alongside the specific quantitative tool provides an initial step forward for thinking about potential harms and the wider impacts of environmental interventions on health.

Research paper 3 provides a methodological framework for quantifying exposures scenarios and handling analytical uncertainty. A novel approach is presented to handle analytical uncertainty associated with imprecise input parameters and the outcome of a model in the context of environmental health impact assessment. This study has highlighted the key processes in the quantification of health impacts during the assessment stage of an HIA. Application of the method is illustrated by a health impact study of ventilation exposures scenarios in England. The case study showed that poor ventilation rates can be a significant contributor to the total annual number of respiratory-related morbidity cases in England. To some extent some aspects of analytical uncertainty are only addressed in research paper 3, in relation to sources of heterogeneity, variability and other aspects of parametric uncertainty identified in the case-study example.

6.3. Main contribution of the thesis

6.3.3. Approaches to quantify uncertainty in environmental health impact assessments

This study investigated quantitative approaches for dealing with uncertainty in environmental health impact assessment. It identified weaknesses in current methods and highlighted gaps in the research literature. Current methods were found to be less amenable for handling uncertainty at a more conceptual level (conceptual uncertainty), particularly when the sources of uncertainty extend beyond issues of parameters or input data. Most random sampling methods such as Monte Carlo simulations assume that there is sufficient data to make suitable assumptions on the statistics of the variability of input parameters. Bayesian methods can address issues regarding limited information on input paramters by the choice of priors which can be also elicited via expert opinions. Other probabilistic techniques can distinguish between uncertainty due random variations or uncertainty due to lack of information on input parameters such as Second Order MC simulation. However, most methods applied in environmental health impact assessment are less tractable to dealing with conceptual sources of uncertainty. The review has highlighted scope for innovation, particularly in the application of quantitative methods to deal with conceptual sources of uncertainty associated with the framing assumptions or conceptualisation of a model. The review concluded that new methods should be continue to be investigated to handle the uncertainty associated with framing assumptions.

6.3.4. Dealing with conceptual uncertainty associated with the framing assumptions

No previous work has addressed quantitatively or semi-quantitatively conceptual uncertainty associated with framing assumptions in the applied field of environmental health impact assessment. [175] This research study attempt to address this gap in the knowledge by assessing the sensitivity of the assessment to the framing assumptions. It then applies this approach to a case study of housing insulation (research paper 2). A key contribution of the study is in the use of complex system modelling approaches to evaluate an intervention. Despite calls among public health researchers in the literature, [79, 176-178] there has been no consistent attempt to apply complex systems modelling approaches in the assessment of environmental health interventions. In this study, a complex system modelling approach was applied using fuzzy cognitive mapping.

In addition, a perturbation analysis was conducted to identify assumed causal factors that are highly sensitive to the framing assumptions in the fuzzy cognitive map (e.g. ventilation, air-tightness, indoor quality etc.). Broader conceptual issues relating to the framing assumptions(rather than parameter estimation issues) are addressed in the construction of the fuzzy cognitive map. The fuzzy cognitive map (FCM) provided the basis for the assumed causal structure in the case study example. In addition, qualitative relations were constructed with the FCM in the supplementary material of chapter 4. One advantage of the approach is that FCM can helps researchers to think more broadly about the likely impacts of interventions by identifying issues relating to framing assumptions. Using FCM seems more natural over other statistically-based approaches as the process consist of providing a list of

potential relevant causal pathways and their assumed relationship. In the absence of more empirical (or statistically-based) data on the potential adverse effects of environmental interventions, the FCM can provide the basis for discussion of what adverse effect can take place. This was demonstrated in chapter 4 with the construction of two causal diagrams based on qualitative assumptions.

6.3.5. Dealing with analytical uncertainty in the inputs and outputs of a model

The analytical framework for quantifying exposure and handling analytical uncertainty was developed through the various steps: (i) selecting the exposure metric and quantifying the evidence of potential health effects of the exposure; (ii) estimating the population affected by the exposure and selecting the outcome measure; (iii) quantifying the health impacts and its uncertainty. The framework used an established method for the characterisation of uncertainty into attributable risk calculations. The study used a fuzzy set method, for the first time in the context of environmental health impact assessment, as an alternative to probabilistic approaches. Limited information to make suitable assumptions on the statistics of the input parameters was assumed in the case study example. The analytical uncertainty was quantified and propagated using fuzzy interval arithmetic. Attributable morbidity burdens were propagated corresponding to different ventilation exposure scenarios in the HIA. The study provided a general framework to help with the quantification of health impacts during the assessment stage of an HIA. Despite the existence of other approaches for the quantification of health impacts within an HIA, [54, 170, 179] it is worth noting that such quantification is still limited in practice. [180] Previous quantitative approaches have not explicitly incorporated uncertainty. [69] This study has attempted to address this limitation by providing an analytical framework

for quantifying exposures and handling uncertainty using a non-probabilistic approach.

In addition, from the specific of the case study, an exposure-response response relationship was derived from a systematic review and a meta-analysis in research paper 3. In the case-study example, limited information was found to assume information on the statistics of the variability of parameters in the propagation of uncertainty. In particular, limited information in the definition of the outcome measure to allow the effect to be quantified with a degree of precision on the outcome. The uncertainty from the effect sizes in the meta-analysis was propagated through a log transformation and fuzzy interval calculation assuming limited or imprecise information in the input parameter instead of random variations. An advantage of the fuzzy approach, in contrast to probability (where all positive values in probability can only represent an unique value of a variable), a fuzzy membership function can measures the degree to which a value can belong to two or more sets. This flexibility for fuzzy assumptions makes the fuzzy method more amenable for handling incomplete information (or imprecise parameters) in the propagation of uncertainty, particular when the evidence comes as a range from different sources that do not completely agree.

6.4. Limitations

While the study has proposed novel approaches to deal with uncertainty in the context of environmental health impact assessment, it has some limitations. The thesis has attempted to clarify and prioritise underlying issues of uncertainty in a framework. The complex set of issues of uncertainty identified in the thesis has been

addressed to a limited extent. The proposed methods only addressed some aspects of uncertainty identified in the thesis. Many issues and complexity arise in the study of uncertainty, and whether a single method could reflect accurately all the complexity of uncertainty in environmental health impact assessment is questionable. The issues of uncertainty in environmental health impact assessment would continue to be a long standing methodological problem for many years in environmental health research. The thesis's framework forms the basis on which some aspects of uncertainty are prioritised in the case studies. However, there are limitations in the methodology that should be recognised. In general, there are limitations in terms of how the new methodology for conceptual uncertainty and for analytical uncertainty can be integrated into a single analytical framework. The scope of the application of the method can vary substantially according to the specific context of application of the case studies in which they arise. This section acknowledges the limitation of the study in relation to each research paper.

6.4.6. Approaches to quantify uncertainty in environmental health impact assessments

This study reviewed uncertainty quantification methods in environmental health impact assessment based on specific search strategies. As with general systematic reviews, studies in the “grey” literature may have not been identified. The emphasis of the review was on the relatively newly established field of HIA, focusing on its applications in environmental health. Therefore, methods applied in the wider literature, and in other areas of environmental health, could have not been identified.

This study also addressed specific issues that could arise in the quantification of health impacts in relation to the framing assumptions. Conceptual uncertainty was only addressed as part of the framing assumptions. Based on this definition, framing assumptions were used to map the causal chains linking an intervention with health outcomes. However, framing assumptions are only a part of conceptual uncertainty. Other aspects of conceptual uncertainty were not fully developed such as (i) the uncertainty in defining the context of the assessment or the boundaries of the system, identified in the framework. In addition, framing assumptions can be considered and interpreted in terms of other concepts (i.e. agenda-setting, etc.) that are not necessarily related to uncertainty.

6.4.7. Conceptual uncertainty associated with the framing assumptions

Research paper 2 applied a novel method based on fuzzy cognitive mapping (FCM). FCM was used to quantify the framing assumptions in the mapping of the causal pathways of an HIA model of housing insulation. On the specific of what the method was intended to address, the method did not cover all possible framing assumptions that could potentially arise in an HIA study of housing insulation. Framing assumptions were identified through a restricted search of the literature in Ovid Medline. Social factors such as housing tenure or composition and socio-economic status were not considered in this study. In practice, a comprehensive systematic review will be necessary for the selection of factors that make up the FCM causal diagram. The studies included in the causal diagram were assumed to be comparable in terms of population intervention, study type and study quality.

In addition, sources of heterogeneity such as place where people are considered exposed alongside occupants' age, specific outcomes and/or seriousness of disease were not fully addressed in the method. The method did not allow specific sources of heterogeneity or sources of statistical significance to be assigned in each node in the construction of the FCM, and this is an important limitation. The FCM allows to emphasise a study "at the system-level" in the mapping of the causal pathways. [79, 105]

For representing uncertainty in the mapping of the causal pathways, the FCM approach can be compared with more statistically-based approaches such as Bayesian Networks (BN) , also known as Bayesian Belief Networks.[181] BN as a graphical causal model, can represent causal assumptions. BN can addresss more appropriately issues of parameter estimation in each node. Each probabilistic values in the nodes can represent the relationship between the factors in the causal pathways. Similar to a BN, the FCM is also a graphical model, however, instead of assuming probabilistic values in the nodes, fuzzy or deterministic values are assumed in the FCM which do not incorporate aspects of statistical significance or parameter estimation in the nodes. Therefore, the interpretation of results in the FCM is semi-quantitatively, as each node in the diagram is ranked in relation to each other .

6.4.8. Analytical uncertainty associated with the parameters of the environmental health impact model

In this study fuzzy set theory was used to characterise uncertainty in the parameters of a health impact model using a case study of housing ventilation exposure

scenarios. While fuzzy sets were used to describe uncertainty due to" lack of knowledge", assumptions in the fuzzy sets were made to define each exposure scenario (e.g. poor, fair or high ventilation rates). These assumptions were based on ventilation guidelines and the Building Research Establishment study and expert opion. [132, 143, 164] As fuzzy sets do not require any assumption about the probability distribution or correlation of the parameters, [20] this could be a potential limitation of the fuzzy set methods, particularly when there is probabilistic information available to make assumptions about the distribution or correlation of the parameters in the propagation of uncertainty.

To a limited extent the complex set of issues of parametric uncertainty from the case study in research paper 3 were only addressed in the fuzzy set method. Lack of understanding or information was assumed to be the primary source of uncertainty in relation to the definition of exposure scenarios, the distribution of the population to the exposure scenarios and the extrapolation of the exposure-response function to different subpopulation, alongside the assumption of a linear threshold above and below a particular value. In case study of research paper 3, expert opinions from the BRE were used to assume the location within the interval in the fuzzy set where the true values were assumed to lie in selecting the fuzzy set membership function for the exposure scenarios. In contrast, some probabilistic approaches can deal with both random variability and imprecise parameters as in the case of second order Monte Carlo simulation or a Bayesian approach (it is important to distinguish the uncertainty initially in both probabilistic approaches). In terms of the computational aspects, both probabilistic approaches and fuzzy approaches in the operation of joint

distributions or operations with membership functions can be challenging when dealing with large number of input parameters.

6.5. Areas of further research

This study has identified some areas of further research in relation to quantifying uncertainty in environmental health impact assessment. The main premise of the framework proposed in this thesis is that it is difficult to simultaneously examine uncertainty from different perspectives. As stakeholders and analysts often have different viewpoints with regard to uncertainty, two main perspective of uncertainty (conceptual and analytical) were chosen as a way of narrowing the examination of uncertainty.

However, it is important to note that the above two perspectives covered are not sufficient to address all the types of uncertainties that may arise at different stages in the assessment. There is indeed an additional stage in which uncertainty can take place in an environmental health impact assessment. This can be defined as the "decision stage" and can potentially be included in the proposed uncertainty framework as a component of a different uncertainty perspective. The decision stage is concerned with how to *process* uncertainty in the outcome of the assessment model for decision-making purposes. Uncertainty could be a normal condition of the policy process for decision makers. Methods from a decision stage in any assessment can be described using the terms "risk management" and "optimisation". Risk management evaluates decision options from a decision-maker point of view. Optimisation can be viewed as a methodology for risk management - where the best options are searched and evaluated when working within a set of constraints.

Optimisation methods could potentially be used to extend this framework to cater for the decision uncertainty perspective. Mathematical programming (MP) methods can be used to maximise or minimise a desired utility,[182-184] often in search of the "best option". In mathematical terms, a desired utility is often called an objective function. In the context of environmental health impact assessment, the objective function can be defined to maximise health gain or minimise health burdens. Optimisation problems, additionally, can be formulated to include constraints such as reducing health inequalities or achieving a "trade-off" between health and other characteristics of the environment.[185] Therefore, potential scope for improvement can exist in the application of uncertainty quantification methods with optimisation techniques as part of an environmental health impact assessment. For example, in the case study illustration of housing ventilation exposure scenarios, optimisation techniques could be used to model the uncertainty of not achieving a safe minimum level of ventilation in homes subject to constraints. Constraints could be defined in terms of housing parameters (e.g. ventilation rates, air tightness, temperature, indoor air quality). These constraints could subsequently be expressed in terms of mathematical inequalities to express relationships between each different indoor factors (physical building parameters). The uncertainty in the model parameter could be incorporated using uncertainty sets for the optimisation. In general, optimisation techniques have been applied in environmental engineering and energy planning for decision-making under uncertainty. [186] Future application of these techniques can be explored in the applied field of environmental health impact assessment.

Another area of research work is dealing with structural uncertainty in environmental health impact assessment models. Structural uncertainty can be explored as part of

the analytical perspective of the framework. Structural uncertainty relates to the configurations chosen in the model, particularly, how parameters and functional equations are specified in the model structure. For example, structural uncertainty can arise in deciding whether an additive or linear term is chosen in the equations of a health impact assessment model. Structural uncertainty can be dealt explicitly using model averaging techniques, where different model specifications or alternative competing models can be explored based on their parameters performance or predictive abilities. Techniques for dealing with structural uncertainty has been explored in other applied field in decision-analytical models.[187]

6.5.9. Scope for further research in the method proposed

Fundamentally, future research work should investigate how best to integrate uncertainty across the two perspectives identified in this thesis (conceptual, analytical). As part of conceptual uncertainty, there is future scope for development when sufficient data is collected for parameter estimation. A fuzzy cognitive map (FCM) approach can be integrated with more statistically-based approaches such as Bayesian Networks (BN).[181] Generating conditional probabilities in a Bayesian Network can prove to be difficult. Therefore, an initial FCM can be constructed to provide the basis of the causal structure of a BN as the process is more direct in the FCM than estimating conditional probabilities. For example, an initial FCM can be constructed when there is insufficient data and then the causal structure can be converted into a BN once sufficient data on parameters are estimated. Broader conceptual issues can be addressed by the FCM as a first step, and conditional probabilities can be addressed by a BN as a second step provided sufficient data is available (or suitable assumptions on the statistics of parameters can be made by

parameter estimations or Bayesian expert elicitation). To assess whether an improper method for propagating the uncertainty might introduce additional uncertainties in the assessments, the propagation of uncertainty can be calculated using fuzzy intervals as a first step without assuming random variations. As a second step, other probabilistic approaches can be conducted as part of analytical uncertainty. For example, in the estimation of an exposure-response relationship if there are insufficient data observations to assume distributions, the fuzzy method can be explored via expert elicitation. Once sufficient data becomes available via expert consensus, a Bayesian approach can be conducted, or if empirical data is available, the exposure-response function can be modelled using generalized additive models with cubic regression splines or other statistically-based approaches. Final outcomes of the model can be given in summary measures of population health such as disability-adjusted life years (DALYs) or others summary measures to allow comparison between policies or exposure scenarios, and comparison with the overall uncertainty between methods.

Futurer research work should also explore how sources of analytical uncertainty and conceptual can be integrated. Techniques such as "global sensitivity analysis" can be used to investigate the effect of changing assumptions or definitions adopted at each perspective.[51] This will help explore how the sources of uncertainty from the problem formulation (conceptual perspective) can impact the uncertainty from the model result (analytical perspective). The idea is to generate different policy alternatives that would have not been considered otherwise. Different alternatives can be explored by using information obtained from the two perspectives. Although it is often difficult to manage "all issues" from all "perspectives" simultaneously, the

integration of uncertainty can be conducted by focusing only on the issues that are shared across each perspective. In this way, we can simplify the integration down to the key cross-cutting issues identified.

6.6. Implications for researchers and policy makers

An important question to researchers and policy makers is to what extent an environmental intervention or exposure can produce health-related changes. Environmental health impact assessment identifies possible health consequences of new environmental policies, intervention or potential scenarios and it is an area of increasing interest to policy makers. It is important to note that most environmental health impact assessments seek to assess the health impacts of an intervention or exposure scenarios *ex ante*. As a result, quantifying the health impacts of an environmental intervention or potential exposure scenarios is particularly important for decision support.

In some assessments, the quantification of health impacts it is still limited in practice, even though it is preferred by policy makers.[180] This limitation is partly due to lack of quantitative information and large uncertainty arising during the assessment stage. Current models are grounded on probabilistic approaches for characterising the uncertainty. These approaches are frequently based on statistical techniques which include Monte Carlo and Bayesian analyses in general.[188-191] Most health risk and impact assessment exercises exclude the selection of the framing assumptions when appraising the uncertainty.[6] Decision makers should be aware that framing assumptions can have a significant impact in the outcome of the assessment.

This study encourages the use of analytical framework to deal with uncertainty when presented with limited information. One important consideration for applied researchers is that lack of probabilistic information should not prevent them from quantifying sources of uncertainty in relation to “lack or limitations in knowledge”. For example, the analysis summarised in Figure 15 (chapter 5) depicts the adjusted risk ratio based on fuzzy interval calculations for each ventilation exposure scenario. The risk ratio associated with each ventilation exposure scenarios describes imprecision (or vagueness) in the definition of risk of poor ventilation rates on health outcomes based on its handling of incomplete information in the input parameters. Although there are sources of parametric uncertainty and imprecision in knowlegde, results are consistent with the view that there is a continuous exposure-response relationship both above and below 0.5 air changes per hour. In addition, results from chapter 5 provide an indication of the potential public health benefits that would be realised in England provided adequate ventilation is achieved. The extent to which results could likely represent a “real effect” with potential caveats about the uncertainty of the initial conditions can be communicated to the policy maker. As such, chapter 5 results can be interpreted as an approximation of the risk and benefits (impacts on respiratory-related conditions) associated with the three ventilation exposure scenarios.

Chapter 4 characterises the framing assumptions that are ultimately captured by a "centrality index" and then determines the sensitivity of the framing assumptions captured by the "causal activity level". The process consists of first characterising the structure (centrality index) and then exploring how the structure affects its function (causal activity level). Chapter 4 results can reveal how potential caveats

about the causal interpretation (causal structure) can potentially affect the analysis. It provides a warning on which indoor factor to consider as they could have the largest impacts on the outcome.

7. Conclusion

The overall aim of this study was to determine how best to deal with quantitative measures of uncertainty in the context of environmental health impact assessment. A primary objective of the thesis was to develop a quantitative framework to deal with uncertainty. The thesis prioritised issues of uncertainty identified in a framework that has given less attention in environmental health impact assessment. Two dimensions of uncertainty were identified in the thesis: "nature" and "location". The "nature" of uncertainty is dealt with using lack of knowledge and propagated in a non-probabilistic domain. The "location" of uncertainty is dealt with using two perspectives: "conceptual" and "analytical". The key assumptions for dealing with the "nature" of uncertainty depend on how the uncertainty is defined in terms of its underlying causes in the propagation of uncertainty. The key assumption for dealing with the "location" of uncertainty depends on how the uncertainty is focused in terms of "where" the uncertainty manifests in the assessment. The review of current uncertainty quantification methods highlighted the need for a broader definition of uncertainty to include conceptual sources of uncertainty associated with the framing assumptions in an environmental health impact assessment. Under the proposed framework, this thesis has identified potential analytical tools as a first step for dealing with analytical and conceptual uncertainty in a non-probabilistic domain. The added value of this research was in the attempt to include, rather than exclude the framing assumptions quantitatively or semi-quantitatively in the appraisal of

uncertainty. This study challenges researchers to think as broadly as possible in the assessment of impacts in environmental health impact assessment. As part of a lesson learned throughout the thesis, four practical recommendations to researchers are given when confronted with large uncertainty in environmental health impact assessment. To conclude this thesis, the four recommendations are given as follows:

(i) "Think as broadly as possible in the assessment of impacts."

(ii) "Do not always focus on parametric sources of uncertainty."

(iii) "Do not always assume random variability when confronted with lack of understanding or limited information."

(iv) "After considering the above three suggestions, continue to model as usual."

Thank you!

Marco Mesa-Frias

References

1. Briggs DJ. A framework for integrated environmental health impact assessment of systemic risks. Environmental Health. 2008; **7**: 61.

2. Knol A, Petersen A, van der Sluijs J, Lebret E. Dealing with uncertainties in environmental burden of disease assessment. Environmental Health. 2009; **8**(1): 21.

3. Walker WE, Harremoës P, Rotmans J, van der Sluijs JP, van Asselt MBA, Janssen P, et al. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment. 2003; **4**(1): 5 - 17.

4. Briggs DJ, Sabel CE, Lee K. Uncertainty in epidemiology and health risk and impact assessment. Environmental Geochemistry and Health. 2009; **31**(2): 189-203.

5. van der Sluijs JP, Janssen PHM, Petersen AC, Kloprogge P, Risbey JS, Tuinstra W, et al. Tool Catalogue for Uncertainty Assessment; 2004.

6. Ramsey M. Uncertainty in the assessment of hazard, exposure and risk. Environmental Geochemistry and Health. 2009; **31**(2): 205-17.

7. Sluijs J, Janssen P, Petersen A, Kloprogge P, Risbey J, Tuinstra W, et al. RIVM/MNP Guidance for UncertaintyAssessment and Communication: Tool Catalogue for Uncertainty Assessment: Utrecht/Bilthoven, Copernicus Institute & RIVM; 2003.

8. Glasbergen P. Learning to Manage the Environment. In: Lafferty W, Meadowcroft J, editors. Democracy and the Environment: Problems and Prospects. Cheltenham: Edward Elgar; 1999. p. 175-93.

9. Fernandez A, Rallo R, Giralt F. Uncertainty reduction in environmental data with conflicting information. Environmental science & technology. 2009; **43**(13): 5001-6.

10. McElhany P, Steel EA, Avery K, Yoder N, Busack C, Thompson B. Dealing with uncertainty in ecosystem models: lessons from a complex salmon model. Ecological applications : a publication of the Ecological Society of America. 2010; **20**(2): 465-82.

11. van der Sluijs J, Craye M, Funtowicz S, Kloprogge P, Ravetz J, Risbey J. Experiences with the NUSAP system for multidimensional uncertainty assessment. Water science and technology : a journal of the International Association on Water Pollution Research. 2005; **52**(6): 133-44.

12. Kosko B. Fuzzy cognitive maps. International Journal of Man-Machine Studies. 1986; **24**(1): 65-75.

13. van Vliet M, Kok K, Veldkamp T. Linking stakeholders and modellers in scenario studies: The use of Fuzzy Cognitive Maps as a communication and learning tool. Futures. 2010; **42**(1): 1-14.

14. Özesmi U, Özesmi SL. Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecological Modelling. 2004; **176**(1-2): 43-64.

15. Xirogiannis G, Stefanou J, Glykas M. A fuzzy cognitive map approach to support urban design. Expert Systems with Applications. 2004; **26**(2): 257-68.

16. Jaulin L, Kieffer M, Didrit O, Walter E. Applied interval analysis. London: Springer-Verlag; 2001.

17. Ferson S. RAMAS Risk Calc: Risk Assessment with uncertain Numbers. Boca Raton, FL: Lewis Press; 2002.

18. Zadeh LA. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems. 1978; **1**(1): 3-28.

19. Dubois D, Prade H. Fundamentals of Fuzzy sets. Boston: Kluwer Academic Publishers; 2000.

20. Smithson M, Verkuilen J. Fuzzy set theory: applications in the social sciences Thousand Oaks, California: SAGE; 2006.

21. Ricci PF, Cox LA, Jr, MacDonald TR. Precautionary principles: a jurisdiction-free framework for decision-making under risk. Human and Experimental Toxicology. 2004; **23**(12): 579-600.

22. Borrego C, Monteiro A, Ferreira J, Miranda AI, Costa AM, Carvalho AC, et al. Procedures for estimation of modelling uncertainty in air quality assessment. Environ Int. 2008; **34**(5): 613-20.

23. McIntyre N, Wheater H, Lees M. Estimation and propagation of parametric uncertainty in environmental models. Journal of Hydroinformatics. 2002; (4): 177-98

24. van der Voet H, van der Heijden GW, Bos PM, Bosgra S, Boon PE, Muri SD, et al. A model for probabilistic health impact assessment of exposure to food chemicals. Food Chem Toxicol. 2009; **47**(12): 2926-40.

25. Orru H, Teinemaa E, Lai T, Tamm T, Kaasik M, Kimmel V, et al. Health impact assessment of particulate pollution in Tallinn using fine spatial resolution and modeling techniques. Environmental Health. 2009; **8**(1): 7.

26. Vlachokostas C, Achillas C, Moussiopoulos, Hourdakis E, Tsilingiridis G, Ntziachristos L, et al. Decision support system for the evaluation of urban air pollution control options: Application for particulate pollution in Thessaloniki, Greece. Science of The Total Environment. 2009; **407**(23): 5937-48.

27. Tainio M, Tuomisto J, Hanninen O, Ruuskanen J, Jantunen M, Pekkanen J. Parameter and model uncertainty in a life-table model for fine particles (PM2.5): a statistical modeling study. Environmental Health. 2007; **6**(1): 24.

28. Kuo T, Jarosz CJ, Simon P, Fielding JE. Menu Labeling as a Potential Strategy for Combating the Obesity Epidemic: A Health Impact Assessment. Am J Public Health. 2009; **99**(9): 1680-6.

29. Fehr R, Mekel O, Lacombe M, Wolf U. Towards health impact assessment of drinking-water privatization--the example of waterborne carcinogens in North Rhine-Westphalia (Germany). Bull World Health Organ. 2003; **81**(6): 408-14.

30. Schram-Bijkerk D, van Kempen E, Knol AB, Kruize H, Staatsen B, van Kamp I. Quantitative health impact assessment of transport policies: two simulations related to speed limit reduction and traffic re-allocation in the Netherlands. Occupational and environmental medicine. 2009; **66**(10): 691-8.

31. Koornneef J, Spruijt M, Molag M, Ramírez A, Turkenburg W, Faaij A. Quantitative risk assessment of CO2 transport by pipelines--A review of uncertainties and their impacts. Journal of Hazardous Materials. 2010; **177**(1-3): 12-27.

32. Özkaynak H, Frey HC, Burke J, Pinder RW. Analysis of coupled model uncertainties in source-to-dose modeling of human exposures to ambient air pollution: A PM2.5 case study. Atmospheric Environment. 2009; **43**(9): 1641-9.

33. Vicari AS, Mokhtari A, Morales RA, Jaykus L-A, Frey HC, Slenning BD, et al. Second-Order Modeling of Variability and Uncertainty in Microbial Hazard Characterization. Journal of Food Protection. 2007; **70**: 363-72.

34. Vardoulakis S, Chalabi Z, Fletcher T, Grundy C, Leonardi GS. Impact and uncertainty of a traffic management intervention: population exposure to polycyclic aromatic hydrocarbons. Science of The Total Environment. 2008; **394**(2-3): 244-51.

35. Liu Y, Guo H, Mao G, Yang P. A Bayesian hierarchical model for urban air quality prediction under uncertainty. Atmospheric Environment. 2008; **42**(36): 8464-9.

36. Li H, Huang G, Zou Y. An integrated fuzzy-stochastic modeling approach for assessing health-impact risk from air pollution. Stochastic Environmental Research and Risk Assessment. 2008; **22**(6): 789-803.

37. Hauck M, Huijbregts MAJ, Armitage JM, Cousins IT, Ragas AMJ, van de Meent D. Model and input uncertainty in multi-media fate modeling: Benzo[a]pyrene concentrations in Europe. Chemosphere. 2008; **72**(6): 959-67.

38. Tainio M, Tuomisto JT, Pekkanen J, Karvosenoja N, Kupiainen K, Porvari P, et al. Uncertainty in health risks due to anthropogenic primary fine particulate matter from different source types in Finland. Atmospheric Environment. 2010; **44**(17): 2125-32.

39. Ragas AMJ, Brouwer FPE, Buchner FL, Hendriks HWM, Huijbregts MAJ. Separation of uncertainty and interindividual variability in human exposure modeling. J Expos Sci Environ Epidemiol. 2008; **19**(2): 201-12.

40. Li J, Huang GH, Zeng G, Maqsood I, Huang Y. An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination. Journal of Environmental Management. 2007; **82**(2): 173-88.

41. Pelekis M, Nicolich MJ, Gauthier JS. Probabilistic Framework for the Estimation of the Adult and Child Toxicokinetic Intraspecies Uncertainty Factors. Risk Analysis. 2003; **23**(6): 1239-55.

42. Roberts S, Martin MA. Bootstrap-after-Bootstrap Model Averaging for Reducing Model Uncertainty in Model Selection for Air Pollution Mortality Studies. Environ Health Perspect. 2009; **118**(1).

43. Barreto H, Howland FM. Introductory Econometrics using Monte Carlo simulation with Microsoft Excel. New York: Cambridge University Press; 2005.

44. Burmaster DE, Anderson PD. Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments. Risk Analysis. 1994; **14**(4): 477-81.

45. Goodman IR, Nguyen HT. Probability updating using second order probabilities and conditional event algebra. Information Sciences. 1999; **121**(3-4): 295-347.

46. Raftery AE, Madigan D, Hoeting JA. Bayesian Model Averaging for Linear Regression Models. Journal of the American Statistical Association. 1997; **92**(437): 179-91.

47. Shwarz G. Estimating the dimension of a model. Annals of Statistics. 1978; **6**(2): 461-4.

48. Akaike H. Information theory as an extension of the maximum likelihood principle. In: Petrov BN, editor. Second International Symposium on Information Theory. Budapest, Hungary: Akademiai Kiado; 1973. p. 267-81.

49. Smith TJ, Marshall LA. Bayesian methods in hydrologic modeling: A study of recent advancements in Markov chain Monte Carlo techniques. Water Resoures Research. 2008; **44**: W00B5.

50. Bárdossy G, Fodor J. The concept of geological uncertainties and new ways of their geomathematical evaluation. Geological Society of America Special Papers. 2006; **397**: 211-5.

51. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global Sensitivity Analysis: The Primer: WileyBlackwell; 2008.

52. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic Sensitivity Analysis Using Monte Carlo Simulation. Medical Decision Making. 1985; **5**(2): 157-77.

53. Lhachimi SK, Nusselder WJ, Boshuizen HC, Mackenbach JP. Standard Tool for Quantification in Health Impact Assessment: A Review. American Journal of Preventive Medicine. 2010; **38**(1): 78-84.

54. Veerman JL, Barendregt JJ, Mackenbach JP. Quantitative health impact assessment: current practice and future directions. Journal of epidemiology and community health. 2005; **59**(5): 361-70.

55. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering & System Safety. 2003; **81**(1): 23-69.

56. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. Journal of the American Statistical Association. 2005; (100): 680-701.

57. Young GA, Smith RL. Essentials of Statistical Inference: Cambridge University Press; 2005.

58. Ferson S. What Monte Carlo methods cannot do. Human and Ecological Risk Assessment: An International Journal. 1996; **2**(4): 990 - 1007.

59. Madill C, Shihab-Eldin A. Assessment of Biofuels: Potential and Limitations: International Energy Forum; 2010.

60. Janssen PH, Petersen AC, van der Sluijs JP, Risbey JS, Ravetz JR. A guidance for assessing and communicating uncertainties. Water Science & Technology. 2005; **52**(6): 125-31.

61. van der Rhee HJ, de Vries E, Coebergh JW. Does sunlight prevent cancer? A systematic review. European journal of cancer (Oxford, England : 1990). 2006; **42**(14): 2222-32.

62. Northridge ME, Sclar ED, Biswas P. Sorting out the connections between the built environment and health: a conceptual framework for navigating pathways and planning healthy cities. Journal of urban health : bulletin of the New York Academy of Medicine. 2003; **80**(4): 556-68.

63. de Blasio A, Giran J, Nagy Z. Potentials of health impact assessment as a local health policy supporting tool. Perspect Public Health. 2012; **132**(5): 216-20.

64. Mindell J, Sheridan L, Joffe M, Samson-Barry H, Atkinson S. Health impact assessment as an agent of policy change: improving the health impacts of the mayor of London's draft transport strategy. Journal of epidemiology and community health. 2004; **58**(3): 169-74.

65. Kemm J. What is health impact assessment and what can it learn from EIA? Environmental Impact Assessment Review. 2004; **24**(2): 131-4.

66. Maire N, Winkler MS, Krieger GR, Divall MJ, Singer BH, Utzinger J. Health impact assessment of industrial development projects: a spatio-temporal visualization. Geospatial health. 2012; **6**(2): 299-301.

67. Dhondt S, Kochan B, Beckx C, Lefebvre W, Pirdavani A, Degraeuwe B, et al. Integrated health impact assessment of travel behaviour: model exploration and application to a fuel price increase. Environ Int. 2013; **51**: 45-58.

68. de Nazelle A, Nieuwenhuijsen MJ, Anto JM, Brauer M, Briggs D, Braun-Fahrlander C, et al. Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment. Environ Int. 2011; **37**(4): 766-77.

69. Fehr R, Hurley F, Mekel OC, Mackenbach JP. Quantitative health impact assessment: taking stock and moving forward. Journal of epidemiology and community health. 2012; **66**(12): 1088-91.

70. Mindell J, Joffe M. Mathematical modelling of health impacts. Journal of epidemiology and community health. 2005; **59**(8): 617-8.

71. DCLG. Building Regulations Explanatory Booklet. London: Department of Communities and Local Goverment; 2003.

72. Wilkinson P, Smith KR, Beevers S, Tonne C, Oreszczyn T. Energy, energy efficiency, and the built environment. The Lancet. 2007; **370**(9593): 1175-87.

73. Barton A, Basham M, Foy C, Buckingham K, Somerville M. The Watcombe Housing Study: the short term effect of improving housing conditions on the health of residents. Journal of epidemiology and community health. 2007; **61**(9): 771-7.

74. Bone A, Murray V, Myers I, Dengel A, Crump D. Will drivers for home energy efficiency harm occupant health? Perspect Public Health. 2010; **130**(5): 233-8.

75. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008; **337**: a1655.

76. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ. 2000; **321**(7262): 694-6.

77. Wilkinson P, Smith KR, Davies M, Adair H, Armstrong BG, Barrett M, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: household energy. Lancet. 2009; **374**(9705): 1917-29.

78. Mesa-Frias M, Chalabi Z, Vanni T, Foss AM. Uncertainty in environmental health impact assessment: Quantitative methods and perspectives. International Journal of Environmental Health Research. 2012; **23**(1): 16-30.

79. Joffe M, Gambhir M, Chadeau-Hyam M, Vineis P. Causal diagrams in systems epidemiology. Emerging themes in epidemiology. 2012; **9**(1): 1.

80. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. Int J Epidemiol. 2010; **39**(1): 97-106.

81. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. BMJ. 2008; **336**(7656): 1281-3.

82. Kitchin RM. Cognitive maps: What are they and why study them? Journal of Environmental Psychology. 1994; **14**(1): 1-19.

83. Wood MD, Bostrom A, Bridges T, Linkov I. Cognitive Mapping Tools: Review and Risk Management Needs. Risk Analysis. 2012; **32**(8): 1333-48.

84. Nakagawa Y, Shiroyama H, Kuroda K, Suzuki T. Assessment of social implications of nanotechnologies in Japan: Application of problem structuring

method based on interview surveys and cognitive maps. Technological Forecasting and Social Change. 2010; **77**(4): 615-38.

85. Axelrod R. Structure of Decision: The cognitive maps of political elites. Princeton, NJ: Princeton University Press; 1976.

86. Giles BG, Findlay CS, Haas G, LaFrance B, Laughing W, Pembleton S. Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. Social Science & Medicine. 2007; **64**(3): 562-76.

87. West DB. Introduction to Graph Theory Prentice Hall; 2000.

88. Chapman R, Howden-Chapman P, Viggers H, O'Dea D, Kennedy M. Retrofitting houses with insulation: a cost–benefit analysis of a randomised community trial. Journal of epidemiology and community health. 2009; **63**(4): 271-7.

89. Howden-Chapman P, Crane J, Matheson A, Viggers H, Cunningham M, Blakely T, et al. Retrofitting houses with insulation to reduce health inequalities: aims and methods of a clustered, randomised community-based trial. Social Science & Medicine. 2005; **61**(12): 2600-10.

90. Jackson G, Thornley S, Woolston J, Papa D, Bernacchi A, Moore T. Reduced acute hospitalisation with the healthy housing programme. Journal of epidemiology and community health. 2011; **65**(7): 588-93.

91. Howden-Chapman P, Matheson A, Crane J, Viggers H, Cunningham M, Blakely T, et al. Effect of insulating existing houses on health inequality: cluster randomised study in the community. BMJ. 2007; **334**(7591): 460.

92. Levy JI, Nishioka Y, Spengler JD. The public health benefits of insulation retrofits in existing housing in the United States. Environmental Health. 2003; **2**(1): 4.

93. Howden-Chapman P, Saville-Smith K, Crane J, Wilson N. Risk factors for mold in housing: a national survey. Indoor Air. 2005; **15**(6): 469-76.

94. Gilbertson J, Stevens M, Stiell B, Thorogood N. Home is where the hearth is: Grant recipients' views of England's Home Energy Efficiency Scheme (Warm Front). Social Science & Medicine. 2006; **63**(4): 946-56.

95. Engvall K, Norrby C, Norback D. Ocular, nasal, dermal and respiratory symptoms in relation to heating, ventilation, energy conservation, and reconstruction of older multi-family houses. Indoor Air. 2003; **13**(3): 206-11.

96. Brugge D, Vallarino J, Ascolillo L, Osgood ND, Steinbach S, Spengler J. Comparison of multiple environmental factors for asthmatic children in public housing. Indoor Air. 2003; **13**(1): 18-27.

97. Rudge J. British weather: conversation topic or serious health risk? Int J Biometeorol. 1996; **39**(3): 151-5.

98. Vandentorren S, Bretin P, Zeghnoun A, Mandereau-Bruno L, Croisier A, Cochet C, et al. August 2003 heat wave in France: Risk factors for death of elderly people living at home. European Journal of Public Health. 2006; **16**(6): 583-91.

99. Hirsch T, Hering M, Bürkner K, Hirsch D, Leupold W, Kerkmann ML, et al. House-dust-mite allergen concentrations (Der f 1) and mold spores in apartment bedrooms before and after installation of insulated windows and central heating systems. Allergy. 2000; **55**(1): 79-83.

100. Fisk WJ, Mirer AG, Mendell MJ. Quantitative relationship of sick building syndrome symptoms with ventilation rates. Indoor Air. 2009; **19**(2): 159-65.

101. Braubach M. Residential conditions and their impact on residential environment satisfaction and health: Results of the WHO large analysis and review of European housing and health status (LARES) study. International Journal of Environment and Pollution. 2007; **30**(3-4): 384-403.

102. Mendell MJ. Indoor residential chemical emissions as risk factors for respiratory and allergic effects in children: a review. Indoor Air. 2007; **17**(4): 259-77.

103. Smith KR, McCracken JP, Weber MW, Hubbard A, Jenny A, Thompson LM, et al. Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial. The Lancet. 2011; **378**(9804): 1717-26.

104. Fisk WJ, Lei-Gomez Q, Mendell MJ. Meta-analyses of the associations of respiratory health effects with dampness and mold in homes. Indoor Air. 2007; **17**(4): 284-96.

105. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. BMJ. 2007; **334**(7591): 455-9.

106. Parry J, Stevens A. Prospective health impact assessment: pitfalls, problems, and possible ways forward. BMJ. 2001; **323**(7322): 1177-82.

107. Lipsey M, Wilson D. Practical Meta Analysis. London: SAGE; 2001.

108. Bueno S, Salmeron JL. Benchmarking main activation functions in fuzzy cognitive maps. Expert Systems with Applications. 2009; **36**(3, Part 1): 5221-9.

109. Keall M, Baker MG, Howden-Chapman P, Cunningham M, Ormandy D. Assessing housing quality and its impact on health, safety and sustainability. Journal of epidemiology and community health. 2010; **64**(9): 765-71.

110. Glytsos T, Ondrácek J, Dzumbová L, Kopanakis I, Lazaridis M. Characterization of particulate matter concentrations during controlled indoor activities. Atmospheric Environment. 2010; **44**(12): 1539-49.

111. Thomson H, Atkinson R, Petticrew M, Kearns A. Do urban regeneration programmes improve public health and reduce health inequalities? A synthesis of the evidence from UK policy and practice (1980–2004). Journal of epidemiology and community health. 2006; **60**(2): 108-15.

112. Chapman R, Howden-Chapman P, Viggers H, O'Dea D, Kennedy M. Retrofitting houses with insulation: a cost–benefit analysis of a randomised community trial. Journal of Epidemiology and Community Health. 2009; **63**(4): 271-7.

113. Howden-Chapman P, Pierse N, Nicholls S, Gillespie-Bennett J, Viggers H, Cunningham M, et al. Effects of improved home heating on asthma in community dwelling children: randomised controlled trial. BMJ. 2008; **337**: a1411.

114. Thomson H, Thomas S, Sellstrom E, Petticrew M. The Health Impacts of Housing Improvement: A Systematic Review of Intervention Studies From 1887 to 2007. Am J Public Health. 2009; **99**(S3): S681-92.

115. Thomson H, Petticrew M, Morrison D. Health effects of housing improvement: systematic review of intervention studies. BMJ. 2001; **323**(7306): 187-90.

116. Howden-Chapman P. Housing standards: a glossary of housing and health. Journal of epidemiology and community health. 2004; **58**(3): 162-8.

117. Bonnefoy X, Braubach M, Davidson M, Roebbel N. A pan-European housing and health survey: Description and evaluation of methods and approaches. International Journal of Environment and Pollution. 2007; **30**(3-4): 363-83.

118. Bone A, Murray V, Myers I, Dengel A, Crump D. Will drivers for home energy-efficiency harm occupant health? Perspectives in Public Health. 2010.

119. Pearl J. Causality: models, reasoning, and inference. New York: Cambridge University Press; 2000.

120. Adam B, Molnar A, Gulis G, Adany R. Integrating a quantitative risk appraisal in a health impact assessment: analysis of the novel smoke-free policy in Hungary. Eur J Public Health. 2013; **23**(2): 211-7.

121. Bhatia R, Seto E. Quantitative estimation in Health Impact Assessment: Opportunities and challenges. Environmental Impact Assessment Review. 2011; **31**(3): 301-9.

122. Lhachimi SK, Nusselder WJ, Smit HA, van Baal P, Baili P, Bennett K, et al. DYNAMO-HIA--a Dynamic Modeling tool for generic Health Impact Assessments. PLoS One. 2012; **7**(5): e33317.

123. Liu HY, Bartonova A, Pascal M, Smolders R, Skjetne E, Dusinska M. Approaches to integrated monitoring for environmental health impact assessment. Environmental health : a global access science source. 2012; **11**: 88.

124. Mesa-Frias M, Chalabi Z, Foss AM. Assessing framing assumptions in quantitative health impact assessments: A housing intervention example. Environment International. 2013; **59**(0): 133-40.

125. DCLG. Code for Sustainable Homes: Technical Guide version 2. London: Communities and Local Government; 2009.

126. Manuel J. Avoiding health pitfalls of home energy-efficiency retrofits. Environ Health Perspect. 2011; **119**(2): A76-9.

127. Stephens B, Carter EM, Gall ET, Earnest CM, Walsh EA, Hun DE, et al. Home energy-efficiency retrofits. Environ Health Perspect. 2011; **119**(7): A283-4.

128. Stymne H, Axel Boman C, Kronvall J. Measuring ventilation rates in the Swedish housing stock. Building and Environment. 1994; **29**(3): 373-9.

129. Zuraimi MS, Tham KW, Chew FT, Ooi PL. The effect of ventilation strategies of child care centers on indoor air quality and respiratory health of children in Singapore. Indoor Air. 2007; **17**(4): 317-27.

130. Wright GR, Howieson S, McSharry C, McMahon AD, Chaudhuri R, Thompson J, et al. Effect of improved home ventilation on asthma control and house dust mite allergen levels. Allergy. 2009; **64**(11): 1671-80.

131. Engvall K, Norrby C, Norback D. Sick building syndrome in relation to building dampness in multi-family residential buildings in Stockholm. International archives of occupational and environmental health. 2001; **74**(4): 270-8.

132. Dimitroulopoulou C. Ventilation in European dwellings: A review. Building and Environment. 2012; **47**(0): 109-25.

133. Sundell J, Levin H, Nazaroff WW, Cain WS, Fisk WJ, Grimsrud DT, et al. Ventilation rates and health: multidisciplinary review of the scientific literature. Indoor Air. 2011; **21**(3): 191-204.

134. Li Y, Leung GM, Tang JW, Yang X, Chao CY, Lin JZ, et al. Role of ventilation in airborne transmission of infectious agents in the built environment - a multidisciplinary systematic review. Indoor Air. 2007; **17**(1): 2-18.

135. Wargocki P, Sundell J, Bischof W, Brundrett G, Fanger PO, Gyntelberg F, et al. Ventilation and health in non-industrial indoor environments: report from a European multidisciplinary scientific consensus meeting (EUROVEN). Indoor Air. 2002; **12**(2): 113-28.

136. Seppanen OA, Fisk WJ. Summary of human responses to ventilation. Indoor Air. 2004; **14 Suppl 7**: 102-18.

137. Norbäck D, Wålinder R, Wieslander G, Smedje G, Erwall C, Venge P. Indoor air pollutants in schools: nasal patency and biomarkers in nasal lavage. Allergy. 2000; **55**(2): 163-70.

138. Nordstrom K, Norbäck D, Akselsson R. Subjective Indoor Air Quality in Hospitals - The Influence of Building Age, Ventilation Flow, and Personal Factors. Indoor and Built Environment. 1995; **4**(1): 37-44.

139. Skov P, Valbjørn O, Disg. The “sick” building syndrome in the office environment: The Danish town hall study. Environ Int. 1987; **13**(4–5): 339-49.

140. Sterling E, Sterling T. The impact of different ventilation levels and fluorescent lighting types on building illness: an experimental study. Canadian journal of public health Revue canadienne de sante publique. 1983; **74**(6): 385-92.

141. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics. 2001; **2**(4): 463-71.

142. Craig R, Mindell J. Health Survey for England - 2010 Respiratory Health. London: NHS Information Centre; 2011.

143. Taylor MP, Morgan L. Ventilation and good indoor air quality in low energy homes. 2011 [cited; Available from: http://www.goodhomes.org.uk/downloads/news/VIAQ%20final%20120220%20-%20PUBLICATION.pdf

144. Zimmermann H. Fuzzy set theory--and its applications. Third ed: Kluwer Academic Publishers; 1995.

145. Maravas A, Pantouvakis JP, Lambropoulos S. Modeling Uncertainty during Cost Benefit Analysis of transportation projects with the aid of Fuzzy Set Theory. Procd Soc Behv. 2012; **48**: 3661-70.

146. Scovronick N, Armstrong B. The impact of housing type on temperature-related mortality in South Africa, 1996–2015. Environmental Research. 2012; **113**(0): 46-51.

147. Rothman K, Greenland S, Lash L. Modern Epidemiology. Philadelphia: Lippincott Williams & Wilkins; 2008.

148. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. American Journal of Public Health. 1998; **88**(1): 15-9.

149. Hänninen O, Knol A. European Perspectives on Environmental Burden of Disease: Estimates for Nine Stressors in Six European Countries. Helsinki, Finland National Institute for Health and Welfare (THL) 2011.

150. Stenberg B, Eriksson N, Hoog J, Sundell J, Wall S. The Sick Building Syndrome (SBS) in office workers. A case-referent study of personal, psychosocial and building-related risk indicators. Int J Epidemiol. 1994; **23**(6): 1190-7.

151. Jaakkola JJ, Miettinen P. Ventilation rate in office buildings and sick building syndrome. Occupational and environmental medicine. 1995; **52**(11): 709-14.

152. Walinder R, Norback D, Wieslander G, Smedje G, Erwall C, Venge P. Nasal patency and biomarkers in nasal lavage--the significance of air exchange rate and type of ventilation in schools. International archives of occupational and environmental health. 1998; **71**(7): 479-86.

153. Oie L, Nafstad P, Botten G, Magnus P, Jaakkola JK. Ventilation in homes and bronchial obstruction in young children. Epidemiology. 1999; **10**(3): 294-9.

154. Milton DK, Glencross PM, Walters MD. Risk of sick leave associated with outdoor air supply rate, humidification, and occupant complaints. Indoor Air. 2000; **10**(4): 212-21.

155. Emenius G, Svartengren M, Korsgaard J, Nordvall L, Pershagen G, Wickman M. Building characteristics, indoor air quality and recurrent wheezing in very young children (BAMSE). Indoor Air. 2004; **14**(1): 34-42.

156. Hagerhed-Engman L, Sigsgaard T, Samuelson I, Sundell J, Janson S, Bornehag CG. Low home ventilation rate in combination with moldy odor from the building structure increase the risk for allergic symptoms in children. Indoor Air. 2009; **19**(3): 184-92.

157. Sun Y, Zhang Y, Bao L, Fan Z, Sundell J. Ventilation and dampness in dorms and their associations with allergy among college students in China: a case-control study. Indoor Air. 2011; **21**(4): 277-83.

158. ONS. Interim 2011-based subnational population projections for England. 2012 [cited 2013 07/04/2013]; Available from: http://www.ons.gov.uk/ons/dcp171778_279964.pdf

159. Tsai DH, Lin JS, Chan CC. Office workers' sick building syndrome and indoor carbon dioxide concentrations. J Occup Environ Hyg. 2012; **9**(5): 345-51.

160. Jaakkola MS, Yang L, Ieromnimon A, Jaakkola JJ. Office work exposures [corrected] and respiratory and sick building syndrome symptoms. Occup Environ Med. 2007; **64**(3): 178-84.

161. Trust ES. Energy efficient ventilation in dwellings - a guide for specifiers. 2006 [cited 2011 24/05/2011]; Available from: http://www.energysavingtrust.org.uk/business/Global-Data/Publications/Energy-efficient-ventilation-in-dwellings-a-guide-for-specifiers-CE124-GPG268

162. Fuentes M. Statistical issues in health impact assessment at the state and local levels. Air quality, atmosphere, & health. 2009; **2**(1): 47-55.

163. Higgins JP. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol. 2008; **37**(5): 1158-60.

164. Dimitroulopoulou C, Crump D, Coward S, Brown B, Squire R, Mann H. Ventilation, air tightness and indoor air quality in new homes: (BR 477) (Building Research Establishment Report): IHS - BRE Press; 2005.

165. Beaula T, Vijaya V. A Study on Exponential Fuzzy Numbers Using alpha cuts. International Journal of Applied Operational Research. 2013; **3**(2): 1-13.

166. Dutta P, Boruah H, Ali T. Fuzzy Arithmetic with or without using the alpha cut method: A comparative study. International Journal of Latest Trends in Computing. 2011; **2**(1).

167. Gao S, Zhang Z. Multiplication Operation on Fuzzy Numbers. Journal of Software. 2009; **4**(4).

168. Serguieva A, Hunter J. Fuzzy interval methods in investment risk appraisal. Fuzzy Sets and Systems. 2004; **142**(3): 443-66.

169. Fink R, Medved S. Health impact assessment of liquid biofuel production. Int J Environ Health Res. 2013; **23**(1): 66-75.

170. Bronnum-Hansen H. Quantitative health impact assessment modelling. Scand J Public Health. 2009; **37**(5): 447-9.

171. McCarthy M, Biddulph JP, Utley M, Ferguson J, Gallivan S. A health impact assessment model for environmental changes attributable to development projects. Journal of epidemiology and community health. 2002; **56**(8): 611-6.

172. Fantke P, Friedrich R, Jolliet O. Health impact and damage cost assessment of pesticides in Europe. Environment International. 2012; **49**: 9-17.

173. Knol AB, Briggs DJ, Lebret E. Assessment of complex environmental health problems: Framing the structures and structuring the frameworks. Science of The Total Environment. 2010; **408**(14, Sp. Iss. SI): 2785-94.

174. Linkov I, Loney D, Cormier S, Satterstrom FK, Bridges T. Weight-of-evidence evaluation in environmental assessment: review of qualitative and quantitative approaches. Science of The Total Environment. 2009; **407**(19): 5199-205.

175. Mesa-Frias M, Chalabi Z, Vanni T, Foss AM. Uncertainty in environmental health impact assessment: quantitative methods and perspectives. Int J Environ Health Res. 2013; **23**(1): 16-30.

176. Sterman JD. Learning from evidence in a complex world. Am J Public Health. 2006; **96**(3): 505-14.

177. Diez Roux AV. Integrating social and biologic factors in health research: a systems view. Ann Epidemiol. 2007; **17**(7): 569-74.

178. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. American Journal of Epidemiology. 2008; **168**(1): 1-8.

179. O'Connell E, Hurley F. A review of the strengths and weaknesses of quantitative methods used in health impact assessment. Public health. 2009; **123**(4): 306-10.

180. Adam B, Molnar A, Gulis G, Adany R. Integrating a quantitative risk appraisal in a health impact assessment: analysis of the novel smoke-free policy in Hungary. European Jounal of Public Health. 2013; **23**(2): 211-7.

181. Sedki K, Bonneau de Beaufort L. Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning. 24th International Conference on Tools with Artificial Intelligence; 2012; Greece; 2012. p. 1035-40.

182. Bertsimas D, Tsitsiklis JN. Introduction to Linear Optimization: Athena Scientific 1997.

183. Svensson E, Berntsson T, Strömberg A-B, Patriksson M. An optimization methodology for identifying robust process integration investments under uncertainty. Energy Policy. 2009; **37**(2): 680-5.

184. Li YP, Huang GH, Nie SL. Optimization of regional economic and environmental systems under fuzzy and random uncertainties. Journal of Environmental Management. 2011; **92**(8): 2010-20.

185. Joffe M. The need for strategic health assessment. European Journal of Public Health. 2008; **18**(5): 439-40.

186. Babonneau F, Vial JP, Apparigliato R. Robust Optimization for Environmental and Energy Planning. 2010. p. 79-126.

187. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing Structural Uncertainty in Decision Analytic Models: A Review and Application of Methods. Value in Health. 2009; **12**(5): 739-49.

188. Xu Y, Hubal EAC, Little JC. Predicting Residential Exposure to Phthalate Plasticizer Emitted from Vinyl Flooring: Sensitivity, Uncertainty, and Implications for Biomonitoring. Environ Health Perspect. 2009; **118**(2).

189. Hickey GL, Craig PS, Hart A. On the application of loss functions in determining assessment factors for ecological risk. Ecotoxicology and Environmental Safety. 2009; **72**(2): 293-300.

190. Greenland S. Sensitivity analysis, Monte Carlo risk analysis, and Bayesian uncertainty assessment. Risk Analysis. 2001; **21**(4): 579-83.

191. Gustafson P, McCandless LC. Probabilistic approaches to better quantifying the results of epidemiologic studies. International journal of environmental research and public health. 2010; **7**(4): 1520-39.