Implementation of Qualitative Uncertainty Guidance: A Worked Example

Ruth Salway, Gavin Shaddick

University of Bath

Introduction

The document 'Guidance on Qualitative Uncertainty Assessment' proposed a three-step procedure for integration qualitative uncertainty assessment within the environmental health impact assessment process of the Heimtsa and Intarese projects. The purpose of this document is to illustrate how this guidance may be implemented in practice, by applying it to a worked example. The emphasis here is on the practical challenges.

The Guidance document identifies three distinct steps:

1. Identification of uncertainty sources;
2. Qualitative characterisation of uncertainty in terms of:
    - direction and magnitude of uncertainty on the results
    - knowledge about the uncertainty source
3. Reporting of qualitative uncertainty

This document is organised as follows:

This document makes a key distinction between performing the uncertainty analysis itself ,and reporting the results. The former should be undertaken in as much detail as possible, and will concentrate on identifying and assessing all sources uncertainty, however small. The reporting of the uncertainty analysis requires a non-technical reader-friendly summary of the results (rather than the process) and will highlight important areas of high uncertainty.

Table of Contents

Example: Health Impact Assessment of NO2 in Bath

To illustrate how to apply the guidance on qualitative uncertainty, this document will be based around a simple worked example. The aim is to assess the health impact of high levels of NO2 in Bath city centre, in terms of estimating the total number of deaths attributable to annual NO2 concentrations higher than the EU objective of 40ugm^{-3} .

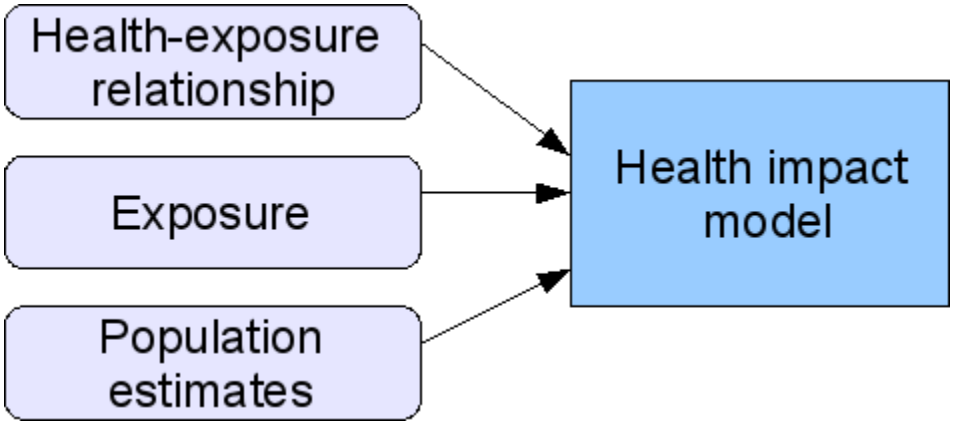

Figure 1: Components of the health impact assessment

Figure 1 illustrates the full analysis, comprising the following components:

- **Health-exposure relationship**

  An estimate of the effect of NO2 on all-cause mortality is taken from the Harvard Six Cities Study (Reanalysis). The relative risk for all-cause mortality (excluding accidents) in over 25s was 1.08 (95% CI: 1.02-1.13) for an increase of 10ugm^{-3} in NO2.

- **Exposure assessment**

  The annual mean NO2 exposure recorded at a single monitoring site (from the UK Air Quality Network data) was 65ugm^{-3} averaged over the years 2005-7. This represents a difference of 25ugm^{-3} above the EU objective of 40ugm^{-3} .

- **Population estimates**

  Population estimates for a specific region in the city centre of Bath are based on the the electoral roll (official list of registered voters); 4283 residents were registered to vote in 2009. The local authority all-cause mortality rate (excluding accidents and infant mortality) for 2005-7 is a rate of 8.9 deaths per 1000.

- **Health Impact model**

  Feeding these figures into a simple formula estimates a total of 5.88 deaths per year in the study region attributable to NO2 levels above the EU objective of 40ugm^{-3} .

The aim is to identify possible sources of uncertainty in this analysis, and assess the effect that their presence could have on the final estimate of 6 pollution-attributable deaths per year.

Step 1: Identifying Sources of Uncertainty

The aim of this step is to identify and list all possible sources of uncertainty, regardless of their perceived importance. The guidance document lists three main sources of uncertainty: scenario, model and parameter uncertainty, and describes several approaches to help generate a list of uncertainties. The important point here is that these are tools to help generate a list of sources of uncertainty; the emphasis should be on listing sources of uncertainty, and not on classifying them. So if a source has already been identified, at this stage it is not that important which category it belongs in (and indeed it may not always be clear which one). If you repeatedly find yourself struggling to assign uncertainty sources that you have already identified into the classification structure, you may find it easier to introduce a category of ***Other*** .

Below is an (incomplete) list of sources of uncertainty for the example. I have chosen to identify uncertainties within each component as this seemed natural to me. As a result, I have included the additional section ***Full Chain,*** which includes sources of uncertainty relating to the analysis as a whole, rather than within any specific component. To facilitate applying the uncertainty analysis across the full chain, I included the components Health-Exposure relationship, Exposure assessment and Population estimates as parameter sources of uncertainty in the health impact model component; uncertainty in these components will feed through.

Since this is only an example, this list does not include every source of uncertainty; sources are given in detail for the health-exposure relationship and more generally for the other components.

Health-exposure relationship	
Scenario	Concentrations measured at monitoring sites are assumed to be equivalent to exposure Individuals are assumed to remain at their postcode of residence there is no allowance for people moving in and out of the study areas over the course of the study Background characteristics such as behaviour may change over the course of the study individuals with different characteristics are aggregated together to produce a common relative risk Multiple monitoring sites are aggregated in some way (typically averaged) It includes over 25s only
Model	The relationship between health and exposure is linear (on the log-mortality scale) and there is no threshold value below which there is no health effect. The relationship between health and exposure is causal; that is, it is directly NO2 concentrations that are responsible for deaths and not some other factor, such as other pollutants Statistical uncertainty in the form of a confidence interval for the RR Extrapolation over time – the study was carried out over 16 years between 1975-1991; the HIA is based on 2005-7 data. Extrapolation over space – the study was carried out in the US, and the HIA is in the UK.
Parameter	There is potential measurement error in the following inputs to the study: ● NO2 concentrations ● Population estimates ● mortality rates
Other	validity of study
Exposure assessment	
Scenario	A single NO2 measurement (near one of the busiest roads in Bath) is taken for the whole of the city centre The mean is taken across 3 years Measurement error in NO2 at the monitor

Population estimates	
	Population at risk estimated from the electoral roll which will exclude children and may under-represent certain subpopulations eg transient population such as students Population was counted manually; subject to error Streets not completely within the study region were excluded population estimates were taken in 2009, compared to 2005-7 for the NO2 and mortality rates exposed population excludes those who work in the area but do not live there (and conversely includes those who live but don't work there) mortality rates are for the whole of the local authority including suburbs and rural areas, not just the city centre study region mortality rates exclude infant mortality but may include other deaths in under 25s.
Health impact model	
Scenario	Assumes that concentrations at monitoring sites are equivalent to exposures
Model	Assumes the RR applies uniformly across all subgroups (eg age, sex, behavioural characteristics) Assumes a causal link between NO2 and health
Full Chain	
Scenario	Suitability of long-term versus short-term effects for HIA An estimate of 6 deaths per year in the future assumes all factors remain constant eg nothing is done to try and reduce high levels of NO2. Suitability of all-cause mortality as an endpoint to capture 'health effects' Suitability of number of deaths attributable to high NO2 as a summary of adverse health effects.

Summary:

- Compile a comprehensive list of all possible sources of uncertainty in the study.
- Use classification categories to help produce a list, but don't get tied up in classifying rather than listing
- Introduce other categories as you feel necessary
- Include components of the full chain to enable uncertainty to be fed through

Step 2: Qualitative characterisation of uncertainty

This step falls into two parts: assessing the magnitude of the effect of each source of uncertainty, and determining the associated uncertainty in the knowledge base (scientific evidence). To ensure that these are assessed consistently we include an additional first step: developing the criteria against which the sources of uncertainty are to be assessed.

Step 2a: Define categories of low, medium and high

Before the sources of uncertainty can be assessed, we must define what we mean by low, medium and high. The guidance document states:

The magnitude of uncertainty is rated low when it is judged that large changes within the source of uncertainty would have only a small effect on the assessment results and when the values of the data sets needed for the assessment are known. A designation of medium implies that a change within the source of uncertainty is likely to have a moderate effect on the results and the values of the data sets needed for the assessment are unknown (completely or partially). A characterisation of high implies that a small change in the source would have a large effect on results and the values of the data sets needed for the assessment are unknown.

The criteria of high, medium and low are context-specific and will therefore depend on the specific case study; criteria developed for one study cannot be applied to a different study. However, within one study (and so within one uncertainty analysis) the same criteria will be applied to all sources of uncertainty, regardless of the different categories or components. This will ensure consistency across the whole study.

To compile such a list of criteria, we propose the following approach:

- Identify clearly the endpoint.

  This endpoint should be the endpoint of the full chain analysis rather than any intermediate point. In this example the endpoint is the estimated number of pollution-attributable deaths. The example concluded that the number of pollution-attributable deaths was 6 per year; we are concerned with how uncertainty could have affected this estimate. For example, it could be that the true answer is 4 deaths per year, but various sources of uncertainty mean that we have estimated it at 6, resulting in a bias of 2.

- Produce a list of all possible effects that uncertainty might have on the endpoint.

  Possible effects include: the answer being correct, being underestimated by 5, being overestimated by a factor 2 etc. The actual effects used (integers, percentages, monetary values etc) , and the way in which they are expressed (eg as an absolute number, or scale factor) will depend on the application. Effects need not be mutually exclusive.

- Assign each of these possible effects to a classification of high, medium and low.

  Where effects overlap, ensure consistency of classification.

A list of possible effects for the example appears in Table 1. Here effects are expressed in integers as the endpoint is number of deaths, and we concentrate on bias caused by uncertainty; that is, the difference between the number of deaths estimated in the study (around 6) and what the true number of deaths might be if we didn't have uncertainty. The case where the truth is that there are no pollution-related deaths but we have estimated 6 deaths is singled out as a special case, although it's also included in the bias within 5-10 category; note that the magnitude category is consistent here.

Effect of uncertainty on the endpoint	**Magnitude Category**
number of deaths is correct to nearest integer	low
bias is within 1-2 deaths	low
Bias is within 3-5 deaths	medium
bias is within 5-10 deaths	high
Bias is greater than 10 deaths	high
True number of deaths is 0 deaths	high

Table 1: Magnitude scale for the effect of uncertainty

The benefits of this table are several .Firstly, by constructing the table independently of the individual list of sources of uncertainty we fix the scale in advance and are less tempted to refine as we go along. Secondly, it moves the emphasis from assessing a specific source of uncertainty as low, medium or high onto determining what the effect on uncertainty could be. Finally, if the table is constructed jointly, it may be applied independently across all components of a study without losing consistency.

This table is used as a look-up table. So for example, consider uncertainty that arises from using monitoring site

concentrations rather than real exposures. Instead of trying to determine whether this source of uncertainty is low, medium or high, we instead concentrate on the effects. Suppose we have evidence to believe that uncertainty from this source could result in underestimating the number of deaths by a factor 2. We simply look up this effect in the table (bias is 10 deaths) and read off the magnitude category as HIGH. This methods helps avoid subdividing categories such as 'lowish, but not as low as the previous one', and also ensures that one isn't constantly returning to earlier sources of uncertainty to re-allocate the categories because of subsequent refinement of the scale.

The same approach can be applied to construct a look-up table for categorising uncertainty related to the knowledge base. Table 2 gives an example; again possible states of the knowledge base are listed and then assigned to a category. It concentrates primarily on support from scientific literature, but includes a couple of other sources of knowledge which it might be useful to consider in practice.

State of the knowledge base	Uncertainty Category
consistent extensive scientific evidence of many different types from many different sources	low
consistent extensive scientific evidence of a single type or from a single source	low
consistent scientific evidence but all suffering from the same limitation(s)	medium
scientific evidence is mixed; the bulk of it supports the conclusion	medium
consistent scientific evidence for related population/scenario but limited for this situation	medium
scientific evidence is limited or inconsistent	high
no external scientific evidence exists; internal analysis to support	high
no external scientific evidence exists; expert opinion is generally consistent	high
nothing is known	high

Table 2: Uncertainty scale for assessing the knowledge base

Step 2b: Magnitude and Direction: how big is the problem?

Once we have a list of all possible sources of uncertainty, and suitable definitions of low/medium/high categories, we apply the uncertainty analysis to each source of uncertainty in turn:

- What do we know about uncertainty from this source?
- Is it likely to result in over or underestimation in the endpoint?
- What is the size of the possible effect?
- Look up this effect in the look-up table and assign a magnitude category to this source of uncertainty.

Determining the magnitude of the effect is more important that determining whether it is over or underestimating; in practice many sources of uncertainty could result in bias in either direction.At this stage, the emphasis is still on assessing uncertainty, not on summarising results, so it is expected that this stage will yield a relatively high amount of text; I anticipate somewhere between one paragraph to one page, per source depending on how important the source of uncertainty is. For this reason, it is not always helpful to force this stage into a matrix grid, which has limited space for the detail necessary.

Detailed example: The relationship between health and exposure is linear (on the log-mortality scale) and there is no threshold value below which there is no health effect.

If the relationship is not linear then depending on what the true relationship is we could either under or over estimate the health-exposure relationship resulting in either an over or under-estimate of the number of deaths. The extent of the bias will depend on what the true relationship is; if the effect on health is likely to be much higher than estimated for high levels of NO2 (which is more likely than for the effect to be lower in this context) this suggest a convex relationship and the assumption of linearity will underestimate the effect for high levels of NO2, and overestimate for low levels. The NO2 levels considered here are relatively high, and so we are more likely to underestimate the health effect resulting an under-estimate of the number of deaths

In this case, the relative risk from the study may be considered an 'average' risk applied to an 'average' population, and only large deviations from linearity will result in large bias in the number of deaths. It is unlikely to affect the final number of deaths by more than 1 or 2 (category: LOW)

A linear-threshold is a particular type of non-linear relationship, where for NO2 exposures greater than the threshold there is a linear relationship between (log) health and exposure, but below that threshold there is no effect. If there is a threshold value then assuming there is not will mean that we will overestimate the relationship for NO2 values below the threshold, and underestimate for those above. In this example, provided that the threshold is below 40ugm^{-3} we will underestimate the health effect resulting an under-estimate of the number of deaths. The magnitude of the bias will depend on what the threshold value is; the lower the threshold, the smaller the bias.

If we assume that a threshold value visible in the Six Cities study would have been identified, we can conclude that any threshold must be less than the observed minimum NO2 of 12ugm-3. A threshold at this level or below would have negligible effect on the results of this HIA, underestimating the number of deaths by around 1 at most (category: LOW)

Step 2c: Knowledge base: how likely is it we have a problem?

Assessing the knowledge base concerns the extent of the problem in this study, and how likely it is that we have uncertainty from this source.

- What does the scientific literature tell us about how likely it is that we have uncertainty from this source?
- What is the evidence – how much? Is it relevant? Is it consistent?
- Look up the state of evidence in the look-up table and assign an uncertainty category to the knowledge base concerning this source of uncertainty.

Detailed example: The relationship between health and exposure is linear (on the log-mortality scale) and there is no threshold value below which there is no health effect.

Most epidemiological studies assume a linear relationship; obvious deviations from this are likely to have been noted. Most evidence supports a nonlinear relationship. (category: LOW)

Numerous epidemiological studies (including indoor, outdoor, cohort, cross-sectional) and toxicology studies have found no evidence to support a threshold effect. (category: LOW)

Summary:

2a: define the categories

- Identify the endpoint: in a full chain analysis, this is the final endpoint of the full chain analysis, not any intermediate point.
- List all possible effects uncertainty could have on the final study estimate of this endpoint
- Classify each as high, medium or low to produce a look-up table
- Repeat for uncertainty about the the knowledge base

2b: direction and magnitude

- For each source of uncertainty in turn, summarise the effect uncertainty from this source is likely to have on the final endpoint.
    - Will it result in overestimation of underestimation? (In some cases it may not be possible to say)
    - What is the effect of this uncertainty?
    - Look up this effect in the look-up table and identify the category.

2c: knowledge base

- summarise the scientific (or other) evidence available
- look up the associated uncertainty category in the look-up table.

Step 3: Reporting Qualitative uncertainty

It is suggested that reporting the uncertainty analysis should be done for each uncertainty category (scenario, model and parameters) within each component

Gavin: not really sure what to do here. As we've applied the uncertainty analysis at the individual source level, to report at uncertainty category level for each component requires combining uncertainties to produce an aggregated uncertainty over several sources, in exactly the way that Denis' document doesn't tell us how to do. I could (i) ignore the guidance and simply report the detailed example from the previous section – table 5, (ii) fudge it and present an example at the right level but without any details of exactly how I did it – table 6 (plus I'll have to actually do it at some level) or (iii) get into the whole how-to-aggregate stuff myself. My temptation is to go for the first option , although that ends up deviating away from the guidance document.... There's also the whole question of Denis' nasty 'qualitative assessment matrix' which I purposefully ahven't used in table 5 but did in table 6...

The uncertainty scales also need to be reported; these can be summarised versions of the look-up tables in Section 2 Note that the earlier more detailed versions were for use as a look-up table to enable the assignment of sources to categories. Here , at the reporting stage, they are simply descriptive to enable someone to understand the uncertainty summary.

Magnitude Category	**Criteria**
Low	bias is within 0-2
Medium	bias is within 3-5
High	bias is greater than 5

Table 3: Uncertainty scale for magnitude

Uncertainty Category	**Criteria**
Low	consistent extensive scientific evidence
Medium	Less extensive or mixed evidence
High	Limited, inconsistent or no evidence

Table 4: Uncertainty scale relating to the knowledge base

Source	Direction	Magnitude	Knowledge base uncertainty
Assumption of linear relationship	either;underestimate more likely	low	low
	Notes: In the event of nonlinearity the relative risk may be considered an 'average' risk; nonlinearity is unlikely to affect the final number of deaths by more than 1 or 2. SUMMARY OF EVIDENCE		
Assumption of no threshold	underestimate	low	low
	Notes: Any threshold is likely to be less than the minimum in the study which would have negligible effect on the number of deaths, underestimating the number of deaths by around 1 at most. There is no support in extensive literature for existence of a threshold value.		

Table 5: Uncertainty analysis: my way

		Dimensions of uncertainty			
Sources of uncertainty		**Direction**	**Level**	**Appraisal of knowledge base**	**Justification**
Health-Exposure Relationship					
Scenario					
Model	Model Assumptions	U	L	L	Extensive evidence supports model assumptions, and any effect is likely to be small.
	Extrapolation		M	L	
Parameters	Measurement error		L	L	Measurement error likely to exist, but only on a very small scale; negligible compared to other errors!

Table 6: Uncertainty analysis: Denis' way