Working Paper for COMEAP Report "Associations of long-term average concentrations of nitrogen dioxide with mortality" published July 2018

Working Paper 2: A viewpoint on using adjusted coefficients for NO_2 and $PM_{2.5}$

J Fintan Hurley

FOREWORD

Background to this Working Paper

By late 2016 it was clear that there were big differences of views within the Committee on the key issue of whether it was both feasible and useful to provide quantitative answers to the questions that the Working Group was tackling. One specific area of difference was whether it was useful, or misleading, to use NO_2 coefficients adjusted for PM; and whether corresponding PM coefficients adjusted for NO_2 could be derived and used also. While these issues were relevant to assessing the impact of policies, discussion within the Committee focused especially on whether or not it was feasible and useful to quantify the mortality burden attributable to the air pollution mixture as a whole. (There was little interest within the Committee as a whole in quantifying the mortality burden attributable to NO_2 itself – that quantification was acknowledged as being both more difficult and less useful than quantifying the burden of the overall pollution mixture.)

By late 2016 the Committee Members who most strongly thought that attempts to quantify mortality burden were not justified had articulated their point of view in a working paper. Those of us who favoured quantification and were working out how it might be done were preoccupied with that effort, to the extent that we had not yet expressed in writing our reasons for supporting quantification. The Working Group Chairman, Professor Roy Harrison, asked me to prepare a working paper on adjusted coefficients and how they might be used, partly to put a point of view on record, partly to inform further discussions of the Working Group and Committee.

The Working Paper drafted in late 2016 and early 2017

The Working Paper that follows is my response to that request. I wrote it in late 2016 and revised it in the first half of 2017. I wrote it as a personal authored piece, because that was easiest in a busy period of Committee and other work. Among the majority of the Committee who favoured quantification, there was and is a range of views on some specific aspects of what can be quantified, and in what way, and how well. Rather than attempt a consensus document, I drafted a 'statement of opinion' about what I considered to be feasible, and with what degree of (un)certainty. Whilst the resultant Working Paper attempts to describe as accurately as I could both my own views and the views of others, the Working Paper is not, nor was it intended to be, a consensus document among those of us who favoured quantification. Of course it draws on the thoughts and opinions of other COMEAP colleagues, in particular Dr. Heather Walton, and for this I am very grateful; but the opinions here are mine, and for these, and for any errors, I take full responsibility.

Subsequent use of the Working Paper

In the latter half of 2017, substantial parts of the present Working Paper were extracted, edited and integrated into the Main Report, as part of describing 'the majority view'. Various drafts of the Main Report, including those aspects integrated from the Working Paper, went through several rounds of comment, discussion and revision, before being finally agreed. The Committee decided that the Working Paper was nevertheless sufficiently useful that it be published as one of several Working Papers supplementing the Main Report.

I then had the choice of whether or not to revise the Working Paper once more, to make it consistent with the Main Report final text. With the agreement of the Working Group, I decided not to undertake a full revision. This was partly for pragmatic reasons: a full revision would have involved substantial work and so would have risked further delay to publication. But also, there is a case for thinking that the original paper, written to inform debate and discussion, is more interesting and useful in its original form, recording views of its time ("warts and all"), than an edited and 'sanitised' version would be. Updated details of coefficients from one study (which subsequently became available to 3 decimal places) have been added, and some minor edits made to make the paper consistent with the final version of the report (e.g. cross-referencing to report chapters and other working papers). However, the text discussing the interpretation and use of the evidence has not been edited.

Status relative to the Main Report

It follows that there are differences between the majority view as described in the Main Report and the earlier, personal, viewpoints described here. It may not need to be said, but in case of difference the Main Report takes precedence. Also, and for the avoidance of doubt, I am happy to say that as a COMEAP Member I am fully signed up to the Main Report.

If nevertheless there are ideas in the Working Paper, and not in the Main Report, which are of interest, you are welcome to contact me fintan.hurley@iom-world.org. Meantime, enjoy or ignore this Working Paper, but always, please, remember its context, purpose and subsidiarity.

Fintan Hurley, Edinburgh, Scotland

February 2018.

A viewpoint on using adjusted coefficients for NO_2 and $PM_{2.5}$

Fintan Hurley

1. BACKGROUND AND CONTEXT

1.1 Mortality and long-term exposure to NO_2

As shown in this report, and elsewhere (e.g. Hoek et al, 2013), there are now numerous cohort studies showing relationships between annual average concentrations of NO_2 and mortality hazards in adults. (Mortality hazards are age-specific risks of dying 'now', among people who have survived to that age. They vary with gender and other characteristics of the population.) These relationships are expressed as concentration-response functions (CRFs) covering the range of annual average NO_2 concentrations studied and are summarised typically as coefficients, with associated confidence intervals (CIs).

In principle these CRFs can be used to help quantify the effect of air pollution on mortality in adults expressed as an effect on, for example, the total time lived (or total years of life lost) by the population, or equivalent number of deaths attributable to air pollution in any one year, or life expectancy at birth. Such quantification is already well established for $PM_{2.5}$ using relationships from cohort studies linking mortality hazards and annual average $PM_{2.5}$ (see e.g. COMEAP 2009, 2010). There is not a corresponding body of experience in using CRFs for mortality and long-term exposure to NO_2.

In its detailed REVIHAAP evidence review (WHO 2013a) and associated HRAPIE proposals for quantification (WHO 2013b), an international working group recommended a coefficient (with associated CI) linking mortality hazards and annual average NO_2, for quantification to inform policy development at EU level. However, WHO (2013a, b) recognised difficulties respectively with interpretation of causality of NO_2 and with the use of the coefficient (it classified it as one with greater rather than lesser associated uncertainties), and indeed an effect of long-term exposure to NO_2 was not included in the main resultant cost-benefit analyses of EU air quality policies (Holland 2014).

1.2 The main report and the role of particulate matter (PM)

The present main report is concerned with how, and indeed whether, given the difficulties, these CRFs / coefficients linking NO_2 and mortality can help provide useful answers to meaningful questions about the effects on mortality of air pollution in general and NO_2 in particular. In discussions within both the main Committee (i.e. COMEAP as a whole) and the NO_2 working group, it soon became clear that in order to understand whether and how quantification in NO_2 might be useful, we needed also to consider the role of other pollutants, especially particulate matter (PM), and especially the mass concentration $PM_{2.5}$ of fine particles.

There are three main reasons for this. All are identified and discussed, sometimes in detail, in the main report.

The first concerns causality. As discussed in the main report and in other Appendices, there are different views on the extent to which CRFs linking mortality and long-term exposure to NO_2 represent a casual effect of NO_2 itself. There is agreement however that, in addition to representing any effect of NO_2 itself, these CRFs reflect also some effect of multiple other pollutants that co-vary with annual average NO_2. One of those co-varying pollutants is PM, which in the context of long-term exposure and mortality is usually represented as annual average $PM_{2.5}$. This raises questions such as: How does correlation with other pollutants, and especially with PM, affect use and interpretation of coefficients in NO_2? In particular, how does it affect attempts to understand and quantify an effect of NO_2 itself?

The second reason is that, as noted, there is now a well-established practice, in the UK and internationally, of using CRFs in annual average $PM_{2.5}$ to quantify the effects of the mixture as a whole, and of particulate matter specifically. This raises questions such as: What implications does quantification using CRFs in $PM_{2.5}$ have for interpretation and use of quantification results using annual average NO_2 coefficients? And indeed, conversely, what implications would quantification in NO_2 have for interpretation and use of quantification coefficients in $PM_{2.5}$? For example, how meaningful is it to add results from both quantifications in order to estimate the effect of a mixture that includes both pollutants, and others?

The third reason is that there is only a relatively small number of cohort studies which report results from two-pollutant models, i.e. models where relationships between mortality and both NO_2 and PM are estimated simultaneously, and so give coefficients for each pollutant adjusted for the other, as well as for many other factors. The main report identifies six cohorts with such 2-pollutant models and summarises their results. Adjusted coefficients are also available from a 7th cohort study. These are based on a 3-pollutant model (i.e. with simultaneous inclusion of and adjustment for ozone as well as NO_2 and $PM_{2.5}$); these will be considered later.

1.3 Differences of views about using adjusted coefficients

These 2-pollutant cohorts and associated adjusted coefficients are potentially an important resource for disentangling the effects associated with NO_2 from those associated with PM. However, as noted both in the main report and in another Working Paper, there are difficulties associated with these adjusted coefficients leading to different views within the Committee about their practical usefulness at this time.

Some Members are of the view that, at present, the adjusted coefficients should not be used for quantification; and Chapter 10 and associated annexes give their reasons for this. Other Members, of whom I am one, think that on balance there is more to be gained by using adjusted coefficients than by not doing so, provided that the limitations as well as the strengths of doing so are described clearly.

1.4 Broad aim and scope of this Working Paper

Against this background, the Chair of the NO_2 Working Group asked me to summarise my views on the use of adjusted coefficients. This Working Paper is my response. I see it as an attempt to answer a generic question:

"To what extent do these two-pollutant models and their adjusted coefficients enable us to provide useful answers to meaningful questions about the effects on mortality of air pollution in general and NO_2 in particular, given (i) the difficulties about causality and (ii) the availability and widespread use of quantification in $PM_{2.5}$, noted earlier?"

1.5 What meaningful questions are to be addressed?

It makes sense to recap briefly the general nature of these "meaningful questions". Drawing on the overall content of the main report (rather than the specific questions identified within it at the start of the report) these are in two main dimensions:

One is familiar from COMEAP (2010): Is the focus of interest on estimating the burden of mortality, i.e. the size of the problem, attributable to current concentrations of NO_2 and/or to the air pollution mixture as a whole? Or is it on the impact of changes, and in particular of policy interventions, which affect annual average concentrations of NO_2 specifically and/or a more general mixture including both NO_2 and $PM_{2.5}$ (as well as other pollutants)?

The other is new, because it is new for COMEAP to consider multiple pollutants simultaneously: Is the focus of interest on estimating the effect of NO_2 itself, or on the effect of a pollution mixture that includes NO_2, PM and other pollutants?

Together these two considerations give rise to four kinds of question about mortality in adults as a result of exposure long-term to air pollution, especially NO_2. What is:

i. The impact of an intervention which reduces NO_2 itself but not emissions of other relevant pollutants;

ii. The burden attributable to NO_2 itself;

iii. The impact of an intervention which reduces (emissions of) a mixture of pollutants, including primary $PM_{2.5}$, as well as reducing NO_2 itself;

iv. The burden attributable to the overall mixture containing $PM_{2.5}$, NO_2 and other pollutants.

So the overall aim can be made more specific: What, if any, is the legitimate role of adjusted coefficients in helping answer these four questions?

1.6 Content and specific objectives

The specific objectives of the present Working Paper are:

1. To recap briefly the evidence base for two-pollutant models of mortality in relation to long-term exposure to NO_2 and PM, and the agreed main qualitative conclusions that can be drawn from it;
2. To consider the difficulties involved in using adjusted coefficients from those two-pollutant models, and the reasons nevertheless for using them in assessment of burden or impact in the UK, in the context of NO_2 and PM;
3. To outline how adjusted coefficients could be used for (i) quantification (whether for calculations of impact or of burden) of the effect of a pollution mixture which includes both NO_2 and $PM_{2.5}$; and (ii) to provide an upper limit of the effect of NO_2 itself (again, whether for calculations of impact or of burden);
4. To discuss the methodology behind any proposed quantification using adjusted coefficients.

2. AREAS OF SIGNIFICANT AGREEMENT AND A MAJOR AGREED QUALITATIVE CONCLUSION

Despite different views in the Committee about using adjusted coefficients for quantification at this time, there has been and is a very significant amount of agreement also. Principally, this has been about the evidence base, the issues it raises, and the qualitative conclusions that can be drawn from it.

2.1 The principle of using adjusted coefficients

The underlying issue is a widespread one in science: If several potential causes are being considered simultaneously as actual or potential explanatory variables for some phenomenon, (i) (how) can we distinguish their separate effects and (ii) how do the several variables interact and/or modify the effect one of another?

It stands to reason that this task is more difficult insofar as some explanatory variables are highly correlated with one another and/or are measured imprecisely.

Statistical methodology offers various kinds of simple and more advanced regression methods to help deal with this issue. These are designed to allow the simultaneous modelling of relationships with multiple coefficients and so to provide coefficients for the association with any one variable that are adjusted for the associations with other variables in the model.

All the coefficients that we use in quantifying the effects of air pollution on mortality, whether from time series studies, from panel studies or from cohort studies, are adjusted to take account of possible confounders. In cohort studies these include both characteristics of the individuals (e.g. age, gender, ethnicity, smoking habit, occupation…) and of the locations (typically cities or parts of cities) where they reside. From this viewpoint, all coefficients are in fact adjusted coefficients; adjusted coefficients are in general ‘a good thing’; there is no problem in principle with using them; and if we didn’t use them epidemiology would have little to tell us about the effects of air pollution at ‘ordinary’ levels.

Within the present report and working papers, we speak of ‘adjusted coefficients’ in the more limited sense of coefficients for NO_2 adjusted for PM, and/or coefficients for PM adjusted for NO_2. As discussed below, there are some particular difficulties with use of adjusted coefficients in this specific context. My understanding however is that, within the Committee as a whole, there was no disagreement about the principle of using adjusted coefficients understood in this more limited way, provided that this can be done with sufficient confidence in the adjusted coefficients and the results they lead to. Indeed, if the adjusted coefficients were considered good enough to be used, for several of us the preferred approach to quantifying the effects of a mixture which included both NO_2 and PM would be to estimate the effect of each pollutant adjusted for the other, and then sum those estimated effects – this seems the most logical approach, if the adjusted coefficients are usable and the underlying data for both NO_2 and PM are available.

2.2 The initial evidence base – six studies identified initially

In the present context, we sought cohort studies which simultaneously modelled the associations with mortality of PM and of NO_2. Logically this is a subset of those studies which contributed to the meta-analysis of single-pollutant NO_2 coefficients. Like the single-pollutant meta-analysis, the work on identifying relevant studies and adjusted coefficients was led and carried out by Richard Atkinson and colleagues at St. Georges, University of London; all on the Committee recognise and appreciate the high quality of that work, whether or not we agreed about how it should be used. Criteria for cohort selection were agreed with the Committee as a whole and resulted in 6 studies being identified (see Table 1, below.)

Table 1: Hazard ratios from single and two pollutant models for NO_2 and $PM_{2.5}$/PM_{10} (HRs are expressed per inter-quartile range (IQR)).

Study	Cohort	Correlation NO_2/$PM_{2.5}$	NO_2 IQR (μg/m^3)	NO_2	NO_2 adjusted for $PM_{2.5}$/ PM_{10}	%[5]	$PM_{2.5}$/ PM_{10} IQR (μg/m^3)	$PM_{2.5}$/ PM_{10}	$PM_{2.5}$/ PM_{10} Adjusted for NO_2	%[5]
Cesaroni et al 2013	Rome	0.79	10.7	1.029 (1.022, 1.036)	1.026 (1.015, 1.037)	10	5.7	1.023 (1.016, 1.031)	1.004 (0.994, 1.015)	82
Carey et al 2013[1]	CPRD	0.85	10.7	1.022 (0.995, 1.049)	1.001 0.959, 1.044)	95	1.9	1.023 (1.000, 1.046)	1.023 (0.989, 1.060)	0
Beelen et al 2014[2]	ESCAPE	0.2-<0.7	10.0	1.015 (0.993, 1.036)	1.007 (0.967, 1.049)	53	5.0	1.070 (1.016, 1.127)	1.060 (0.977, 1.150)	14
Fischer et al 2015[3]	DUELS	0.58[4]	10.0	1.027 (1.023, 1.030)	1.019 (1.015, 1.023)	29	2.4	1.019 (1.016, 1.022)	1.010 (1.007, 1.013)	46
HEI 2000[4]	ACS CPS II	-0.08	81.4	0.95 (0.89, 1.01)	0.90 (0.84, 0.96)	105	24.5	1.15 (1.05, 1.25)	1.22 (1.11, 1.33)	-42
Jerrett et al 2013	ACS CPS II	0.55	7.7	1.031 (1.008, 1.056)	1.025 (0.997, 1.054)	19	5.3	1.032 (1.002, 1.062)	1.015 (0.980, 1.050)	53

Explanatory Notes: 1: $PM_{2.5}$ results –personal communication; 2: Based on 14 cohorts in which correlation between NO_2 and $PM_{2.5}$ was less than 0.7 (figures to 3 decimal places provided by personal communication are reported here; these were not available to us during our initial consideration), HRs expressed per 10 μg/m^3 NO_2 and 5 μg/m^3 $PM_{2.5}$ 3: PM_{10}; 4: HR (95% CI) for min-max range of average concentrations in fine particulate cohort (41 cities); 5: % reduction in HR)

This Table was the basis of a lot of discussion within the NO_2 working group and the Committee as a whole. We noted the following.

- There is very high correlation between (annual average) NO_2 and $PM_{2.5}$ in two of these studies (Cesaroni et al. and the UK cohort of Carey et al.). This limited their usefulness in trying to distinguish the effects associated with the two pollutants, individually but also jointly.

- The NO_2 results of HEI 2000 are strange and counter-intuitive, both before and after adjustment for $PM_{2.5}$. This is a between-city study with very little (in fact a slightly negative) correlation between NO_2 and $PM_{2.5}$. As such it is different from the other studies which used finer scale modelling. Also, its results were superseded by a later paper (Turner et al., 2016) from the same ACS cohort.

- Fischer et al. (2015) has significant limitations and significant strengths. The strengths included cohort size (more than 7 million subjects) and resultant relatively narrow CIs for the estimated coefficients. The main limitations are that there was only limited adjustment for potentially confounding non-pollution factors and that PM was represented as PM_{10} rather than $PM_{2.5}$.

- There was a lot of variation between studies in the extent to which adjusted coefficients for NO_2 and for PM from 2-pollutant models differed from the corresponding unadjusted coefficients from single-pollutant ones. However the more extreme changes occurred in the two high-correlation studies and in HEI 2000.

- The values of annual average NO_2 and $PM_{2.5}$ used in those studies to represent exposure were necessarily to some extent subject to 'error' (a mixture of bias and imprecision) and the effect of this 'error' on adjusted and unadjusted coefficients is unclear. (Some such 'measurement error' is unavoidable in cohort studies.)

These and other aspects were summarised in the Interim Statement of December 2015 and have been discussed in greater detail in the main report.

2.3 Major qualitative conclusion

Nevertheless, and as described in the main report, consideration of the evidence from these six studies led to a major agreed conclusion, which was summarised as follows in Paragraph 4 of the Summary of the Interim Statement of December 2015:

"Further analysis to date has suggested that within the limited number of individual epidemiological studies that examine the effects of long-term exposure to both NO_2 and $PM_{2.5}$, the combined effect of NO_2 and $PM_{2.5}$ estimated using coefficients where each is adjusted for the effects of the other, is either similar to or only a little higher than what would be estimated for either $PM_{2.5}$ or NO_2 alone, using unadjusted single-pollutant coefficients."

Very briefly, the basis of this finding is as follows. Take any of the six studies and compute (i) an effect of unadjusted NO_2; (ii) an effect of unadjusted PM; and (iii) a combined effect of adjusted NO_2 and adjusted PM, where these effects are computed based on the inter-quartile range (IQR) of NO_2 and of PM in that study. Results showed that, across the six studies, the sum of effects from the adjusted coefficients is in general a little, but only a little, higher than the greater of the two estimated unadjusted effects. However, whether the 'greater of the two unadjusted effects' relates to unadjusted NO_2 or unadjusted PM varied by study.

2.4 Some implications of the qualitative conclusion

This qualitative conclusion has several important implications. The Interim Statement itself highlighted one of these: “This suggests that using a single pollutant coefficient for NO_2 and a single-pollutant coefficient for $PM_{2.5}$ and adding the results, would give an overestimate of the combined effects of the two pollutants.” (Interim Statement, Summary, Para 4.)

More generally, it implies that, in the present context of mortality and long-term exposure to NO_2 and $PM_{2.5}$, the effect of the overall mixture is not seriously under-estimated by using only a single pollutant and associated unadjusted coefficient as an indicator of the mixture as a whole, compared with quantifying in both pollutants, using adjusted coefficients, and then adding the results. This provided reassurance that the current methodology (c/f COMEAP, 2010), of using only $PM_{2.5}$ as a basis for quantifying both mortality burden and impact of long-term exposure to air pollution, was not seriously misleading and in particular that, even when associations with NO_2 were taken into account, the estimated “size of the air pollution problem” in the UK was unlikely to be markedly higher than that estimated by COMEAP 2010.

The different results from different studies in Table 1 show however that it is unclear a priori which of the two single pollutants, NO_2 or $PM_{2.5}$, would give results which best approximate that of the mixture as a whole. Since both are likely to be under-estimates, presumably whichever gives higher results has least under-estimation in it. Finding out which one gives the higher results may need to be done empirically in attempts to answer particular questions.

Finally, the qualitative statement also allowed those of us who favoured using adjusted coefficients nevertheless to support recommendations for assessment of impacts that are based on single-pollutant unadjusted coefficients, even for quantifying the effect of a mixture, because the difference in results between the two approaches would not be of any great practical significance. (The advantage of using unadjusted coefficients was that the corresponding results were likely to be more widely accepted, within the Committee and more widely, than results based on adjusted coefficients.)

3. GIVEN THESE RESULTS, CAN AND SHOULD WE USE ADJUSTED COEFFICIENTS TO HELP ANSWER THE QUANTIFICATION QUESTIONS OF INTEREST?

3.1 Difficulties in using the adjusted coefficients

There were and are several barriers to using coefficients for NO_2 adjusted for PM, either alone or together with coefficients for PM adjusted for NO_2. One such difficulty was the high correlation between $PM_{2.5}$ and NO_2 in two of the six studies of Table 1, because it is not possible to estimate reliably the effects of explanatory variables individually when the study location involves highly correlated and/or imprecisely measured explanatory variables.

A related difficulty is the uncertainty in transferring, for use in the UK, the results of cohort studies in particular locations, with specific correlation structures between NO_2, $PM_{2.5}$, other pollutants and other non-pollutant explanatory variables, to other contexts with presumably different pollution mixtures and a different mixture of non-pollution factors also.

These difficulties of transferability are exacerbated when/if quantification involves estimating effects on mortality at pollutant concentrations outside the range of what has been studied in the original underlying cohorts.

There were also difficulties in combining information across studies into some kind of average adjusted coefficient for NO_2.

And finally, there was what seemed an insurmountable difficulty: How would we get a suitable adjusted coefficient for mortality and long-term exposure to $PM_{2.5}$ without reviewing afresh the relevant $PM_{2.5}$ literature, something which had not been envisaged and had not been planned?

3.2 Why then use adjusted coefficients?

There was broad agreement about the nature of these difficulties; different views about their importance and the extent to which they could be overcome led eventually to different views on whether or how the adjusted coefficients could be used in a COMEAP Report. A viewpoint that adjusted coefficients for NO_2 and $PM_{2.5}$ should not be used, at least at this time, is given by Atkinson et al. (Chapter 10).

Others of us remained interested in using the adjusted coefficients for quantification. Our primary and shared interest was methodological – we wanted to see what was involved in implementing an approach based on adjusted coefficients, despite the difficulties summarised above and elsewhere about their use. Health Impact Assessment (HIA) of the effects of air pollution has over the years made progress by quantifying in the face of uncertainty, while being transparent about the limitations as well as the strengths of doing so. This, over time, leads to improved methods, whereas not quantifying, because of uncertainties, tends not to move the area forward.

And the issues of multi-pollutant estimation are here to stay, even if we are in the early stages of addressing them. There is a need to see how they can be used in HIA, and we wanted to avoid setting a precedent of looking on adjusted coefficients as unusable, unless we really thought that they were. And we thought it unlikely that COMEAP would in the near future have a better opportunity to find out – we have no strong indications of a significant increase either in the evidence base of available and relevant cohort studies or in methodological developments that would overcome the difficulties we all agreed were present.

Finally, for most of us who favoured quantification there was a practical aspect also. We thought that the difficulties with adjusted coefficients could be overcome sufficiently that the results from applying them could be used. We were curious to know how the qualitative conclusion of the Interim Report might translate into quantitative estimates of how big is “only a little higher than what would be estimated for either $PM_{2.5}$ or NO_2 alone”?

4. RE-CONSIDERATION OF THE EVIDENCE BASE

4.1 Studies included for estimating adjusted coefficients

Those of us interested in implementing adjusted coefficients agreed that the adjusted coefficients from Cesaroni et al. (2013) and Carey et al. (2013) were, in practice, uninformative because of the very high correlation between NO_2 and $PM_{2.5}$ (see earlier) and so were excluded, whereas (from the viewpoint of correlation between NO_2 and PM), the other four studies of Table 1 seemed usable.

We excluded also results from HEI (2000) because clearly there was something unusual going on there about NO_2 and indeed a more recent study based on the same American Cancer Society (ACS) cohort had become available (Turner et al., 2016). This newer study gave results that were much more in line with other studies with regard to NO_2; the decision to exclude HEI 2000 was therefore not a difficult one. We did not however use Turner et al. (2015) instead of HEI (2000) – its adjusted coefficients were not detailed enough to allow inclusion and it had been published after the cut-off date for the literature review of the present report.

As well as including Beelen et al. (2014) and Jerrett et al. (2013), we decided to continue to use Fischer et al. despite its limitations (see earlier), because the number of available studies was small, the authors had done good work with the available data, and the study itself was a very large one.

Finally, we included results from a fourth study (Crouse et al., 2015). This had been included in the meta-analysis of single-pollutant unadjusted coefficients. However it had not been included with the six studies of Table 1, because the adjusted coefficients it reported were based on a 3-pollutant model (i.e. with ozone also) and not simply a 2-pollutant one. With so few studies available, we decided to include it: while different from the other (2-pollutant) studies, it certainly meets the need of providing coefficients for adjusted for other pollutants, and it had been published before the cut-off date for the present work

This gives a revised table, Table 2, based on four studies, three of which had appeared in Table 1 with identical data.

Table 2: Summary results from four cohorts eventually used for quantification using adjusted coefficients[1].

Study	Cohort	Correlation $NO_2/PM_{2.5}$	NO_2 IQR ($\mu g/m^3$)	NO_2	NO_2 adjusted for $PM_{2.5}$/ PM_{10}	%	$PM_{2.5}$/ PM_{10} IQR ($\mu g/m^3$)	$PM_{2.5}$/ PM_{10}	$PM_{2.5}$/ PM_{10} Adjusted for NO_2	%
Beelen et al 2014[2]	ESCAPE	0.2-<0.7	10.0	1.015 (0.993, 1.036)	1.007 (0.967, 1.049)	53	5.0	1.070 (1.016, 1.127)	1.060 (0.977, 1.150)	14
Fischer et al 2015[3]	DUELS	0.58[4]	10.0	1.027 (1.023, 1.030)	1.019 (1.015, 1.023)	29	2.4	1.019 (1.016, 1.022)	1.010 (1.007, 1.013)	46
Jerrett et al 2013	ACS CPS II	0.55	7.7	1.031 (1.008, 1.056)	1.025 (0.997, 1.054)	19	5.3	1.032 (1.002, 1.062)	1.015 (0.980, 1.050)	53
Crouse et al 2015	CanCHEC	0.40	15.2	1.052 (1.045, 1.059)	1.045 (1.037, 1.052)	13	5	1.035 (1.013, 1.049)	1.011 (1.003, 1.020)	68

Explanatory Notes: 1: Coefficients expressed per IQR except for Crouse et al. (2015a), which used per mean minus 5th percentile rather than IQR, and Beelen et al. (2014), which used per 10 $\mu g/m^3$ NO_2 and 5 $\mu g/m^3$ $PM_{2.5}$ 2: Based on 14 cohorts in which correlation between NO_2 and $PM_{2.5}$ was less than 0.7 (figures to 3 decimal places provided by personal communication are reported here; these were not available to us during our initial consideration),), 3: PM_{10};

4.2 Completeness of these studies as forming the evidence base

We are confident that these studies represent the available evidence base of NO_2 adjusted for PM published before 5 October 2015, because they are derived from a literature search of publications that was comprehensive in its search for cohort studies of mortality and NO_2 published before that date

There had been reservations about attempting, within the present work on NO_2, to derive any updated coefficients in $PM_{2.5}$, whether unadjusted or adjusted for NO_2, because no comprehensive literature search had been carried out to identify all relevant cohorts and find relevant coefficients. This is a valid concern for any new meta-analysis of single-pollutant unadjusted $PM_{2.5}$ coefficients. However, any cohort providing both $PM_{2.5}$ coefficients adjusted for NO_2 and NO_2 coefficients adjusted for PM, is likely to have been identified in the initial literature search for cohorts linking mortality with annual average NO_2; i.e. it ought to be that the four studies now identified are either the entire relevant evidence base, or a major part of it.

Consequently we think that the set of four studies tabled here is a reasonable basis for deriving and using adjusted coefficients, both for NO_2 adjusted for PM and for $PM_{2.5}$ adjusted for NO_2. Also, importantly, there was now much greater coherence between these four studies in what happened to both pollutants after adjustment than there was previously between the six studies included originally (Table 4.1). (The percentage reductions in coefficients to three decimal places from Beelen et al. are somewhat dissimilar from the reductions in the other three studies; these figures were provided by personal communication and were not available to us during our initial consideration).g.

5. OPTIONS FOR DEVELOPING AN ADJUSTED COEFFICIENT

5.1 Generic approach

There were two main options in developing a generic form of adjusted coefficient using data from any of the four studies. One is to use the adjusted coefficients from that study directly, i.e. use directly the coefficients in Table 4.2, with associated CIs. The other is to take, for each of NO_2 and PM, the % by which the coefficient is reduced on adjustment, and apply that % reduction to the relevant pooled unadjusted coefficient from a suitable meta-analysis.

These two approaches have different strengths and weaknesses, in respect of representativeness and/or transferability, and of providing estimates of uncertainty.

Representativeness and/or transferability

There are always issues about the transferability to the UK of relationships from where primary studies were carried out. Using the second approach, i.e. transferring only % reduction (or, equivalently, % 'retained'), limits the extent to which the particularities of any of the four studies determine the value of the final adjusted coefficient. This is because the only study characteristic which we transfer is an estimate of the effect of adjustment, expressed as % reduction in the coefficient due to adjustment; the underlying size of the effect from the particular two-pollutant study is not used, except insofar as it contributes to the meta-analysis of the unadjusted coefficient. That meta-analysis is based on a much greater and much more representative body of evidence about the underlying size of the effect than can be provided by any of the four studies of Table 2; and so from this viewpoint the second strategy is preferable.

Uncertainty estimates / confidence intervals

Direct use of the adjusted coefficients from a particular study gives not only a pair of adjusted coefficients, but also an associated pair of CIs. We did not have a means of deriving a valid CI for an adjusted coefficient derived using the second approach, i.e. by applying a % reduction to an unadjusted coefficient. Such an adjusted coefficient is the product of two component parts which can reasonably be considered as independent of one another, because the study which provides the % reduction has at most a limited effect on the meta-analysis which provides the unadjusted coefficient. Consequently a CI for the adjusted coefficient can be fairly easily derived if we have a valid CI for each of the component parts.

Now we do have a valid CI for one of these components, the underlying unadjusted coefficient, because for each of NO_2 and $PM_{2.5,}$ the meta-analysis which provided the underlying unadjusted coefficient provided an estimated CI also.

We did not, at least initially, have a corresponding CI for the % of the unadjusted coefficient retained after adjustment, in the selected two-pollutant study. However, development work, reported in Working Paper 3 led to derivation of an approximate CI for this component also.

Option selected

We opted for the second strategy, of transferring % reductions only, on the grounds that it was more important to maintain a more representative estimate, albeit with limited knowledge of its true CI, than it was to base the entire quantification on a single study, with whatever its particularities, even though we then had valid estimates of associated CIs. This second strategy was also the approach preferred by the Committee as a whole when the use of an adjusted coefficient, for NO_2 only, was first considered. For Fischer et al. (2015), it meant treating a % reduction on adjustment in PM_{10} as if it were a % reduction in $PM_{2.5}$ – again, this was considered preferable to omitting the study, or to using directly an adjusted coefficient in PM_{10}, and possibly applying an adjustment factor to 'translate' this into an estimated $PM_{2.5}$ effect.

5.2 Choosing an average value for % reduction in NO_2 coefficient on adjustment for PM

There are circumstances where the main focus is on NO_2 adjusted for $PM_{2.5}$, not on $PM_{2.5}$ adjusted for NO_2. This could arise, for example, for effects close to source of policies that change NO_2 itself but not the mixture as a whole, or quantifying the mortality burden attributable to NO_2 itself (though the Committee chose not to attempt such a quantification). This raises the question: (How) Can some representative value of % reduction in NO_2 be obtained, from across the four studies, as a step towards deriving an adjusted NO_2 coefficient? We considered three options, the first two of which had been discussed by the Committee as whole earlier in its work.

One approach, and a natural one if feasible, was to do a formal meta-analysis of % reduction across the four studies. However, the Committee as a whole had accepted that at present a meta-analysis of % reduction could not be done validly.

Another option was to select % reduction from a particular study - COMEAP (2009) had based its risk coefficient for mortality and long-term exposure to $PM_{2.5}$ on the results of Pope et al. (2002), despite results being available from other cohorts also. The study considered was Fischer et al. (2015) which had a dominant influence on a limited earlier attempt at meta-analysis of % reduction across six studies (though this was of questionable validity – see above). However (and contrary to Pope et al., 2002) there are significant methodological limitations to Fischer et al. (2015). These have been already noted. For these, and other, reasons, we did not wish to use only the estimated 29% reduction from Fischer et al.

Considering the distribution of values of % reduction in and other characteristics of the four studies, we instead opted to use an overall % reduction of 20%, as an informal representative value.[1]

[1] [1] At the time of our discussions, the single- and two-pollutant coefficients reported by Beelen et al. (2014) from the 14 cohorts in which correlation was <0.7 were only available to us to 2 decimal places as published: unadjusted HR 1.01 (0.99-1.04) per 10 $\mu g/m^3$ annual average NO_2; HR adjusted for $PM_{2.5}$ 1.01 (0.97-1.05). Tables 1 and 2 include these coefficients reported to 3 decimal places (subsequently kindly provided to us by the study authors). These indicate a larger reduction on adjustment for $PM_{2.5}$ than is suggested by the published coefficients.

6. ADJUSTED COEFFICIENTS AND NO_2 ITSELF

6.1 Interpretation of the adjusted NO_2 coefficient

As has been discussed repeatedly throughout this report, our understanding of causality of NO_2 implies that if a coefficient in NO_2 from epidemiological studies is used to estimate an effect of NO_2 itself, the result includes some effect also of co-varying pollutants. Consequently it is an overestimate and should be regarded as an upper limit of the real effect of NO_2 itself. Strictly, the degree of over-estimation is unknown, but it will be greater if an unadjusted coefficient rather than an adjusted one is used (because an adjusted coefficient removes some of the correlated effect of PM and gives a lower result). This suggests use of an adjusted coefficient.

Was it possible to be informative about where the true value would lie? Discussions with Defra had suggested that, if a range of 0-100% of the adjusted coefficient was given as a recommendation, analysts would pragmatically tend to use the mid-point (50%) as the 'best' estimate. This was considered not unreasonable in real terms also. Neither extreme (0% or 100%) was regarded as a plausible estimate: There was evidence from time-series and mechanistic studies suggesting that NO_2 was causal to some extent, so an option of zero effect could be ruled out; and use of 100% implied that, after adjusted, co-varying pollutants had no residual effect, which seemed similarly unrealistic.

Later discussion in the main Committee suggested an effect of NO_2 itself of 40% of an *unadjusted* coefficient, with an associated plausibility range of 25%-55% - see main report. This proposal worked well for those of us who favoured using an adjusted coefficient, in that it gave a symmetric plausibility interval around a central 50% value of the adjusted coefficient, as had been discussed

6.2 Use for impact assessment

Worked examples of using the adjusted NO_2 coefficient for assessing a policy intervention that affects NO_2 itself but not other pollutants of the mixture, are included in the main report.

6.3 Use for burden assessment

We do not, at this time, have any worked examples of using the adjusted NO_2 coefficient to estimate the mortality burden of long-term exposure to NO_2 itself. The option of doing so was taken "off the table", in order to streamline the NO_2 work, at a time when there were substantial differences of views about the use of adjusted coefficients, the assessment of mortality burden, and the causality of long-term exposure to NO_2 itself; and before the idea of 40% reduction of unadjusted coefficient had been developed. Our view is that it is reasonable to estimate the mortality burden of NO_2 itself, with suitable *caveats* about the approximations and uncertainties necessary to do so.

7. ADJUSTED COEFFICIENTS AND THE EFFECT OF A MIXTURE

7.1 Using pairs of % reduction from the same study – doing it 4 times

As can be seen above, it was not straightforward to find a meaningful average for % reduction on adjustment even for a single coefficient, but a value of 20%, was selected for % reduction in NO_2 adjusted for PM.

It is not difficult arithmetically to propose a corresponding % reduction in PM, on adjustment for NO_2; the resulting value is an estimated 45% reduction in PM, on adjustment for NO_2; but similar issues arise about the meaning and use of such an average. This idea is neither discussed nor used in the main report.

Additionally, and perhaps more importantly, we had reservations about using such a pair of average reductions, rather than a pair of values of % reduction from any one particular study. This is because the particularities of the correlation structure in any one study affect *simultaneously* the adjustment of NO_2 for PM and the adjustment of PM for NO_2; and so adjustment using a pair of coefficients, or a pair of values of % reduction, *from the same study*, somehow preserves something of that mutual dependence which may be lost if average values across the four studies are used. In addition, using an average loses potentially useful information about the extent to which applying a pair of % reductions from each of four studies does, or does not, give similar results.

7.2 Use for impact assessment

Worked examples using this methodology for assessing a policy intervention that affects a pollution mixture including, but not limited to, NO_2.

7.3 Use for burden assessment: methods and results

The methodology has been applied to estimate the mortality burden of the overall mixture. Additional details of the methods, within the broad framework outlined here, and details of the results, are given in Working Paper 3, where similarities and differences in results are also discussed.

Working Paper 3 also considers, in the light of these results, the general question of whether the mortality burden has been estimated sufficiently well for it to be used to inform, both for policy makers and for the general public, the debate on what is the annual mortality burden attributable to air pollution in the UK. In that context, the Discussion that follows focuses on methodological issues of using adjusted coefficients, in the light of the methods that have been proposed.

8. DISCUSSION OF METHODOLOGICAL ISSUES

The two previous Sections highlight that it is possible to use adjusted coefficients to address all four of the questions of interest identified earlier (Section 1.5). This Discussion considers some of the uncertainties and limitations in using adjusted coefficients in the present context (of quantifying the effects on mortality of long-term exposure to NO_2 and PM, and specifically about their use for estimation of mortality burden) guided at least in part by the issues identified by the minority group.

Highly correlated pollutants

Broadly speaking, the issue has been addressed already via selection of studies, which excluded the two most highly correlated studies. It is of interest that these exclusions, together with that of HEI (2000) reduced substantially the between-study differences in what happens on adjustment.

Possible interaction between $PM_{2.5}$ and NO_2

In principle this is an issue and in practice we don't know how serious it is. There is little that we can say definitively because we do not know whether or not there is an interaction and if there is, how important it is, because none of the four studies reported results about it. It is nevertheless possible to make some remarks.

One aspect, is whether or not any interaction is statistically significant; i.e. we can be confident that it is real. Also important is whether or not it is large, relative to the estimated sizes of the adjusted coefficients estimated without interaction. Put differently, at the concentrations of the pollutants in question, does taking an interaction into account matter much? If a study is large enough, it is possible for an interaction to be statistically significant without affecting much the final answers – and indeed conversely, it is possible for an interaction to have a substantial effect on answers without being statistically significant.

Additionally, if there really is a statistically significant and substantial interaction, then unadjusted coefficients will be affected also – the value of an unadjusted coefficient will be some (weighted) average value, averaged over the concentrations of the 'other' pollutant, not included in the model. The statistical modelling tradition however is to have an explanatory model no more complex than it needs to be, and if an interaction in the present context really was a very important issue, it is curious that none of the four underlying studies, with experienced authors, reported having modelled it or tested for it.

The lack of statistically significant difference between the combined effect from a two-pollutant mixture estimate and the effect as estimated using a single-pollutant model

As noted earlier, the Interim Statement of December 2015 noted that "the combined effect of NO_2 and $PM_{2.5}$ estimated using coefficients where each is adjusted for the effects of the other, is either similar to or only a little higher than what would be estimated for either $PM_{2.5}$ or NO_2 alone, using unadjusted single-pollutant coefficients". Indeed, as noted in the Main Report and Working Paper 1, the overlap of CIs suggests that the differences between the combined effect from a two-pollutant mixture estimate and the effect as estimated using a single-pollutant model are not statistically significant. This means that the results are compatible with there being no difference between the two-pollutant and the single-pollutant estimates of the effect of a mixture; i.e. the viewpoint that there is no real difference between the results is *tenable*, in the light of the data. However a viewpoint being tenable doesn't make it the only or best viewpoint and certainly the results are also compatible with the combined or aggregated result being somewhat higher than the single-pollutant one.

This closeness of results does however raise the question: if one were to quantify the effect of a mixture using a single-pollutant rather than a two-pollutant model, which single pollutant should be used? It seems that there is no general answer. The results of the four individual studies (based on effects as estimated at the IQRs of NO_2 and PM) are not convincing about which single pollutant is the better one to base single-pollutant estimation on. It may well be, for example, that between NO_2 and $PM_{2.5}$ the preferred single pollutant for assessing the impact of the mixture from local transport policies is not the same as that for estimating the burden of air pollution in the UK.

Given these uncertainties, a reasonable strategy, if a single-pollutant quantification were to be used, is to quantify both in (unadjusted) NO_2 and (unadjusted) $PM_{2.5}$, and choose the larger value, on the grounds that either single-pollutant estimate is likely to under-estimate rather than over-estimate the mixture effect. This was, eventually, the view of the majority of the Committee.

What are the implications of all of this? For quantification where approximate answers are sufficient, or resources are scarce, we suggest doing single-pollutant analyses in both NO_2 and PM (using unadjusted coefficients) and choosing the one that gives higher values. For more detailed modelling, and where resources permit, then two-pollutant modelling seems preferable.

Differential exposure assessment/measurement error and different spatial scales

Measurement error denotes the differences between what we hope to measure (in this case, personal exposures of individuals attributable to outdoor air pollution) and what in fact we do measure (in this case, measured or modelled long-term average concentrations of NO_2 and PM near a subject's residence, based on a limited number of monitors city-wide). In the present context, differential measurement error arises when the gap between what we wish to measure and what we do measure is different for NO_2 and for PM. With differential measurement error there is a danger of 'transfer' of apparent effect from one pollutant to another; the expected direction of 'transfer' is from the more weakly measured pollutant (i.e. the one with greater 'measurement error') to the more strongly measured one.

Typically, within cities there is more spatial variation in (annual average) concentrations of NO_2 than in (annual average) $PM_{2.5}$. Then, unless particular resource is put into measuring NO_2, using city-specific values to represent the 'exposures' of all residents in a city will imply greater 'measurement error' in NO_2 than in $PM_{2.5}$. Also, it has been noted that penetration indoors of $PM_{2.5}$ is more effective (i.e. more similar to outdoor concentrations) than for NO_2 – this would add to the 'measurement error' in using outdoor concentrations as a marker of personal exposure. Factors such as these may at least in part explain why in a major study like ACS / HEI 2000, which uses between-city contrasts in pollution, there is a clear $PM_{2.5}$ effect but not an NO_2 one. In within-city analyses, however, with resultant 'better' characterisation of NO_2 concentrations experienced by individuals, an NO_2 effect shows through more clearly.

In principle this affects unadjusted as well as adjusted coefficients. Intuitively however the effect on adjusted coefficients is likely to be greater.

Transferability of relationships between locations - from where the studies were carried out, to use of the coefficients for quantification in the UK

This is a key issue. Transferability depends on the nature and extent of the differences, between the study location (and its pollution and other characteristics, including population) and corresponding characteristics of the location where the results are being applied, in this instance the UK. As indeed there are many kinds of difference between the characteristics of locations of the primary studies and those of the various countries and regions of the UK, where the results will be applied. The final sub-sections of this Chapter consider some of these issues in detail. Here we make some general overview remarks about transferability.

Firstly, it is reasonable to note initially that it follows from the above that there are transferability issues for single-pollutant / unadjusted coefficients also. So the question isn't

“Are there transferability issues?” It is: “Is transferability for adjusted coefficients so much worse than for unadjusted ones that we can transfer unadjusted coefficients but not adjusted ones?” That’s the criterion we need to have when we consider whether or not adjusted coefficients are usable.

Note also that this is not just a statistical issue about coefficients; it is a scientific issue about when necessarily different circumstances (e.g. between where primary studies are carried out and where the results are applied) are so different that inferences from one cannot reasonably be applied to another. Rejecting transferability raises the question: “What if anything can we therefore ever say, quantitatively, about the effects of pollution in a particular context (here, the UK), based on evidence from studies conducted elsewhere?” And indeed, what can we say reliably, even qualitatively?

The fact that rejection of transferability has serious consequences for what we can know, in the future as well as now, does not mean that we should support transferability at all costs. It does mean that we need not to reject it without very good reasons.

Thirdly, transferability depends also on exactly what summaries of the original study results are transferred. In the present context, the effects of transfer are mitigated by transferring pairs of adjusted coefficients from the same study. They are mitigated by using % reduction rather than the actual adjusted coefficients. And as discussed below, there are additional safeguards in doing the burden estimates 4 times, with (presumably) different transferability issues for each of the four studies.

Finally, transferability depends on how those transferred summaries are used. And again, in the present context, it seems that there is less uncertainty in using pairs of adjusted coefficients to jointly estimate the effect of a mixture, than in using individual adjusted coefficients (or pairs of adjusted coefficients) to estimate separately the effects associated with the NO_2 or $PM_{2.5}$. The requirements for valid estimation of the effects of the pollutants separately are stronger / stricter than the requirements for valid estimation of the overall mixture, in that it is possible to be confident about the estimated effect of a mixture while retaining doubts about the value of the individual components which are added together to give the mixture estimate.

Differences between locations in pollution mixtures, including in the correlation structure between pollutants

Certainly there are differences between locations in the pollution mixtures and in the correlation structure between pollutants; in the present context that is between where the studies were carried out and the UK, where we wish to apply the results. At first sight these differences appear to be a problem, raising questions about transferability. But there are advantages also. For example it seems (Carey et al., 2013) that the correlation in the UK between long-term exposure to NO_2 and to PM is high, indeed so high as to make it very difficult, or impossible, to disentangle the effects of NO_2 and PM based on UK data only. We may however be able to make progress by examining the effects in locations and populations where the correlation is weaker, and there is a greater chance of separating the effects; and then apply that knowledge to the UK situation.

This possibility is of course strengthened if we have several studies elsewhere. Presumably these have different mixtures and different correlation structures one from another also, as well as differences with / from the (various countries and regions of the) UK. Consequently we can examine the extent of variation in results from using adjusted coefficients from *each* of these different primary studies. If they give similar results, then presumably the differences in correlation structure are not having a dominant effect on results. In the present context we have only four external studies from which to draw conclusions, so our perspective is limited, but having adjusted coefficients from four studies does give at least some perspective on consistency of results.

Extrapolating beyond the range of the data in the original studies

The choice between quantifying only above a cut-point reflecting the lowest concentrations where effects have been demonstrated (which, for cohorts associating mortality with long-term exposure to NO_2, COMEAP had earlier assessed as 5 µg/m^3 annual average NO_2) or quantifying down to some lower value, possibly zero, has exercised the Committee for many years. The core issues apply to unadjusted as well as to adjusted coefficients. As before, the question for us is: "Are the issues as they apply to adjusted coefficients clearly more severe than those as they apply to unadjusted ones?"

The core issue is that extrapolating beyond the range where effects have been demonstrated involves additional uncertainties compared with estimating above a cut-point only. But the reason for doing so is that extrapolating beyond the range of the data is necessary in order to answer a meaningful and relevant burden question, i.e. about how big is the size of the problem caused by air pollution.

In a little more detail: As an answer to the question "What is the burden of NO_2 above 5 µg/m^3 NO_2?" restricting the quantification to concentrations >5 µg/m^3 annual average avoids making any assumptions about the behaviour of the concentration-response function at lower concentrations, i.e. it avoids extrapolation.

However the burden question, as usually understood, isn't about NO_2 concentrations >5 µg/m^3, it's about the burden associated with NO_2 at all concentrations of NO_2. And if the quantification of NO_2 burden at NO_2 concentrations >5 µg/m^3 is offered as an answer to the usual burden question, in effect it assumes that the burden at < 5 µg/m^3 NO_2 is zero, i.e. that an annual average of 5 µg/m^3 is a genuine threshold for the effects on mortality of long-term exposure to NO_2. Now this is both a strong assumption and an unrealistic one. Treating as zero the effect associated with NO_2 at < 5 µg/m^3 NO_2 doesn't avoid the issue of extrapolation that the use of a cut-point is intended to avoid. Instead, it is a strong assumption about the behaviour of the curve at concentrations lower than where we currently have reasonable evidence.

And it is an unrealistic assumption for at least two reasons. Firstly, the evidence on causality of NO_2 (or pollutants associated with it) does not suggest that there is a genuine population threshold or safe level. And secondly, experience from research on PM and on NO_2 show that attempts to treat "the lowest concentration at which effects have been demonstrated" as a genuine population threshold or safe level have, over time, always been proved to be wrong, because new studies have demonstrated effects at even lower concentrations.

Under the circumstances, quantification down to zero via simple extrapolation seems a much more realistic option though with, as acknowledged, some added uncertainties because of extrapolation.

Given then that there is a need to extrapolate burden estimates down to zero, what then of the question: “Are the uncertainties as they apply to adjusted coefficients clearly more severe than those as they apply to unadjusted ones?” We are not convinced that the issue is a major one, in that annual average concentrations of NO_2 and PM in cities in the UK are not very different from those where cohort studies have been carried out, and the relevant core studies have not used a threshold or cut-point model. Extrapolation would be a bigger issue if estimates of burden were dominated by effects in locations where annual average NO_2 is less than 5 $\mu g/m^3$ and, fortunately for the methodology though, unfortunately for the residents, this is not the case.

9. FINAL REMARKS

Given the evidence, the uncertainties and the complexity of the issues, there is certainly legitimate reason for differences of views about what approaches are not only best, but usable, or not usable, at all. And there are good reasons why these issues matter.

Perhaps the key one is that air pollution is a major public health issue and it matters that we have a correct understanding, or at least the best available understanding, of the role of different key regulated pollutants, and that we can communicate that understanding, together with its strengths and weaknesses, in a way that limits the chances of well-meaning mis-interpretation or of selective quoting by others to support a pre-determined position.

But the issues matter as well from the viewpoint of the methodology that we use. In an era when expertise is intentionally side-lined and policy is sometimes based on hurt emotion rather than assessment of the evidence, it is important that we care for the methodological tools available to help us draw out the policy implications of the relevant science that is done. Here some of us wanted to avoid the danger of the Committee pushing the evidence beyond what the evidence reasonably could support. Others wanted to avoid the danger of the Committee setting aside without strong enough reason adjusted coefficients, the fruits of complex multi-pollutant modelling, one of the key emerging tools available for tackling the issues of multi-pollutant sources.

We differed in our sense of these dangers. Given that there are well-informed and able people on both ‘sides’ of the arguments, it’s probably too early to know what’s right. Time will tell – it will be interesting sometime to look back with what we hope will be the benefits of a better understanding of the methodology.

In the meantime it is reassuring that different approaches (i.e. using unadjusted and using adjusted coefficients in quantifying the mortality effects of long-term exposure to pollution mixtures containing both PM and NO_2) lead to somewhat similar practical conclusions about the size of air pollution as a public health problem in the UK, even if the routes by which it may be controlled seem, at least for now, more complex than they did some years ago, when it seemed that all that was necessary was ‘simply’ control of population exposure to human-made fine particles ($PM_{2.5}$).

REFERENCES

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A. S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N. L., Ostenson, C. G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K. T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K., Wang, M., Bueno-De-Mesquita, B., Sugiri, D., Kramer, U., Heinrich, J., De Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Kunzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M. Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Katsoulis, M., Trichopoulou, A., Brunekreef, B. & Hoek, G. 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet, 383, 785-95.

Carey, I. M., Atkinson, R. W., Kent, A. J., Van, S. T., Cook, D. G. & Anderson, H. R. 2013. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. Am.J.Respir.Crit Care Med., 187, 1226-1233.

Cesaroni, G., Badaloni, C., Gariazzo, C., Stafoggia, M., Sozzi, R., Davoli, M. & Forastiere, F. 2013. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. Environmental Health Perspectives, 121, 324-331.

COMEAP 2009. Long-term exposure to air pollution: effect on mortality. Committee on the Medical Effects of Air Pollutants

COMEAP (2010). The Mortality Effects of Long-Term Exposure to Particulate Air Pollution in the United Kingdom, Committee on the Medical Effects of Air Pollutants.

Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., Van Donkelaar, A., Martin, R. V., Villeneuve, P. J., Jerrett, M., Goldberg, M. S. & Pope Iii, C. A. 2015. Ambient PM2. 5, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environmental health perspectives, 123, 1180.

Fischer, P. H., Marra, M., Ameling, C. B., Hoek, G., Beelen, R., De Hoogh, K., Breugelmans, O., Kruize, H., Janssen, N. A. & Houthuijs, D. 2015. Air Pollution and Mortality in Seven Million Adults: The Dutch Environmental Longitudinal Study (DUELS). Environ Health Perspectives, 123, 697-704.

HEI 2000. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality: A Special Report of the Institute's Particle Epidemiology Reanalysis Project, Health Effects Institute.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B. & Kaufman, J. D. 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ Health, 12, 43.

Holland, M. 2014. Cost-Benefit Analysis of Final Policy Scenarios for the EU Clean Air Package. http://ec.europa.eu/environment/air/pdf/TSAP%20CBA.pdf.

Jerrett, M., Burnett, R. T., Beckerman, B. S., Turner, M. C., Krewski, D., Thurston, G., Martin, R. V., Van Donkelaar, A., Hughes, E., Shi, Y., Gapstur, S. M., Thun, M. J. & Pope, C. A., 3rd 2013. Spatial analysis of air pollution and mortality in California. Am J Respir Crit Care Med, 188, 593-9.

Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, and Thurston GD. (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*. **287**: 1132–1141
Turner, M. C., Jerrett, M., Pope Iii, C. A., Krewski, D., Gapstur, S. M., Diver, W. R., Beckerman, B. S., Marshall, J. D., Su, J., Crouse, D. L. & Burnett, R. T. 2016. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. American Journal of Respiratory and Critical Care Medicine.
WHO 2013a. Review of Evidence on Health Aspects of Air Pollution - REVIHAAP Project: Technical Report. World Health Organization, Copenhagen.
WHO 2013b. Health risks of air pollution in Europe – HRAPIE project. Recommendations for concentration–response functions for cost–benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen: World Health Organization.