THE ROYAL SOCIETY

Research priorities in air pollution and health - recommendations from a Royal Society workshop

Introduction

This paper provides recommendations for enhanced approaches to research on the health effects of air pollutants. It is based on the outcome of an international workshop on research priorities in air pollution and health hosted virtually by the Royal Society on 14th and 15th September 2020.

The objective of this workshop was to identify new approaches to research on the health effects of air pollutants that will help to better understand the relative contributions to observed health effects from individual pollutants from exposure to a combination of air pollutants. Improved understanding of relative contributions from individual pollutants would allow policy measures for air pollutant control to be better focussed and yield the most cost-effective solutions to minimise impacts upon public health.

The workshop took the form of four plenary talks by experts centred on a specific research question which were followed by breakout groups, each of which was provided one of those four research questions and sub-questions to focus their discussions. A summary of the workshop discussion can be found in annex 1, and the agenda and the attendees list are in annexes 2 and 3, respectively.

The target audience for this synthesis includes research funders, policymakers, and researchers developing proposals for work in this field.

Recommendations

General

- The focus of the workshop was to identify and prioritise research approaches to quantify the relative contributions to health outcomes of individual pollutants within the mixed pollution climate outdoors and indoors. The workshop discussion led to an overarching recommendation that more research is needed that integrates toxicological and epidemiological research activity, while taking account of the best methods for exposure estimation.
- It is critical to recognise that the use of traditional multi-pollutant model approaches leads to difficulties in interpretation and challenges in producing the information required for further analyses of health impacts. There is a need to develop new multi-pollutant model approaches and to employ recent innovations in epidemiological studies.
- Future toxicological research should take account of possible interactions between components of pollutant mixtures, including particle components.
- Future research proposals should harmonise toxicological test methods. This would facilitate comparisons between studies and help in elucidating the effects of specific exposures and the ranking of toxicity of different components of airborne particles.
- Human challenge studies are to be encouraged. The development of a high-quality UK chamber facility for such a purpose is highly recommended.
- Quantification of the effects of air pollutant exposure in the development and maturation of organs in children is of high importance.

Omics[1]

- A focus on hypothesis-free approaches (metabolomics, transcriptomics, proteomics) should be encouraged, especially in longitudinal cohort studies where early signs of disease can be identified.
- Discovery and exploitation of new biomarkers of exposure and effect might offer considerable new opportunities to advance the field of air pollution effects research, as well as the use of transgenic humanised mice in mechanistic studies.
- There is a need to clarify the predictive capability of in vitro test methods such as Oxidative Potential (OP) for effects in human populations, recognising that different OP methods may be predictive of different outcomes.

Modelling

- Reconstruction of past exposures of established cohorts for a range of pollutants, taking full account of changes to particulate matter composition, may facilitate valuable epidemiological studies.
- Air quality models with a high spatial resolution (metres) should be further developed and validated, and made widely available to the research community.
- There is considerable scope to use changes in the air pollution climate (long and short term) resulting from emissions reductions to elucidate the effects of specific pollutants on health.
- Research into the effects of exposure misclassification in epidemiological models, including methods for correction of effect transfer due to differential measurement error is a high priority.

Indoor air quality

- There is an urgent need for comprehensive information on indoor air quality in the UK and its connection with housing characteristics and personal lifestyle factors.
- Studies of the health impacts of indoor emissions from sources such as cooking and cleaning agents are a high priority as the evidence base on this is not well developed.
- Researchers should seek to exploit low-cost sensors in the evaluation of spatial patterns of air pollution and personal exposure, whilst recognising the current limitations of these devices.

---

[1] Omics refers to the collective technologies used to explore the roles, relationships, and actions of the various types of molecules that make up the cells of an organism.

Annex 1 - Detailed discussion overview

Research question A: Can either toxicology or epidemiology give clear and consistent insights into the relative toxicity of particulate matter of differing source, composition and/or size distribution in human populations?

Background

There are major divergences between the findings of toxicological and epidemiological studies relating to the relative toxicity of the components of airborne particulate matter. These may be largely explicable by considering the differences between the two approaches.

Epidemiology studies exposures to variables in particle mixtures by using long-term data on exposure rate and repeated exposures. Toxicology studies relative toxicity in experimental models, often in simple exposure systems and short duration experiments. It is difficult to compare the data resulting from toxicology and epidemiology studies as the impact of exposure is not determined in the same way. Using the same exposure metrics (e.g. composition, size distribution and source) in toxicology and epidemiology in the lab and in the field would be helpful but is not easily achieved.

Particle toxicity can be influenced by source, chemical composition and size distribution (which are not wholly independent variables) and it is important to ensure that the markers and endpoints under investigation are relevant to the disease. It may therefore be useful to focus on specific independent endpoints and not just inflammation or oxidative potential. In vitro tests such as measurement of oxidative potential and animal studies have a role to play; additional analytical approaches are needed to better translate the findings to human exposure.

Discussion

Are the results of in vitro tests (e.g. oxidative potential) predictive of toxic effects in human populations?

- In vitro tests are relatively quick and easy, particularly in ranking toxicity. They provide insights into exact biomolecular pathways of toxicity, not isolated to oxidative stress, which can indicate mechanisms involved in human toxicity. However, the models lack complexity; for example, they have limited ability to study cell-cell interactions and systemic metabolic processes.
- Although identifying the toxicity of individual components of air pollution is relatively straightforward when using in vitro models (though not so easy for studying mixed exposures), it remains important to test the emerging hypotheses from in vitro studies using in vivo models, to extrapolate to human exposure.
- There is a need for more informative in vitro assays and more sophisticated in vitro systems. More sophisticated in vitro models are currently being developed. They include multicellular co-culture models of structural and immune cells as well as lung slices with aerosol exposures and blood flow. They use highly relevant primary mammalian and human cells and tissues, with significant potential for future studies.
- In toxicology, it is important to utilise realistic exposures and relevant models of optimal human health as well as models of disease, in order to predict mechanisms involved across susceptible groups (e.g. asthma, young and elderly) for a range of exposure scenarios.
- Oxidative potential depends on many factors, not just the nature/mixture of the environmental exposure, but also temporal responses and susceptibility. Use of appropriate in vitro and in vivo models (e.g. acute, chronic; showing resistant and susceptible traits that could be introduced by genetic manipulation) that exhibit relevant bio-physiological characteristics would provide more holistic data.

- There is also a need to agree on what endpoints are truly associated with exposure and the resulting cellular bioreactivity using standardised and state of the art analyses; harmonisation of these analyses is needed.

Can the results of studies with animal models be meaningfully extrapolated to human populations?

- Extrapolating the results of animal model studies to human populations has uncertainties. They can establish the relevance of in vitro data to human populations, providing a pathway to controlled human studies in humans to link toxicology with epidemiology. Importantly, animal studies can examine single and mixed components of air pollution, as well as other fundamental variables contributing to risk factors of environmental pollutants, for example, the impact of oxygen tension. The use of genetically modified and humanised animal models would increase the significance of such studies.
- Toxicokinetics, toxicodynamics and a better understanding of what happens when particles and/or associated molecules translocate across the respiratory membrane to another organ (possibly inducing secondary toxicity, an underexplored area) can be better assessed by using in vivo models.
- To optimise use of animals and improve the overall outcome and information gained, in vivo studies could be designed to address a range of questions, analysed by a number of groups and research centres with complementary interests and expertise, where all the tissues and organs are studied and nothing is wasted. In this context, a centralised bio-bank for animal tissues would be a useful resource for environmental toxicology.

Can human challenge studies offer meaningful answers on relative toxicity?

- Human challenge studies offer meaningful answers on relative toxicity but tend to be limited to simple exposures of single components of air pollution mixtures under acute conditions. Much insight into chronic effects could be gained if such studies were also performed, under controlled circumstances, to better understand the relationship between acute and chronic exposures.
- These types of studies lend themselves to crucial state-of-the-art analytics that potentially identify novel, previously unknown, mechanisms of action of the pollutants and their impact on human health. Furthermore, finding accurate biomarkers of effect is key to understanding the degree to which acute conditions extrapolate to chronic conditions.
- More multi-centre studies would be useful, with a focused integration between epidemiology and experimental work (using the more sophisticated and relevant analytical strategies). A few high-quality inhalation chambers, across nations, which utilise consistent inhalation procedures, would improve quality and insights gained to accelerate progress in this area and help developing hypotheses to address precise questions.
- Limitations include the ethics surrounding human challenge studies, which are expensive and lengthy. Furthermore, previous studies have not accounted adequately for age and how acute responses play out across age group and aging; an area that should be addressed in the future.

Why do epidemiological studies not give a coherent picture of the most toxic components?

- Epidemiological studies do not always track the components of air pollution exposure, and how they interact with each other, as people move between places. They use blunt endpoints and miss the subtle endpoints that could be used to better interpret and track human studies.
- Another issue is that epidemiological studies are retrospective with little predictive power to relate the findings to ongoing and new environmental changes and how these might affect human health. One example is the increase in airborne microplastics, the toxicity of which remains unclear. Another is the changing environment in the face of new policies and controls of traffic emissions and climate change. Thus, studies that depend on old datasets or unreliable outcome measures such as death certificates are not very useful.

- Currently, many studies rely on particle mass, whereas composition, location and seasonal variations in environmental pollution also matter. The question of relative toxicity of different exposure profiles is complex and particularly challenging for epidemiological methods, which typically lack adequate data to link specific exposure patterns with health outcomes. There is a need for more accurate biomarkers of exposure, which can be linked to pollutant and dose.
- The role of epidemiological studies generally isn't to provide predictions or to determine toxic components and precise effects, but rather to give a coherent picture of the overall impacts of environmental pollution on health.
- Overall, current epidemiology lacks quality information and data on the biological impact of exposure. There is an urgent need to combine conventional measures of exposure with new techniques (such as –omics) and other approaches, not necessarily hypothesis based, to identify novel markers of exposure, and markers of mechanistic processes involved, to better evaluate and monitor the impact of exposure to components of air pollution on health.

Why do toxicology and epidemiology give apparently different insights into the most toxic components of airborne particles?

- It is important to note that there are some fundamental agreements between toxicologists and epidemiologists on the health effects of primary combustion particles, which demonstrate toxicity across every experimental model and population study. Atmospheric composition is an important field of convergence between the two approaches. Yet mass of fine and coarse airborne particles is a blunt tool in assessing public health impact and to effectively control environmental exposure.
- The focus and methods of epidemiology and toxicology are different, and each provides different insights. Epidemiology provides a macro view of total exposure over time, the difference between outdoor and indoor air quality, and exposure from transport, rather than the exact composition of the air pollution. The difference in insight from epidemiology often reflects a lack of specificity and a focus on biomarkers which are not always informative. It was noted that sources are an important aspect, perhaps more so than particle size.
- Toxicology provides a micro view, mostly of acute exposure, and focuses on the individual sources and controlled combinations of exposures. Epidemiology is dependent on understanding toxicity to identify meaningful associations with the exposures and is reliant on biomolecular science and toxicology to form hypotheses and to establish risk-associated exposures.
- For better integration, harmonisation across the two fields is needed. Identifying the best experimental model system to enable this is an ongoing process, as discussed above. Specific standard measures that identify indicators of toxicity/exposure are needed. Nanotoxicology has been establishing toxicological insights and standards based on the ability of designer particles engineered to address specific questions, including size, surface properties, shape and chemistry.
- Epidemiologists and toxicologists need to work together to provide a better understanding of the relationship between real exposure conditions and toxicological consequences. Epidemiological studies should be conducted so that toxicologists are involved from the very beginning to optimise the outcome and more accurately determine the mechanisms connecting what is inhaled to health effects, which is what is needed to control the exposure to, and health effects, of air pollution.

Research question B: Are there innovations in epidemiological methods which would allow a clear differentiation of the effects of closely correlated pollutants upon public health (e.g. nitrogen dioxide and PM2.5)?

Background

Air pollution epidemiology is a relatively “blunt tool”, with a number of challenges including limitations in relation to the availability of exposure data, reliable data sources for health outcomes, erroneous cause of death certification, and a major problem in differentiating the individual effects of closely correlated pollutants. The latter includes differentiating the effects of multiple pollutants associated with a particular source type (e.g. road traffic) and those from different sources that may be similarly affected by meteorological factors.

Some pollutants, such as $PM_{2.5}$ and nitrogen dioxide, and nitrogen dioxide and ozone, are closely interlinked and there have been no clear successes so far to isolate the health effects of each. The interaction between multiple pollutants, and the subsequent correlation between measured and modelled levels of pollution, is a major problem when using traditional epidemiological methods that are unable to accurately differentiate the health impacts associated with individual pollutants. Further complications include changes in the relationships between pollutants, and differences in demographics and human susceptibility to air pollutants, over both space and time.

The workshop explored the difficulties associated with traditional methods, the possible use of newer, innovative approaches related to causal analysis and the need for future research and methodological development that would enable the public health impacts of closely correlated pollutants to be investigated.

Discussion

Are there alternatives to the current design of multi-pollutant epidemiological models which can resolve the effects of closely correlated pollutants?

One of the important recent innovations in epidemiological modelling (of the effects of air pollution on health) has been in the development of models for causal inference in which statistical techniques, e.g. propensity scores, are used to simulate a randomised controlled trial in observational data. Observational studies are susceptible to confounding bias by factors that are associated with both the exposure and outcome of interest and failing to account for them can lead to bias. Traditionally, allowing for the effects of confounders has been achieved by including potential confounders as covariates in a regression model. However, this can lead to model misspecification and, with highly correlated multiple pollutants, will lead to issues of multi-collinearity.

1. Propensity Score Analyses

One approach is to try to represent the randomisation element (of individuals to treatment/exposure or control/not exposed) of a controlled trial and to try and incorporate this within an observational analysis. Propensity scores are the probability of exposure (or not, in the case of a binary variable) based on potential confounders. They can be found using logistic regression models and the idea is that these probabilities provide a summary of the information related to why someone is exposed (e.g. to high air pollution).

Propensity scores can then be used within analyses of health risks (outcome models) in which unbiased estimates of the effect of exposure should be attainable by comparing (the health outcomes of) subjects who had similar probabilities of being exposed. The scores can be used to match individuals (by the definition of the propensity score, two subjects with similar propensity scores should also be similar on all important covariates) and health outcomes in the two groups (exposed/not) compared. They can also be used as weights, or as covariates (alone or with other covariates) in the outcome (regression) model. However, it is noted that as with traditional methods, propensity scores can only control for the known measured confounders.

Whereas standard models result in coefficients that are conditional on all the other covariates, propensity score models measure the marginal effect, which is preferable in health impact assessments. Propensity score models can be applied to already-analysed datasets and can help with multiple pollutants. For example, possible correlation and non-linearity, between $PM_{2.5}$ and nitrogen dioxide can be controlled by weighting samples to ensure independence between the two pollutants. By manipulating the linearity and different possible interactions to get balance plots that show essentially zero standard deviations between covariates and the exposure, the models would achieve independence of all the measured covariates or other pollutants.

2. Positivity

In traditional methods, causal modelling would examine what would happen to an outcome if an exposure were reduced from A to B, keeping all covariates constant. However, for some pollutants it is not feasible to reduce from A to B given the covariates; for example, it is not realistic to lower nitrogen dioxide exposure without lowering $PM_{2.5}$ exposure. Positivity maintains that, given all covariates, it has to be possible for the individuals to have experienced all possible exposures and if not, the data should be trimmed. Previous analysis attempted to examine causal marginal estimates of an effect separating nitrogen dioxide and $PM_{2.5}$ impacts and had to exclude around 30% of records to maintain positivity.

3. Matching

With matching, the analysis uses observations where one exposure remains constant and another varies but the observations are matched on all other potential confounders, e.g. season or demographics. The use of matching reduces the sample size and consequently results in a loss of power. However, by not matching, there is significantly more variation to control for, which limits the ability to detect an effect due to the widening of confidence intervals.

4. Machine Learning Methods

Machine learning methods can progress causal modelling further if non-linearity, interactions and variations due to external factors such as temperature are influencing the relationships. Machine learning is designed to capture the complex relationship patterns in propensity score models and adjust. Examples of machine learning methods include Random Forests, Gradient Boosting, Support Vector Machines and Neural Networks.

5. Mixture Methods

Mixture methods such as the Bayesian Kernel Machine Regression can be used to identify interactions and non-linearities in the exposure-response, and Bayesian variable selection can help eliminate poor predictors. However, full separation is still an issue especially where there are complex interactions.

6. Nature-based: Differences in Differences

In some cases, nature forces exposure A to be independent of exposure B, i.e. instrumental variables that produce variations in exposure that are independent of covariates. These scenarios often occur with sudden policy changes that will break correlations. Some published examples have looked at industrial plants in the US after EPA (United States Environmental Protection Agency) regulations imposed the installation of scrubbers, and the introduction of low-sulphur fuels. There are some UK examples that are yet to be researched.

Are there viable techniques for correcting for effect transfer in current models?

- Effect transfer occurs when one pollutant has more exposure error than another. One way identified to address this it to estimate the spatio-temporal variation of exposure error for multiple pollutants. This would require robust models for exposure error. Methods such as inverse variance weighting and the spatial SIMEX methodology could then be used for correcting for measurement error. However, there will be considerable computational challenges associated with using them with multiple pollutants and where there is substantial spatial and temporal variation. Research is ongoing to integrate these error models with propensity scores and is currently not well developed.

- Overall, to correct for effect transfer in current models, simplifying assumptions are made and therefore there is a research priority for the development of computationally efficient methods to address measurement error in more complex situations.

Why do epidemiological studies give such diverse results for the relative toxicity of individual components of airborne particulate matter?

- Even with the most advanced analysis methods, there are many pollutants that are currently not monitored, measured or modelled and so cannot be controlled without doing an experimental randomised control trial, and the results will still have an unknown level of confounding bias.
- Current study designs are unable to account fully for possible pollutant interactions.

Can current epidemiological approaches generate consistent exposure-response functions for ultrafine particles?

- Reviews of health effect studies of ultrafine particles reveal a wide range of health outcomes and coefficients but suffer from poor consistency. Some of the potential causes of inconsistencies were noted as potential exposure misclassification, measurement error due to spatial heterogeneity, and/or difference in spatial smoothing of the data, as well as different source-related mixtures of particles.

Can any components of airborne particles be considered as not contributing to toxicity?

- Overall, it was agreed that all airborne particles should be regarded as toxic, as toxicity appears to have a consistent dependence on the mass concentration of particles, but also because particles may become toxic through interactions with others.

Research question C: Current estimates of the health burden due to air pollutant exposure deriving from cohort and time series studies quantify only the impacts of outdoor pollution. How can the impacts of indoor sources be evaluated?

Background

Ninety percent of human intake as mass per day/year/life is air, and outdoor air can be as little as only five percent of this intake. Depending on ventilation rates, around 25% of the pollution in indoor air derives from outdoors. Throughout the day, people are exposed to different amounts of particles at different time periods, with cooking, solid fuel heating and candles cited as example sources relevant to indoor air quality. Particles found indoors range from coarse mode (e.g. soil, sea spray, pollen) to fine (e.g. combustion soot, outdoor secondary particles) to ultrafine (e.g. combustion nanoparticles).

Exposure varies across occupations, time activity patterns and age groups, with retired people at home being exposed to a greater fraction of domestic indoor air over a day as compared to a biking postman or a bus driver. Individual exposure is highly personalized and to quantify exposure, we would ideally need measurements of all relevant parameters over long time periods, including individual housing characteristics, and the habits of people that may be driving much of the exposure differences.

In indoor air research, the 'old suspects' that may cause harm include household pests, endotoxins, and moulds. The studies presented by the speaker highlighted that households with higher detectable allergens and endotoxins had negative health outcomes i.e. increased risk of asthma as compared to households with no allergens, and people who are exposed to nicotine at home have a higher risk of getting asthma. Key unanswered questions include whether differences in particle size and composition cause different toxic effects, what the important drivers of adverse effects indoors are, and who the most susceptible groups are?

A class of pollutants of relevance to indoor air quality are volatile organic compounds (VOCs), which are reported to be consistently higher in concentrations indoors than outdoors. A study presented showed that the total amount of VOCs emanating from a product (such as a building after painting) decay with time, but at a much higher rates if ventilation is effective. It is thus essential to allow time for adequate ventilation and decay of these pollutants in environments such as newly painted rooms to minimize poor health outcomes associated with exposure.

Another study of interest highlighted the influence of clothing on uptake of semi-volatile organic compounds (SVOCs). Even for non-occupational environments, wearing clothing that has adsorbed/absorbed indoor air pollutants can increase dermal uptake of SVOCs by substantial amounts relative to bare skin. Key questions regarding organic compounds include: what are the indoor drivers of concentrations of organic compounds, both primary sources and secondary products via indoor oxidation? Who are the susceptible groups? What are the effects beyond allergy?

There are limited data available on the contribution of cleaning products to indoor air pollution. It was noted that the use of oxidative air cleaners (e.g. ultraviolet radiation, ozone) and air fresheners (containing reactive fragrances) together can be problematic as it creates reactive species which are more likely to cause adverse health effects.

An intervention study in which air cleaners were put into apartments of elderly people revealed that microvascular function improved two days after the installation of the device which reduced PM2.5 particles from 8 to 4 µg/m3. This implies that improving the quality of indoor air among high risk groups with high levels of exposure can result in improved health status.

Discussion

How can the effects of indoor pollution sources such as cooking and secondary particles from cleaning products be quantified in human populations?

- In comparison with outdoor pollution, indoor air pollution is much less well characterised, especially in the UK. In the UK, only very limited observational studies have been undertaken to characterise the indoor mixture. It was agreed that more studies (including routine monitoring analogous to outdoors) are needed.
- Indoor air studies should make use of sensor technologies including personal monitors to gain insight into the temporal variability of indoor pollution. These types of monitors are good at assessing particle number and some progress has been made towards simple, lower cost assessment of specific pollutants. However, ongoing issues of reliability, calibration and inter-study comparability were raised.
- Further research should explore a wider range of potential source products, particularly for VOCs which are embedded in a huge range of different materials in the home. Major areas include personal care products, cleaning products, soft furnishings and decorative materials.
- There is a need to address other chemical sources of less volatile organic compounds such as phthalates, and other factors such as allergens and moulds.
- In classifying sources and effects, it is was agreed that it is prudent to consider a wide range of building types and to cover the full range of socio-economic groups. There is an opportunity to explore citizen science approaches to generate data in these areas which maybe a cost-effective way of building up larger datasets. Humidity could be used as a marker for some exposures (i.e. moulds) to help characterise certain aspects of indoor air pollution.
- More could be done to consider the effects of human factors upon exposures, such the recent trend in the UK to open plan living i.e. bringing cookers into the living rooms and its influence on indoor air quality, and the increased consumption of solvent-containing products.
- Emission source data must be combined with more accurate statistics on behaviour and consumption (i.e. frequency of wood-burner, candles, personal care product etc). This could be collected by using approaches such as phone apps, although it was noted that ethical issues would need to be addressed. Information on the amounts of items purchased by consumers (e.g. personal care products/air fresheners) could provide useful data for exposure assessment.

How important is the chemical composition and size distribution of indoor particles in influencing the toxic hazard?

- Chemical composition clearly matters as does size distribution and thus needs to be characterised as far as possible. There is a large gap in observations in the UK. It simply isn't possible to say what the dominant chemical components of UK indoor air are at the moment, or indeed even what 'typical' UK indoor concentrations would be.
- The need for standard particles or standard products to aid toxicity testing studies was discussed, highlighting that toxicology studies make much use of standard particles, e.g. NIST DEP particles. The potential benefits of collecting indoor relevant particle samples to create sets of standard particles could be considered, but concerns were raised about the use of standard materials for comparison as they tend to age and change reactivity which could affect overall accuracy of the tests conducted.
- Next generation toxicity testing systems should be developed further. Toxicity testing systems should be improved going from simple in vitro cultures to organoids/organ-on-a-chip systems thereby addressing lung, brain and other target systems. These could use cells under different conditions and susceptibilities to address the range of population responses.
- Chamber studies could be used more to explore effects of indoor sources (e.g. by selective filtration), although it was recognised that these are generally useful for short term studies and not so relevant for longer-term effects. The use of chamber studies to try and identify potentially useful biomarkers

was considered an area to be further explored. Samples from the UK Human Biomonitoring Study could be used.

Can the effects of indoor exposure to particles of outdoor origin be reliably quantified? Is this actually what current epidemiological studies using outdoor air quality data are able to represent?

- It may be possible to undertake cohort studies using information on the characteristic of homes including ventilation as a way to address the importance of outdoor sources for effects indoors.
- Many of the uncertainties associated with quantifying the balance of indoor vs outdoor sourced particles arise because of the lack of speciated chemical information on particles in homes. Well-established source apportionment methods can only be applied once a well-characterised set of particulate matter sources and ambient concentrations are available.

Research question D: Are the currently available public health statistics optimal for elucidation of health effects, and are there any changes to current practice which would be beneficial?

Background

Greater availability and use of a number of environmental and public health datasets could help elucidate the health effects of various air pollutants. However, a number of these datasets are not routinely available. These include air quality data (e.g. the UK lacks a long-running record of $PM_{2.5}$ concentrations) as well as highly spatially resolved datasets that could help interpret health outcomes such as deprivation, smoking habits and other lifestyle characteristics.

Access to data depends on country-specific policy and regulatory approaches to data governance and use (e.g. in terms of privacy protection), and the systems in place for data collection and sharing. Some countries can provide useful examples to draw from for creating an appropriate data sharing infrastructure, e.g. Denmark.

Air quality data collection and monitoring has improved in the UK since the Clean Air Act (1956) which led to the establishment of a network of monitoring stations for black smoke and sulphur dioxide concentrations. Mapping historical exposures to air pollutants is important for understanding associations between air pollution and health outcomes due to evidence showing that past exposure to air pollutants has long-term effects on mortality. Where measurement records are missing, modelling can provide estimates of past exposure. It is technically possible to model exposure at high levels of spatial and temporal resolution. However, one of the current unknowns remains the extent to which modelled exposure reflects the dose. Personal exposure studies could provide further clarity on this.

The COVID-19 pandemic led to increased media and political attention to the issue of air pollution, and has enabled greater and faster data sharing through the National Health Service, the Office for National Statistics and large research consortia than ever seen before. These may have long-term positive consequences for the UK data infrastructure, and access to relevant public health statistics for air pollution research and policy.

Other structural challenges for UK air pollution monitoring include: the fact that monitoring stations are mainly configured for compliance and regulatory purposes but not to estimate the pollution where people live; and the lack of datasets on indoor pollution, which evidence suggests might be responsible for 75% of individual exposure to air pollution.

Discussion

How well can past exposure histories preceding reliable measurements be reconstructed for a range of pollutants?

- Despite the absence of reliable measurements of pollutants in the past, there are a number of ways to estimate these. One of them is to use modelling to create historical pollution maps, building on work using the EMEP4UK model for sulphur dioxide. This is facilitated by the fact that the pollution climate of the UK atmosphere has dramatically changed over time, and that there is some good past data available. Data which could be drawn from include: black smoke monitoring data to examine the relationship with e.g. particulate matter, elemental carbon, black carbon etc., the paper archive from early 20th century, and historical information, which could help perform semi-quantitative adjustments.
- The particulate matter mix has changed over time, for example with the increase and subsequent decline of diesel, and that the balance of components has therefore changed.

Are there existing cohort studies which could be extended to provide insights into the effects of air pollution exposure?

- There are a large number of cohorts available (Wellcome Trust, Imperial College London, UK Bio-bank etc.) including birth cohorts (e.g. 1946 and millennium cohorts). Initiatives like that of the Wellcome Trust bringing together longitudinal cohorts are a good opportunity to link this with air pollution data. This could be further leveraged. Many of the other cohorts available are already being exploited, but with further impetus more could be linked with air pollution data.
- Administrative cohorts/routine data could also play an important role. These provide wide coverage – though may lack individual data – but could ideally complement studies with smaller numbers/detailed individual data. An issue is that differences in age of different cohorts might influence results, and it is therefore important to adjust by age when comparing results.

Is the current temporal resolution of routine health statistics adequate for air pollution studies?

- Temporal resolution is not a problem because a) the overall research interest is in studying long-term effects rather than short-term effects b) data is recorded daily and can be aggregated easily, which makes currently available levels of temporal resolution sufficient and fit for purpose.
- One of the difficulties when handling data coming from routine sources is that sometimes there are confidentiality restrictions over the combination of having both temporal and spatial data. The main problem therefore is spatial granularity. This is often due to ethical issues regarding identification of individuals. Accessing individual primary care data (e.g. general practitioner and hospital data) is particularly problematic despite it being crucial for research.

Are there valuable statistics currently being collected which are either unavailable or not being utilised for air pollution research?

- In terms of air quality data, it would be helpful for local authorities to share their air quality data, which is not always readily available, and for more information regarding speciation of particulate matter (e.g. chemical components, ultrafine particles for which measurements stations are currently lacking in numbers) and Volatile Organic Compounds to be available.
- Access to activity and movement data could allow modelling of a person's exposure in a dynamic total movement sense. This data could be obtained from crowd sourcing – despite potential biases – or apps such as Strava, or the Population 24/7 project that aims at developing a 24-hour gridded population model of the UK.
- A potential model for sharing environmental health data could be the US NHANES. Creating a UK equivalent and providing a more systematic access to data on obesity, smoking, physical activity could be helpful. This could be collected through ad-hoc surveys, in addition to existing databases, and offer significant potential if it is linked with fine spatial information.

Are there public health datasets which could be collected (at reasonable cost) which would facilitate new studies in air pollution and health?

- Making data to help investigate inequalities (e.g. data on ethnicity, deprivation etc) more readily available could help enhance research to increase understanding of the interactions between air pollution and health inequalities. Recently, in the context of the COVID-19 pandemic, exposure to air pollution was investigated as a potential modifying factor and highlighted inequalities in exposure between various ethnic and social groups.
- There is a lack of data on exposure to air pollutants in indoor environments, in part due to the absence of a national system of monitoring in England. The English Housing Survey provides information related to buildings and occupancy but does not collect information on occupants' behaviours. This is

needed to help understand how behaviour influences exposure and could fill important gaps in understanding of how variations in indoor environments determine health outcomes.

Would a reconfiguration of the current monitoring network offer any advantages for health effects research?

- As currently configured, the monitoring network does not reflect population exposure well. This might need to be addressed if a Population Exposure Reduction Target is introduced under the Environment Bill. Making data available as population-weighted exposures would be helpful.
- There is increasing interest in looking at the effects of different particle components and size fractions and only a few sites across the country where data on these can be obtained. However, these sites, where there are direct measurements, will be crucial for validating the spatial models that will provide such a breakdown for use in health studies
- Some aspects that the monitoring network could better capture include local pollution hotspots and air pollutant concentrations in rural areas, which can experience high levels of ozone and wood smoke. However, long-term measurement stations and datasets are crucial in identifying trends, so caution should be exercised when considering reconfiguring monitoring networks.

Is there a research need to have air pollution data available at appropriate finer scales for health studies? Might there be disadvantages if the same scale and model was used for exposure assessment in all studies?

- It might be easier to differentiate the effects of different pollutants or components in studies which assess exposure at a fine spatial scale. Currently, fine scale modelling is more readily available in some locations (notably London) than others, leading to regional differences in the ability to undertake epidemiological studies. Although high resolution (metres) air quality models have been developed, they are not widely available to the health effects research community.

  While there is agreement that a centrally available resource that provides access to such data would be helpful, it should be borne in mind that a) every study is different and has specific requirements, b) using the same model or exposure data risks replicating errors. This resource should therefore be a repository and not favour the use of any specific model or data.

Annex 2 – Workshop agenda

14 Sept 2020	Session 1: Setting out the challenge (3h30)
	Chair of session 1: Prof Roy Harrison FRS
14:00 -14:20	**Welcome and introduction to the meeting (context, objectives, format)** (20 min) Prof David Fowler FRS
14:20 – 17:20	**Key science questions to be addressed by the workshop** 20-30 min talk + 10 min Q&A per speaker. Each speaker will introduce one science question as listed below. a. Can either toxicology or epidemiology give clear and consistent insights into the relative toxicity of particulate matter of differing source, composition and/or size distribution in human populations? **Dr Roel Schins, Leibniz Research Institute for Environmental Medicine, Germany** b. Are there innovations in epidemiological methods which would allow a clear differentiation of the effects of closely correlated pollutants upon public health (e.g. nitrogen dioxide and PM2.5)? **Prof Joel Schwartz, Harvard University, US** **15 min break** c. Current estimates of the health burden due to air pollutant exposure deriving from cohort and time series studies quantify only the impacts of outdoor pollution. How can the impacts of indoor sources be evaluated? **Prof Torben Sigsgaard, Institute of Public Health, Aarhus University, Denmark** d. Are the currently available public health statistics optimal for elucidation of health effects, and are there any changes to current practice which would be beneficial? **Prof Anna Hansell, University of Leicester, UK**
17:20 – 17:30	**Conclusion of session 1 and what to expect at the next session** (10 min) Prof Roy Harrison FRS

15 Sept 2020	Session 2: Identifying and discussing research priorities (3h30)
	Chair of session 2: Prof Terry Tetley
14:00 -14:10	**Welcome and objectives of session 2** (10 min) Prof Terry Tetley
14:10 – 15:40	**4 parallel breakout sessions, each addressing one of the questions introduced during session 1** (1h30) Breakout groups will receive guidance on how to structure their discussion. Each group will have a chair and a rapporteur.
15:40-15:55	**Break**
15:55-17:10	**Reporting from breakout sessions + discussion** (1h15) Prof Terry Tetley
17:10-17:30	**Conclusion of the meeting and next steps** (20 min) Prof David Fowler FRS

Annex 3 – Attendees list

Name	Affiliation
Alison Gowers (breakout group rapporteur)	Public Health England
Ally Lewis (breakout group Chair)	University of York
Anna Hansell (speaker)	University of Leicester
Ben Konnert	The Royal Society
Boudewijn Dominicus	Natural Environment Research Council
Cathryn Tonne	Instituto de Salud Global, Spain
Christopher Carlsten	University of British Columbia, Canada
David Fowler FRS	Formerly Centre for Ecology and Hydrology
David Rhodes	Public Health England
Elizabeth Surkovic	The Royal Society
Fan Chung	Imperial College London
Frances Bird	The Royal Society
Francesco Forastiere	Imperial College London
Frank Kelly (breakout group Chair)	Imperial College London
Gavin Shaddick (breakout group Chair)	University of Exeter
Graham Campbell	Medical Research Council
Heather Walton (breakout group rapporteur)	Kings College London
Helene Margue	The Royal Society
Ian Mudway (breakout group rapporteur)	Imperial College London
James Musisi	The Royal Society
Joel Schwartz (speaker)	Harvard University, US
Johan Ovrevik	Norwegian Institute of Public Health, Norway
John Newington	Department for Environment, Food & Rural Affairs, UK
Jorge Bernadino de la Serna	Imperial College London
Liza Selley	Medical Research Council
Mark Miller	The University of Edinburgh
Marta Blangiardo	Imperial College London
Martin Clift	University of Swansea
Martin Williams (breakout group Chair)	Imperial College London
Matthew Loxham	University of Southampton
Mireille Toledano	Imperial College London
Paul Wilkinson	London School of Hygiene and Tropical Medicine
Per Everhard Schwarze	Norwegian Institute of Public Health, Norway
Rachel Smith (breakout group rapporteur)	Public Health England
Rodger Duffin	University of Edinburgh
Roel Schins (speaker)	Leibniz Research Institute for Environmental Medicine, Germany
Roland Wolf	University of Dundee
Roy Harrison FRS	University of Birmingham
Shaun Brace	Department for Environment, Food & Rural Affairs, UK
Shema Bhujel	The Royal Society

Steffen Loft	University of Copenhagen, Denmark
Stephanie Wright	Imperial College London
Stephen Holgate	University of Southampton
Suzanne Bartington	University of Birmingham
Terry Tetley	Imperial College London
Tim Gant	Public Health England
Torben Sigsgaard (speaker)	Aarhus University, Denmark
Ulla Birgitte Vogel	National Research Centre for the Working Environment, Denmark
Zorana Andersen	University of Copenhagen, Denmark