The Academy of Medical Sciences

Correlational methods

Summary of a workshop conducted for the Academy of Medical Sciences working group on identifying the environmental causes of disease

June 2007

The Academy of Medical Sciences

The Academy of Medical Sciences promotes advances in medical science and campaigns to ensure these are converted into healthcare benefits for society. Our Fellows are the UK's leading medical scientists from hospitals and general practice, academia, industry and the public service.

The Academy seeks to play a pivotal role in determining the future of medical science in the UK, and the benefits that society will enjoy in years to come. We champion the UK's strengths in medical science, promote careers and capacity building, encourage the implementation of new ideas and solutions – often through novel partnerships – and help to remove barriers to progress.

Contents

Introduction

This report provides a summary of the Academy of Medical Sciences' workshop on '*Correlational (non-experimental) methods*', which took place at 10 Carlton House Terrace, London on 19 June 2007. The workshop is part of a wider evidence gathering and consultation process conducted by the Academy's working group on identifying the environmental causes of disease that published their final report in November 2007. Copies of the report are available from: **http://www.acmedsci.ac.uk/publications**

The project has been generously supported by the University Hospital Association.

The aims of the workshop were to:

- Engage and consult with stakeholders.
- Seek views on draft principles prepared by the working group.
- Consider how the principles might apply to different stakeholders groups.
- Inform the conclusions and recommendations of the working group.

The Academy is most grateful to the speakers, facilitator and discussants for their thoughtful presentations and stimulating remarks. This report was prepared by Laurie Smith, secretariat to the non-experimental methods working group.

Report of the proceedings

1. Background: Professor Sir Michael Rutter CBE FRS FBA FMedSci

Almost daily there are reports of research showing a new cause for a medical condition. However, few of these claims are confirmed by further research and occasionally new studies find the opposite. What is more, many risk factors for disease, such as toxins or environmental pollutants cannot be tested experimentally in humans because it would be unethical or impractical. Both of these challenges lead to the questions around how we judge reported causes and risk factors for disease: how should we decide what to believe?

Most medical conditions, such as cancer or heart disease, have many causes (risk factors) that operate along multiple causal pathways. When investigating the causes of disease it is therefore necessary to study components of causal influences that are somewhat influenced by chance rather than some hypothetical single basic determinative cause. In addition, many risk factors involve prior causal pathways. Thus smoking is a causal component for lung cancer that operates fairly early in the initiation of disease, but whether an individual decides to smoke will be influenced by factors such as social pressures, the availability of cigarettes, genetics and advertising.

There are numerous reasons why an observed statistical association between a putative risk factor and a disease does not imply causation. The disease may be mediated genetically rather than by the environment. Exposure to the risk factor may be due to social selection that is not random. It may be that the disease causes the risk factor, rather than the reverse.

Alternatively there may be the influence of some unknown 'third variable'. Research into the causes of disease has to be able to deal with these and other possibilities.

Some scientific methods are better suited to investigating the causes of complex diseases than others. Laboratory experiments can be inappropriate because it may be unclear which variable to manipulate and which to control. Similarly, randomised controlled trials (RCTs), are useful when investigating the efficacy of interventions, but less suitable for understanding risk factors. Animal models are a very valuable tool, although care is needed when extrapolating results to man. Natural experiments, such as twin studies, are particularly helpful because they offer the opportunity to study individual variables that normally go together. Finally, there are correlational (non-experimental) studies, where researchers make observations without manipulating variables in the hope of finding causal associations.

There are many examples where correlational research has been used to identify the causes of disease. These include the links between:

- Smoking and lung cancer.
- Sulphonamides and treatment of puerperal sepsis.
- Socioeconomic status and health outcome.
- Prenatal alcohol exposure and foetal alcohol syndrome.
- Thalidomide and abnormal foetal limb development.
- Physical/sexual abuse of children and mental disorders.

- Folic acid and neural tube defects.

In most of these instances, the outcome was rare in unexposed individuals (i.e. those who had not been exposed to smoking, sulphonamides, thalidomide etc), but there was a large excess risk in those who had been exposed.

When considering the causes of disease it is important to remember that risk is not a unitary concept. There are differences between relative risk, which compares risk in two different groups of people; absolute risk, the risk of an event over a period of time; and attributable risk, the difference in risk between those who are and are not exposed. For instance, the relative risk of a woman over the age of 40 giving birth to a child with Downs Syndrome is sixteen times that of a younger women. However, the absolute risk is about one in a hundred, because Downs Sydrome is rare even amongst older women. The attributable risk at the population level is also low because most women who have children are younger than 40, so older women make up only a small fraction of women having children overall. The different types of risk are often not well understood by professionals, patients or the public.

The correlational studies listed above have survived rigorous testing, often due to sharing the following features:

- Large effect or rare/unusual outcome.
- Detailed attention to alternative non-causal explanations.
- Use of multiple research designs, including natural experiments.
- Causal inference tested in multiple differing populations.
- Animal models and human experiments supported by biological mediation processes.

There have also been misleading findings from correlational studies, such as the purported protective effect of Hormone Replacement Therapy (HRT) against cardiovascular disease, the link between the Measles, Mumps and Rubella (MMR) vaccine and autism, and consumption of caffeine in pregnancy and low birth weight. Most of the misleading claims are based upon single, weak, small-scale pilot studies with inadequate controls and/or highly specialised samples. Moreover, misleading conclusions are most likely when there is a strong possibility of selection or indication bias. Correlational studies are not well placed to investigate small effects, although these may be important especially at the population level. Tellingly, it has been estimated that about half of all presentations at conferences never appear in peer-reviewed scientific journals, but may still be picked up by the media

During the course of their inquiries the working group has identified six overarching issues:

- No single study ever proves a general cause.
- Identified causes almost always involve components of causation.
- Context effects, such as gene-environment interaction, are important.
- Clinicians and policymakers have to act on the evidence available at the time and that requires judgements as well as ‘facts’.
- Identification of the mediating mechanisms is always helpful but it is rare to be able to delineate these early in the research process.
- All research is provisional in that it solves some problems but opens many new questions.

When acting upon correlational research, it is important to be clear that decisions

either to take action or not may both have significant consequences.

Discussion

The draft principles proposed by the working group were discussed by the participants. Some felt that the principles could be clearer and more positive in tone. It was suggested that the guidelines could be tailored differently for each of the targeted stakeholder groups. There was a strong consensus that 'non-experimental methods' was a more appropriate title for the project

It was noted that several of the examples discussed concern interventions rather than causes of disease. For instance, numerous correlational studies indicated that HRT was protective against cardiovascular disease. A subsequent RCT showed that it was a risk factor indicating that multiple studies may not stop misleading results. On the other hand, women could choose whether to take HRT. And those that did were systematically different from those who did not. The mismatch between the correlational and experimental evidence can therefore be explained by indication bias. RCTs are currently underway to determine whether HRT is protective and have not yet reported. A clear goal of the project should be to restore some of the credibility to non-experimental research in light of high profile reversals.

It must be remembered that even a single RCT rarely provides sufficient evidence to change policy and practice. This is why many such as the National Institute for Clinical and Public Health Excellence (NICE) place particular emphasis on meta-analyses. Some also advised caution since RCTs can sometimes give the right answer to the wrong question.

Over forty years ago the renowned medical statistician Sir Austin Bradford Hill FRS set out guidelines to help make the distinction between association and causation in medical research. These were clearly framed as matters of judgement. While the guidelines have stood the test of time some aspects have been more durable than others. For instance, biological plausibility seems to be less important than once thought. It was suggested that the Bradford Hill guidelines should be discussed more explicitly in the summary, conclusions, principles and recommendations of the final report.

The participants noted that the relationship between science and society has changed dramatically since the 1960s when Bradford Hill published his guidelines. Institutions, funders and researchers now seek to promote the results of their research via press offices and through the media. The public are less trusting of authority. Researchers are formally evaluated through the Research Assessment Exercise. With more universities comes greater competition for funds. All these changes justify another look at correlational research.

It was noted that caution is needed when putting medical science in the realm of belief rather than knowledge. There should be a clear distinction between what patients and the public know and what they believe. Patients and the public sometimes view causes dichotomously, something does or does not, rather than probabilistically, which is more useful for understanding complex diseases.

2. Researchers: Professor Kay-Tee Khaw CBE FMedSci

Most of our beliefs, practices and polices are based on observational (correlational) data. If it is possible to explain why something happened then it may also be possible to predict whether it will happen in the future. This allows action to be taken to alter the course of events. A key question is therefore: 'how much evidence do we need before taking action?'

There are many examples of ecological associations, those measured at the aggregate rather than individual level, between variables and health outcomes. These include:

- Per capita consumption of cigarettes in 1930 and crude death rate for lung cancer among men in 1950 in different countries.
- Per capita cigarette sales and coronary heart disease mortality in US states in 1960.
- Meat consumption and colon cancer incidence in various countries.
- Fat consumption and breast cancer mortality in different countries.
- Sunlight and breast cancer incidence in different countries.
- Pairs of nesting storks and births over time in Sweden.

In order to decide which of these associations are causal it is necessary to determine the extent to which each can be explained by chance, confounding or bias. The Bradford Hill guidelines are helpful when making such judgements.

The relationship between infant sleeping position and Sudden Infant Death Syndrome (SIDS) offers an interesting case study. From 1954 to 1988 an increasing proportion of paediatric textbooks recommended frontal sleeping for infants, possibly to avoid risk from them choking on their own vomit. In 1988 an overview of largely case-control studies, instigated because of rising rates of SIDS in much of the developed world over the previous 15 years, reported that prone sleeping carried a substantially increased risk. 'Back to Sleep' campaigns in several countries were followed by a fall in the rate of SIDS. The value of understanding the relationship between sleeping position and SIDS is illustrated by the estimated 11,000 SIDS deaths in England and Wales between 1974 and 1991 attributable to harmful health advice to sleep in the prone position.

The example of SIDS offers some important lessons:

- It usually requires a body of evidence to change health policy.
- Consistent findings from different populations with different confounding and biases are needed to demonstrate replicability.
- It is necessary to rely on natural experiments, continuing surveillance and monitoring because in practice it is never possible to be sure of the causality of complex interventions and their unexpected effects.

RCTs can minimize, but not exclude, confounding and bias. But most health questions cannot be investigated using this method for ethical or practical reasons. Although RCTs are thought to be the 'gold standard' for internal validity many question their external validity. It is therefore necessary to rely on other types of research design, with their inherent strengths and limitations. The totality of evidence from all sources should be brought together in order to reach sensible conclusions. Judgement, accompanied by quantitative assessments and sensitivity analyses of the consequences of different

options, is needed in order to decide what action or policy is justified by the available evidence.

Correlational research, as with all science, is a dynamic process. New cohorts, exposures and interactions are always needed as there is always uncertainty.

To quote Sir Austin Bradford Hill FRS:

'All scientific work is incomplete – whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have or to postpone the action that it appears to demand at a given time'

Discussion

Concern was expressed that correlational methods are sometimes used when RCTs are appropriate and possible. For example, multiple correlational studies indicated that certain vitamins reduce mortality. It was argued by some that this evidence was sufficient for action. Subsequent RCTs demonstrated that the opposite was the case: some vitamin supplements increase mortality. Nevertheless, some researchers continue to support the correlational findings.

It was suggested that the project should delineate when correlational studies and RCTs are appropriate. For instance, while RCTs are the 'gold standard' for evaluating the efficacy of interventions they are less suitable for the investigation of rare adverse events. Researchers should use the method most appropriate for the research question.

It was noted that the draft principles and recommendations currently focus mainly on internal validity. RCTs tend to be more internally valid while correlational methods tend to be more externally valid. External validity is particularly important in the social sciences and of relevance to the media, policymakers, patients and the public, although it sometimes receives insufficient attention with regard to RCTs.

Many agreed that journal editors and peer-reviewers are a key stakeholder group involved in the communication of correlational research who should be offered distinct recommendations. Peer-reviewers are sometimes selected for their expertise in a topic area rather than a methodology so may not be best placed to judge correlational research.

The point was raised that the working group should engage those primary care practitioners as these are the gatekeepers of medical knowledge for many communities. Patients and the public put great weight on the views of medical practitioners with whom they have had personal contact. Primary care practitioners need to be equipped with accurate information in order to best engage with patients.

Two factors that determine whether research is translated into policy are the strength of the evidence and timing. The knowledge required to justify causal inference and the knowledge required to justify policy decisions are different. Causal inference can be determined using scientific evidence while policy decisions are matters of judgement that are only partly informed by science. There was some concern that researchers are being asked to act as public health officials when considering the relationship between science and policy.

Research is based upon data that needs to be stored, preserved and shared. Data

sharing allows better meta-analysis and replication, which can help avoid misleading causal inferences. It was noted that free access to data can help prevent or solve cases of scientific fraud. The recently established Institute to Evaluate International Health Programs, which was funded by the Gates Foundation, has already promised to make its data publicly available. Other funders are beginning to follow suit. Public bodies often have large quantities of unused data that can be utilised for comparative advantage as is being done in Western Australia. It was argued that making datasets available would only reveal computational errors. However, in some circumstances it might also reveal differences of interpretation.

3. Policymakers: Dr Geoff Mulgan

Medical science is more uncertain than many believe. Policymakers therefore need to exercise judgement when making decisions that are informed by scientific evidence. A particular challenge is that research evidence can change over time. For instance, studies in the 1990s on California's Greater Avenues for Independence (GAIN) welfare to work programme provided strong evidence that prompted changes in policy. Nine years later the same studies demonstrated the opposite was the case. This time the results received little attention.

The use of research by policymakers can be divided into five components:

- Structures
- Processes
- Methods
- Culture
- Money

A number of structures, such as Departmental Chief Scientific Advisors and the Chief Medical Officer, exist within government to provide scientific advice to policymakers and act as the guardians of good method. Many scientific issues cut across departments but many structures do not. This challenge is exacerbated by the adoption of the Rothschild Principles that separated research from policy. Researchers should be embedded as an integral component of government policy teams.

Key processes in government include review, spending, strategy and policy. Review should be designed into evaluations of policy. Government departments, such as the Treasury, should be encouraged to own the principles set out in the working group's final report to ensure that they are taken forward.

Prior to the methodology discussed by previous speakers there exist causal maps: abstract, coherent, representations of causal relations among events. Policymakers often do not make causal maps. Understanding of causal maps, along with the need for evaluation, should be incorporated into the training of senior civil servants.

The knowledge generated by medical research is often tentative. Researchers should therefore exercise caution and not over sell their results. Scientific evidence is one of many factors that policymakers need to consider when making decisions. In a democracy this means that science may be trumped by other considerations. Time pressure is of particular importance. Researchers often operate over months or years while policymakers are sometimes required to act over hours or days. The need to act quickly can mean decisions need to be made despite great scientific uncertainty.

The policy decisions to which scientific evidence contribute often have cost implications. In some cases there may not be sufficient resource to enact a policy that has scientific backing. For instance, although there is evidence that statins reduce cardiovascular disease it would be too expensive to prescribe them to the whole population.

Researchers and policymakers are becoming increasingly aware of the gaps in the understanding of medicine. A central message is therefore the need to acknowledge that policy is informed, rather than based upon, evidence. This echoes the conclusions of the 2005 House of Commons Science and Technology Select Committee report on scientific advice in policymaking.

Discussion

Evidence generated using different methods suffer from different sorts of uncertainty. While no form of evidence is perfect some types of evidence, such as that generated by scientific method, is more certain that others, such as hunch or anecdote.

There is increasing pressure for policymaking to become more transparent. The Human Genetics Advisory Commission, Food Standards Agency and US Food and Drugs Administration have already opened many of their expert advisory meetings to the public. NICE will do the same shortly. That said, some expressed concern that such measures are only likely to engage a relatively narrow section of public. The ultimate aim is to have a scientifically literate public who would be able to critique correlational research.

In rapidly advancing fields of medicine there may be a very narrow window of opportunity in which to conduct RCTs before the technology becomes redundant. It may therefore be necessary to rely on observational data. For example, more than three quarters of new diagnostic imaging techniques have been accepted into clinical practice on the strength of observational studies. Organisations such as the Health Technology Assessment (HTA) are now looking to fast track research into key questions to allow more time for analysis.

4. Media, patient and the public: Dr Richard Horton FRCP FMedSci

The results of correlational research often become a matter of urgent public concern. On 7 June 2007, the New York Times and Wall Street Journal ran full-page adverts from GlaxoSmithKline (GSK) addressed to 'avandia patients'. The company agreed that it had 'conducted an unprecedented number of clinical trials' concerning the safety of its drug. But it also acknowledged that 'recent press coverage about the safety of avandia' had caused 'confusion and concern'. And it affirmed that 'further information regarding potential heart-related risks is currently under review by the FDA.'

The occasion for this extraordinary reaction was a fast-tracked paper in the New England Journal of Medicine by Steven Nissen and Kathy Wolski. In a meta-analysis of 42 clinical trials, they found that the odds ratio for myocardial infarction was 1.43. This result pushed them to conclude that rosiglitazone (avandia) presented a 'significant increase' in risk. An accompanying editorial questioned the whole rationale for prescribing rosiglitazone. It called for 'regulatory action' by the FDA.

GSK tipped into crisis. Its share price fell by 13%. Prescriptions declined by 20%. Newspapers raised disturbing parallels with Vioxx. Congress called for an inquiry. The FDA was accused of negligence.

The case against rosiglitazone was precarious. No single randomised trial had shown a consistent risk attached to the drug. Nissen and Wolski admitted that their result was based on limited access to trial data and a relatively small number of events. The odds ratio could have been affected by small changes in the classification of events. The trials that were examined were not intended to explore cardiovascular outcomes. Those outcomes did not have adjudicated cardiovascular events. The studies were small and short term. Uncertainty was great. Even the authors called their report 'less convincing than a large prospective trial.'

GSK joined the debate by publishing the results of a large observational study. In a managed-care database of over 30 000 patients, they reported that a composite cardiovascular endpoint was 1.75 events per 100 patient-years for rosiglitazone and 1.76 events per 100 patient-years for the non-rosiglitazone group – a very definite non-urgent, non-significant risk. These data were buttressed by interim results from an ongoing clinical trial. Here was a public war between methodologists, experimental and non-experimental.

This case-study raises questions for researchers, editors, sponsors, regulators, journalists, and patients. For researchers, what is their responsibility when considering how to present controversial findings in public? For editors, should they fast-track such papers, with all the attendant additional publicity? For sponsors, how should they conduct debates about risk? For regulators, what is the threshold for taking action when new risks are discovered? And what weight should be placed on the different types of methods that yield such risks? Journalists face particular difficulties. They are dammed if they do and damned if they don't report news. As the pivot between researchers and the public, they face extraordinary responsibilities. And finally, what are patients to do? GSK published advice for patients, as did newspaper columnists. But the picture was one of utter confusion.

Despite a history of repeated scares and confusions, lessons appear not to have been learned.

Discussion

It was argued that guidelines for journalists have a limited effect how the media reports correlational research. At the heart of the problem is an inevitable tension between the interests of editors, who wish to sell copy, and researchers, who wish to communicate their results. This matter is further complicated by the increasing reliance of patients on the internet which is less subject to central controls than traditional media.

As guidelines for journalists seem to have had a limited effect it was suggested that the wider scientific community should ensure its messages are properly communicated. Specialist science and medical journalists are likely to be a great ally of researchers in achieving this goal. Some thought lessons might be learned from the media tactics employed by many NGOs. Others noted that NGOs sometimes focus more on advocacy than research evidence. It may be that the only way forward is to offer helpful guidance that may only be used by some journalists.

It was suggested that the working group present their guidelines as the result of a

conversation amongst the scientific community about the challenges and good practice in communicating correlational research to make them more newsworthy. Journalists might then become interested in how the conclusions of the project relate to them.

Press Officers are crucial to the communication of correlational research. Five years ago few were employed in academia or science. Now most organisations have some press function. It was noted that Press Officers are frequently judged by the recognition they achieve for their host organisation rather than how accurately they communicate science. Guidelines for Press Officers may be helpful alleviating this challenges. Another suggestion was that senior scientists should review press releases for accuracy before they are issued. It might also be helpful to encourage scientists to take the time to explain the merits and demerits of particular studies to satisfy the desire of journalists for information. Funders too could have an important role to play by making accurate communication of results a condition of grants.

Peer-review currently principally considers the scientific merit of research rather than its wider social impact. It was therefore suggested that journals should be more cautious when publishing research that is likely to be of great public interest. Increasingly journals are fast tracking papers. On one hand, this may be helpful by ensuring evidence becomes available more quickly. One the other, it may mean research is less subject to scrutiny.

It was mentioned that the project should offer principles that apply directly to the patients or the public that might help them ask the right questions. Interpretive tools and risk assessment techniques might also be helpful.

Appendix I programme

19 June 2007

10 Carlton House Terrace, London SW1Y 5AH

13:00	Lunch	
13:30	Welcome	Dr Geoff Watts FMedSci, Freelance Science and Medical Journalist
13:35	Session 1: Introduction	Sir Michael Rutter CBE FRS FBA FMedSci, Academy of Medical Sciences
13:45	Initial questions	Led by contributions from: Sir Michael Rutter CBE FRS FBA FMedSci
13:50	Session 2: Principles as applied to researchers	Professor Kay-Tee Khaw CBE FMedSci, University of Cambridge
14:00	Discussion	Facilitated by Dr Geoff Watts FMedSci
14:40	Session 3: Principles as applied to policymakers and funders	Dr Geoff Mulgan, The Young Foundation
14:50	Discussion	Facilitated by Dr Geoff Watts FMedSci
15:30	Tea/ Coffee	
15:50	Session 4: Principles as applied to the media, patients and the public	Dr Richard Horton FRCP FMedSci, The Lancet
16:00	Discussion	Facilitated by Dr Geoff Watts FMedSci
16:40	Concluding remarks	Sir Michael Rutter CBE FRS FBA FMedSci
16:55	Thanks	Dr Geoff Watts FMedSci
17:00	Close	

The Academy are most grateful for the **University Hospital Association's** support of this project.

Appendix II workshop delegates

Susan Barber, Cancer Research UK
Sally Brearley, Patient's Forum
Tracey Browne, Sense About Science
Professor Nancy Cartwright FBA, London School of Economics
Dr Lee-Ann Coleman, British Library
Dr Peter Coleman, Stroke Association
Professor Rory Collins FMedSci, University of Oxford
Professor Sally Davies FMedSci, Department of Health
Professor Philip Dawid, University College London
Simon Denegri, Association of Medical Research Charities
Professor Adrian Dixon FMedSci, Royal College of Radiologists
Professor Pat Doyle, The London School of Hygiene and Tropical Medicine
Dr Alan Doyle, Wellcome Trust
Fiona Fox, Science Media Centre
Dr Fiona Godlee, British Medical Journal
Dr Trish Groves, British Medical Journal
Professor Graham Hart, University College London
Dr Aroon Hingorani, University College London
Professor Dave Leon, London School of Hygiene and Tropical Medicine
Dr Georgie MacArthur, Academy of Medical Sciences
Mary Manning, Academy of Medical Sciences
Dr Peter Marsh, Social Issues Research Council
Professor Klim McPherson FMedSci, University of Oxford
Professor Jonathan Meakins, University of Oxford
Dr Liz Miller, Health Protection Agency
Dr Helen Munn, Academy of Medical Sciences
Professor Catherine Peckham CBE FMedSci, Institute for Child Health
Katrina Nevin-Ridley, Wellcome Trust
Sir Michael Rawlins FMedSci, National Institute for Clinical and Public Health Excellence
Professor Barnaby Reeves, University of Bristol
Dr Andrew Russell, Association for Spina Bifida and Hydrocephalus
Laurie Smith, Academy of Medical Sciences
Roger Steel, INVOLVE
Hazel Thornton, University of Leicester
Professor Peter Weissberg, British Heart Foundation

Academy of Medical Sciences
10 Carlton House Terrace
London, SW1Y 5AH
Tel: +44(0)20 7969 5288
Fax: +44(0)20 7969 5298

Email: info@acmedsci.ac.uk
Web: www.acmedsci.ac.uk
Registered Charity No. 1070618
Registered Company No. 3520281