A spline for the time

Joel Schwartz

'Nowhere in the Bible does God say the Laws of Nature must be linear' Enrico Fermi

Indeed, it seems almost necessary that some nonlinearities exist. Most biological processes are under feedback control, for example, which generally implies nonlinearity. The limitation of probabilities to range between zero and one essentially implies S shaped curves. So nonlinear relations seem a fact of life. Of course any continuous curve is well approximated by a straight line within a neighbourhood. Hence, if the effect is not too large, and the range of exposure is likewise limited, linear dose-response relations can be observed even when the underlying phenomenon is not linear.

When epidemiologists primarily dealt with exposure variables such as whether or not one got one's drinking water from the Southwark and Vauxhall water company, such issues hardly mattered. Today, one is more likely to be examining a continuous predictor, such as the role of blood pressure or low-density lipoprotein (LDL) cholesterol in predicting myocardial infarctions (MIs). How should these predictors be handled?

A reasonable first question is does it matter? The answer is clearly yes. Public health interventions must be lined to the shape of the dose-response curve to be effective. For example, figure 1 shows a penalised cubic spline curve I fit to model the difference from expected birth weight in 400 000 live births in Massachusetts versus the number of cigarettes smoked per day by the mother during pregnancy. If the nonlinear association observed here is confirmed, it would imply that telling smoking mothers to cut back to half a pack a day or less would accomplish little. Most of the benefit derives from reductions from 10 or fewer cigarettes per day to zero. Again, the Tobacco industry long argued that the association of secondhand tobacco smoke with MIs was implausible because the slope of the dose-response curve was much higher than for active smoking.

Correspondence to Joel Schwartz, Harvard School of Public Health, 401 Park Drive, P.O. Box 15677, Suite 415L West, Boston 02215, USA; jschwrtz@hsph.harvard.edu

Recently, Pope *et al*[1] showed that there was a continuous, nonlinear association across a wide range of exposures to smoke particles starting, with ambient air pollution at the low end, and concluding with heavy smoking at the high end. The association became less steep with increasing dose across all three exposures (ambient particles, secondhand smoke, active smoke).

Failure to deal with potential nonlinearities could completely obscure an association as well. The association between today's temperature and today's deaths in urban areas is U shaped, with higher deaths on both very cold and very hot days. The best linear association could well have a slope of zero in some locations.

In light of this, one is left with the question of how. One common approach is to divide the range of exposure into quarters or fifths and use dummy variables for each level. Such an approach is capable of capturing nonlinearity, including U shaped curves. It is sometimes argued that it does so without making any assumptions about the shape of the dose-response curve. That is not quite true. It does not make any assumption of the size or even direction of the jump between categories, but it makes a strong assumption about the 'shape' of the dose-response curve. It assumes it is a step function, with no association between exposure and response within category, and sudden jumps between categories. And failure to allow for that within category association, between exposure and response, reduces power if there is a true effect. Is it possible to do better? One obvious alternative, endorsed in the paper of Castaldi *et al*,[2] is to use a piecewise linear curve, instead of a piecewise constant curve. Piecewise fit such as these are called splines in statistics. Figure 2 shows an example of these alternatives. In this instance, data has been taken from the US First National Health and Nutrition Examination Survey, a probabilistic sample of the US population. The figure shows two alternative fits of systolic blood pressure to age. The first modelled age as quarters, and shows some signs of nonlinearity, in that the jump between the first and second quarters is smaller than the subsequent jumps. This is a piecewise constant model. The second fit also divides the age range, in this case at the median, and fits a piecewise linear model. This model captures increases in blood pressure within category of age, which is now believed to occur, and uses one fewer degree of freedom to boot. Hence, the approach taken by Castaldi *et al*[2] in this issue is a reasonable approach to potentially nonlinear dose-response. Further, it is reasonable to speculate that a gene-environment interaction may be more important in one part of the nonlinear dose response curve, and again, by missing this, one may obscure important relations. This is indeed what they report.

Nonlinear dose response curves have an uncomfortable relation with interactions. To see this imagine two scenarios. In both cases, the decline of lung function with age is being examined, and the distribution of ages in this sample differs by sex. If in truth, there is a linear decline in first second forced expired volume (FEV_1) with age, which differs between men and women, and the prevalence of women in the sample increases with age, then an

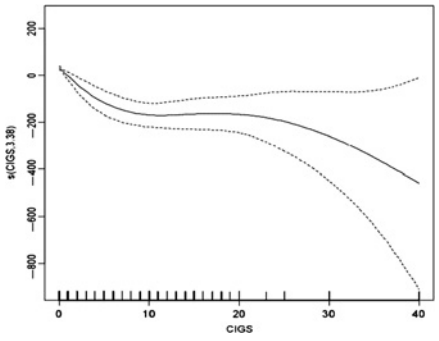

Figure 1 A penalised cubic spline curve modelling the difference from expected birth weight in 400 000 live births in Massachusetts versus the number of cigarettes smoked per day by the mother during pregnancy

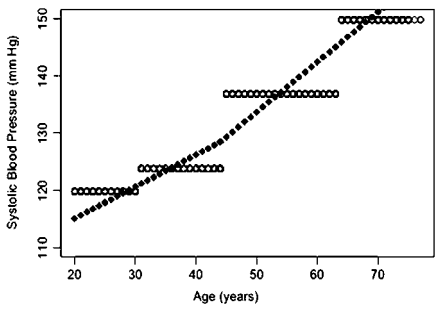

Figure 2 Two alternative fits of systolic blood pressure to age. Data has been taken from the US First National Health and Nutrition Examination Survey.

Thorax: first published as 10.1136/thx.2010.154195 on 26 May 2011. Downloaded from http://thorax.bmj.com/ on April 24, 2024 by guest. Protected by copyright.

analysis of the full sample will seem to reveal a changing slope with increasing age—that is, a nonlinear age association. Here an interaction masquerades as a nonlinearity. Alternatively, if the true association with age is nonlinear, but the same among men and women, an analysis with a linear age term and an interaction by sex will likely find a significant interaction. This has led many to believe nonlinear modelling is a bridge too far. However, the problem in the above example is that the exposure (age) is associated with the effect modifier (sex). In the paper of Castaldi *et al*[2] they have been careful to examine whether smoking was associated with the genotype they studied as a modifier, and report no association. And for gene-environment interactions this is less likely than for other modifiers. Moreover, there are ways to deal with the association. In the above example, for instance, one could test whether the association with age was not linear within sex.

In modern statistical software, testing for nonlinearity, and fitting dose-response models that do not assume constant slopes is straightforward. It is time to stop making a default assumption that continuous predictors have linear associations with health outcomes.

Competing interests None.

Provenance and peer review Commissioned; not externally peer reviewed.

Accepted 20 December 2010
Published Online First 26 May 2011

Thorax 2011;**66**:841–842.
doi:10.1136/thx.2010.154195

REFERENCES

1. **Pope CA3rd,** Burnett RT, Krewski D, *et al*. Cardiovascular mortality and exposure to airborne fine particulate matter and cigarette smoke: shape of the exposure-response relationship. *Circulation* 2009;**120**:941–8.
2. **Castaldi PJ,** Demeo DL, Hersh CP, *et al*. Impact of non-linear smoking effects on the identification of gene-by-smoking interactions in COPD genetics studies. *Thorax* 2011;**66**:903–9.

Preventing adolescents' uptake of smoking

Tim Coleman,[1] Linda Bauld[2]

Smoking is the principal preventable cause of ill health worldwide.[1] It not only affects smokers themselves but is also extremely harmful to non-smokers who inhale environmental tobacco smoke (ETS).[2] In non-smoking adults, ETS exposure causes lung and other cancers, ischaemic heart disease, stroke, chronic obstructive pulmonary disease, asthma and other respiratory illnesses.[3] Perhaps less well known by the general public is the threat that ETS inhalation poses for infants and children; there are strong associations with children's asthma, lower respiratory tract infection, sudden infant death syndrome, middle ear infection and bacterial meningitis.[3 4] Knowledge of the health risks of ETS exposure have led many countries, including the UK, to introduce laws that prohibit smoking in indoor public spaces like bars and pubs (smoke-free legislation).[5] Advocates of smoke-free laws successfully argued for these on the basis of protecting non-smokers who might work in or visit smoky environments. Consequently, as tobacco smoke has been completely eliminated from most situations in which adult non-smokers might encounter it, they are well protected from ETS. No similar protection exists for the children of smokers. Most of children's exposure to tobacco smoke occurs domestically[3 4] but, internationally, there are no smoke-free laws which forbid smokers from 'lighting up' in their homes when children are present. Children's domestic ETS exposure therefore remains an important public health concern which, as Leonardi-Bee and colleagues[6] show in this issue of *Thorax*, is even more harmful than was previously thought. The authors show that children's exposure to ETS from parental smoking has a pervasive inter-generational behaviour-modelling effect such that the children of smokers are much more likely to become smokers themselves. Their systematic review and meta-analysis collated findings from 58 epidemiological studies investigating associations between parental smoking and the subsequent development of established smoking in offspring. When both parents smoke, the risk of their children becoming smokers almost triples; if only one parent or a sibling smokes the risk is lower but, even then, children are between 1.75 times and twice as likely to become smokers than those not exposed to parents' or siblings' smoking. The consistency of findings from individual studies comprising the review is striking; virtually all component studies reported a positive association between parental/sibling smoking and children's subsequent uptake of the habit, lending strong support to the authors' conclusion that associations are probably causal. Of course, smoking is strongly associated with household psychosocial problems[7 8] and it is possible that, for some young people, these issues have more of an impact on their future smoking behaviour than parental smoking itself. Nevertheless, it seems very unlikely that the normalising impact of persistent parental smoking within the home would have no effect. Parental smoking in the home therefore has direct, substantial and immediate impacts on children's health from inhaled ETS and also, in those children who become adult smokers as a consequence of learned smoking behaviour, it has serious longer-term indirect effects mediated by their future smoking. Leonardi-Bee and colleagues call for 'radical changes in public policy and behaviour and in the acceptability of smoking in places where children are present'. However, while arguments for eliminating smoking in the presence of children are compelling, quite how this could be achieved remains unclear.

Legislative changes that curtail widespread behaviours need both robust public support (to ensure that new laws are obeyed) and effective compliance mechanisms (to ensure that breaking new laws has a reasonable chance of incurring penalties). There was strong public support for smoke-free laws before these were introduced in the UK and public

[1]Division of Primary Care, UK Centre for Tobacco Control Studies, University of Nottingham Medical School, Queen's Medical Centre, Nottingham, UK; [2]Stirling Management School, UK Centre for Tobacco Control Studies, University of Stirling, Stirling, UK

Correspondence to Tim Coleman, University of Nottingham, School of Community Health Sciences, University Hospital, Nottingham NG7 2UH, UK; tim.coleman@nottingham.ac.uk

Thorax: first published as 10.1136/thx.2010.154195 on 26 May 2011. Downloaded from http://thorax.bmj.com/ on April 24, 2024 by guest. Protected by copyright.