

Epidemiology

# Smoothing is soothing, and splines are fine

K Steenland

Commentary on the paper by Eisen *et al*  
(*Occup Environ Med*, October 2004)\*

Eisen and colleagues have provided a good example of the use of smoothing splines in a thorough analysis of exposure-response data, for a study of lung cancer in relation to silica exposure.<sup>1</sup> Exposure-response data are increasingly important for two reasons.

First, as noted by Bradford Hill, a positive exposure-response provides support for a causal interpretation of an association. In the case of silica and lung cancer, evidence of a positive-exposure response in several studies has provided important support for the original 1997 IARC judgement that silica is a class I (definitive) carcinogen. That judgement has remained controversial because in some studies the exposed population has not had a higher lung cancer rate than the non-exposed comparison group. Some have argued that this may be because the surface properties of silica change in different settings and may have different toxicities, so that in some cases silica may not increase lung cancer risk. However, the explanation may simply be that in some cohorts there were not enough highly exposed subjects. Our own exposure-response analysis of 10 silica exposed cohorts (60 000 workers) indicated that indeed there is a positive exposure-response for silica but that the increase in risk is seen primarily at higher exposures, and the overall slope of the exposure-response curve is relatively low compared to classic lung carcinogens such as nickel and asbestos.<sup>5</sup> This relatively low slope may be the reason why it has been difficult to show that silica does indeed cause lung cancer.

Second, exposure-response data provide necessary data for regulators to conduct quantitative risk assessment. Regulators want to know the amount of excess risk incurred due to exposure at different levels, and only good exposure-response data can answer this

\*Eisen EA, Agalliu I, Thurston SW, *et al*. Smoothing in occupational cohort studies: an illustration based on penalised splines. *Occup Environ Med* 2004;**61**:854–60.

question. For example, the US OSHA typically sets limits based on a level of exposure over a working lifetime which permits an excess risk of at most 1 per 1000 (0.1%) above background risk. The background lifetime risk of lung cancer is about 5%. OSHA would therefore seek a permissible limit for silica exposure which would allow a lifetime risk of no greater than 5.1%. The current limit for silica exposure is 0.1 mg/m<sup>3</sup>. It is clear from exposure-response analyses of silica, including the paper by Eisen and colleagues,<sup>1</sup> that the current limit is too permissive. Several analyses indicate that the lifetime risk of lung cancer after 40 years of exposure at the standard results in an excess risk of the order of 1–2% rather than the goal of 0.1%.

Splines and other types of smoothing functions are a middle ground between traditional categorical analyses of exposure, which avoid any parametric assumptions and let the “data speak for itself”, and the also traditional parametric analysis which uses exposure as a continuous variable in a model in which—sometimes without realising it—the investigator imposes a shape on the exposure-response curve. For example, in using logistic or Cox regression, the investigator is assuming the log of the rate ratio is a linear function of exposure. Yet this assumption requires justification, as this model may not fit the data, and a thorough search for the best model, with the best fit to the data, needs to be conducted.

Categorical analyses have their own limitations. The investigator must choose the number and placement of the cutpoints defining the categories, which may be arbitrary and which may heavily influence the apparent “shape” of the exposure-response. Furthermore within each category a categorical analysis assumes that there is a single exposure effect—that is, the rate ratio is constant across the exposure category, an obviously false assumption when the category is reasonably wide.

Smoothing functions do not impose a particular form of a simple parametric

model on the data, yet avoid some of the pitfalls of categorical analysis by being less dependent on the choice of cut-points and by providing a continuous curve which is not a step function. They are primarily useful graphically, for seeing the shape of the exposure-response curve. The shape of the curve may help provide a hint for choosing the best simple parametric model which will provide a concise summary of the exposure-response and be useful for quantitative risk assessment. One type of smoothing function, splines, may themselves be used for quantitative risk assessment, because they permit a quantitative estimation of risk for any specific level of exposure.

The idea of smoothing functions stems from using a simple moving average of “y” across local regions of “x”, often a weighted average in which the centre points in the region have more weight than the outermost points, and the average is calculated for one region after another as one moves across the x-axis. This produces a smooth curve in which the investigator imposes minimal constraints on the shape of the curve. Such curves have a long history, including, for example, the common moving average of the stock market calculated across time. These are non-parametric curves, in that there is no simple function with a few parameters which can summarise the curve.

Splines are an extension of this idea in which a regression of “y” on “x” is carried out in each local region as one moves across the “x” axis. Cubic splines are one common type of spline in which the effect measure (for example, the log of the rate ratio) is regressed on a cubic function of exposure (the x-axis), across several different regions or categories of exposure, spanning the entire range of exposure. A single smooth curve across these regions is then produced. Penalised splines, as used by Eisen and colleagues,<sup>1</sup> are another variant of splines in which there is a penalty for rapid change in slope of the curve in any given region of the x-axis. At this point it is not clear whether they offer any particular advantage over more traditional cubic or quadratic splines. The details of the difference between different types of spline functions need not overly concern investigators, as long as they understand the basic idea. Essentially the software increasingly makes spline functions available to investigators, although the epidemiologist may need a statistician’s help for such programming.

One important point of the analysis by Eisen *et al* is the influence of outlier observations, in this case the influence of two non-cases with very high

exposure values. These two controls resulted in the downward shape of the exposure-response curve in the highest regions of exposure. Eisen *et al* analyse their data with and without these two outliers. Without them, the curve tends to continue to increase at the highest exposures. They note that the shape of the curve at the low and medium dose region does not really change whether the outliers are included or not. This is important because it is this relatively low and medium dose region that in practice is of importance to risk assessors. It is quite likely that measurement error is greater in such extreme high dose regions where there are little data. An alternative approach, also illustrated by Eisen *et al*, is to consider the log of exposure rather than exposure itself. Taking the logs tend to reduce the influence of the highest exposures. A log transformation of exposure tends to

result in curves in which the rate ratio tends to stop increasing or plateau at the highest exposures, a phenomenon which seems consistent with data observed for a large number of occupational carcinogens.<sup>4</sup> There are a number of plausible reasons for such a plateau, including mismeasurement at highest exposures, an exhaustion of susceptibles at high exposure, and saturation of biological pathways.

Clearly it is incumbent on epidemiologists to collect as good exposure data as possible. But then the job is not over. We must use the rich exposure data to the fullest in our exposure-response analyses, and new analysis techniques have become available for this.<sup>2 3 6</sup>

*Occup Environ Med* 2005;**62**:141–142.  
doi: 10.1136/oem.2004.017194

Correspondence to: Dr K Steenland, Emory University, Rollins School of Public Health,

1518 Clifton Road, Atlanta, GA 30322, USA;  
nsteenl@sph.emory.edu

## REFERENCES

- 1 Eisen EA, Agalliu I, Thurston SW, *et al*. Smoothing in occupational cohort studies: an illustration based on penalised splines. *Occup Environ Med* 2004;**61**:854–60.
- 2 Harrell F, Lee K, Pollock B. Regression models in clinical studies: determining relationships between predictors and response. *J Natl Cancer Inst* 1988;**80**:1198–202.
- 3 Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology* 1995;**6**:356–65.
- 4 Stayner L, Steenland K, Dosemeci M, *et al*. Attenuation of exposure-response curves in occupational cohort studies at high exposure levels. *Scand J Work Environ Health* 2003;**29**:317–24.
- 5 Steenland K, Mannetje A, Boffetta P, *et al*. Pooled exposure-response and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an IARC multi-centric study. *Cancer Causes Control* 2001;**12**:773–84.
- 6 Steenland K, Deddens J. A practical guide to exposure-response analyses and risk assessment in occupational epidemiology. *Epidemiology* 2004;**15**:63–70.

## Gulf War syndrome

# Gulf War illnesses ... the story is still alive and well

N Greenberg

Commentary on the paper by Kelsall *et al*  
(*Occup Environ Med*, December 2004)\*

The story of Gulf War Syndrome (GWS) just won't go away. It's healthier now than ever. The controversy about whether a Gulf War specific syndrome exists and what may have caused it is one of claim and counter claim. In the UK the Ministry of Defence (MOD) has always strongly denied that anything amounting to a "Gulf War syndrome" exists, though it has accepted that some veterans have become ill and have been paid benefits accordingly. However, disgruntled veterans of the Gulf War have had a recent boost to their battered morale when a recent privately funded inquiry in the UK heard evidence in relation to the subject. The witnesses included scientists, senior military commanders, and veterans. Although the Law Lord who is presiding over the inquiry has yet to

publish his report, it is hard to see how the findings will not fan the flames rather than douse the fire. Another source of flammable material in the debate was the recent report by the US Research Advisory Committee on Gulf War veterans' illnesses which concluded that "a substantial proportion of Gulf War veterans are ill with multisymptom conditions not explained by wartime stress or psychiatric illness".<sup>1</sup>

There can be no doubt that both physical and psychological ill health is associated with military service during the 1991 Gulf War. In common with studies from the UK and US,<sup>2–4</sup> the recent *OEM* paper by Kelsall and colleagues<sup>5</sup> shows that Gulf War veterans report a myriad of symptoms 2–3 times more commonly than service personnel that did not deploy. We also know that deploying to the Gulf War can impair quality of life and health perception. However, one of the key remaining unanswered questions concerns the issue of causation.

Kelsall and colleagues' Australian cohort report experiencing numerous

exposures that are associated with a high prevalence of self-reported health symptoms.<sup>5</sup> Other studies have found similar associations.<sup>2 3</sup> Those who claim that GWS exists implicate numerous environmental, military, and psychological stressors as being the "toxic" agents. Chemical weapons, depleted uranium, and multiple immunisations have all been put forward as causative agents of long term damage to the health of military personnel. Clearly if one could link war zone exposures to subsequent ill health, this would add significant weight to the argument in favour of GWS. There are, however, substantial difficulties in accepting that the reported associations are wholly valid.

Firstly, most of the studies that have examined ill health in GW veterans have been carried out many years after the end of the war. As well as the basic problem of remembering facts from an event many years ago, such memories are likely to be subject to recall bias. There is good evidence to suggest that when recalling hazards experienced during previous events, that recall is heavily influenced by current self-rated health perception. In other words, if you feel unwell when you are asked about possible exposures some years ago, you are more likely to recall that you were indeed exposed to suspected "toxins" than if you currently feel well.<sup>6</sup> Unfortunately, gaining objective evidence of exposures is difficult to say the least. For example, when a group of UK service personnel who reported receiving vaccines in preparation for deployment to the Gulf War, but did not in fact deploy, had their records examined some years later, most showed that no vaccines of any sort were in fact administered.<sup>7</sup>

\*Kelsall HL, Sim MR, Forbes AB, *et al*. Symptoms and medical conditions in Australian veterans of the 1991 Gulf War: relation to immunisations and other Gulf War exposures. *Occup Environ Med* 2004;**61**:1006–13.

Staying on the issue of exposures, the Kelsall *et al* study finds, like others, that thinking that one has been exposed to chemical weapons is correlated with reporting symptoms of ill health.<sup>5</sup> However, a series of expert review committees and panels have failed to be convinced by the suggestion of widespread use of chemical weapons during the Gulf War. If they are indeed correct, then how can we marry the fact that many people believe that they were exposed to chemical weapons when in fact they were not even used? Another anomaly is the Canadian experience using pyridostigmine bromide (PB) tablets, a prophylactic agent against some forms of chemical weapons and another “toxic” candidates for GWS. Two of the three ships they sent to the Gulf used PB prophylaxis, and one did not. Yet rates of illness were identical between the three ships.<sup>6</sup> There are numerous other examples of exposures linking to the development of illnesses that simply do not make sense.

There is another, rather more important bias that works to muddy the already murky waters yet further: participant bias. As Kelsall’s group clearly points out, those who went to the Gulf have a clearly defined event in their mind to which they can attribute the onset of their symptoms.<sup>5</sup> As a rule people search for meaning, and for many unwell veterans their Gulf service serves as the focus for their symptoms. Media fuelled speculation may have reinforced this belief, but once established, the belief is not prone to change. Indeed Hotopf and colleagues<sup>9</sup> followed up a cohort of people who

believed they had GWS, and found that four years later they remain unwell, although interestingly the health gap between the Gulf and the control (Era) group was narrowing. There can be no doubt that whatever the reason, a substantial number of those who went to the Gulf report poor health many years after the end of the war.

If, as it appears, psychological mechanisms may be responsible for the mismatch of perceived health status and actual physical functioning, then it might seem logical to assume that psychological factors could be to blame. Many studies, including that by Kelsall’s group, find an increase in the prevalence of post-traumatic stress disorder (PTSD) in Gulf veterans. However when the absolute risks are examined, the resultant attributable risk remains small—just 3% in Kelsall’s cohort.<sup>5</sup> So although clearly relevant, PTSD alone cannot explain the increase in reported ill health.

It is clear that the extensive research into GWS has shown conclusively that a substantial number of GW veterans report feeling unwell. This finding has been replicated by scientists in numerous countries and cannot be doubted. However, research aiming to identify the elusive causative factor or factors has yet to deliver. Although no doubt there will be many more publications examining the health of those who served in the Gulf War, given the difficulties in verifying exposures and the difficulty in conducting research so many years after the end of the war, it is hard to see how this complicated jigsaw will ever be completed. Perhaps it is now time to examine

what can be done to assist those who are suffering, and to continue to ensure that future conflicts leave those who serve without a legacy of ill health.

*Occup Environ Med* 2005;**62**:142–143.  
doi: 10.1136/oem.2004.017855

Correspondence to: Dr N Greenberg, Surgeon Lieutenant Commander Royal Navy, King’s Centre for Military Health Research, Academic Dept of Psychological Medicine, GKT School of Medicine & Institute of Psychiatry, Weston Education Centre, Cutcombe Road, London SE5 9RJ, UK; sososanta@aol.com

Competing interests: none declared

## REFERENCES

- 1 Shane S. Chemicals sickened ‘91 Gulf War veterans, latest study finds. *New York Times* 15 October, 2004.
- 2 Unwin C, Blatchley N, Coker W, *et al*. The health of United Kingdom servicemen who served in the Persian Gulf War. *Lancet* 1999;**353**:169–78.
- 3 The Iowa Persian Gulf Study Group. Self-reported illness and health status among Persian Gulf War veterans: a population-based study. *JAMA* 1997;**277**:238–45.
- 4 Barrett DGGC, Doebbeling BN, Clauw DJ, *et al*. Prevalence of symptoms and symptom-based conditions among Gulf War Veterans: current status of research findings. *Epidemiol Rev* 2003;**24**:218–27.
- 5 Kelsall HL, Sim MR, Forbes AB, *et al*. Symptoms and medical conditions in Australian veterans of the 1991 Gulf War: relation to immunisations and other Gulf War exposures. *Occup Environ Med* 2004;**61**:1006–13.
- 6 Wessely S, Unwin C, Hotopf M, *et al*. Is recall of military hazards stable over time? Evidence from the Gulf War. *Br J Psychiatry* 2003;**183**:314–22.
- 7 Greenberg N, Iversen A, Hull L, *et al*. Vaccination records in Gulf War veterans. *J Occup Environ Med* 2003;**45**:219.
- 8 Anon. *Health study of Canadian Forces personnel involved in the 1991 conflict in the Persian Gulf*. Ottawa: Goss Gilroy Inc, 1998.
- 9 Hotopf M, David AS, Hull L, *et al*. Gulf war illness—better, worse, or just the same? A cohort study. *BMJ* 2003;**327**:1370.

## Ergonomics

# Exposure assessment in ergonomic epidemiology: is there something specific to the assessment of biomechanical exposures?

A Leclerc

Commentary on the papers by Heinrich *et al* (*Occup Environ Med*, December 2004)\* and Svendsen *et al* (*Occup Environ Med*, January 2005)\*\*

In recent issues of *OEM*, the authors of two articles in “ergonomic epidemiology” stress several necessary qualities of exposure data: they must be accurate and precise;<sup>1</sup> and the method of expo-

sure measurement must be reliable.<sup>2</sup> All epidemiologists in occupational epidemiology would agree with that: absence of systematic or random error is important, stability of the measure if repeated

under identical conditions is important too. Among the expected qualities of exposure data one could add “relevant”; we expect that the exposure data are consistent with what is known (or suspected) about the mechanisms underlying their effect on disease. This is less obvious than it seems; for example, in many situations one can wonder whether the relevant exposure is that of today, or that of last week, or that of 20 years ago, or the cumulative exposure over the last 20 years. Another quality (or limit) has to do with feasibility. If exposure assessment, at an individual level, is very expensive (in terms of money or time), alternative solutions have to be found if the study sample is large.

All this is common to all the fields of occupational epidemiology. However, is there something specific to the assessment of biomechanical exposures? Are the problems met when recording the time spent with the upper arm elevated above 90° very different from those met when assessing the level of past

exposure to (for example) formaldehyde or electromagnetic fields, considered as potential carcinogens?

At least the history of research in these domains is different: a decade ago, in the field of occupational risk factors for musculoskeletal disorders, some questions were not discussed: there was no debate on how to use the expertise of ergonomists or specialists in biomechanics for epidemiological studies; and there was almost no research activity on other methodological aspects such as comparisons between questionnaires and direct observation. These questions were almost absent from the first PREMUS (Prevention of Work-related Musculoskeletal disorders) conference in 1992 in Stockholm. One exception was an abstract by Winkel and colleagues from the MUSIC project, who had evaluated a questionnaire estimating physical workload.<sup>3</sup> The situation was very different in occupational cancer research. The concept of JEM (job exposure matrix) covering an array of chemical substances, job titles, and industries, was described as early as 1980.<sup>4</sup> Ten years ago cancer epidemiologists applied specific methods for sampling and analysis of exposure data.<sup>5</sup> In the presentation of the results of a European concerted action on the retrospective evaluation of occupational exposures in epidemiology in 1993, Goldberg and Hémon gave more than 50 references discussing the measurement of occupational exposures in epidemiological studies.<sup>6</sup> A reason for that early development of methods for exposure assessment could be that, for cancer, difficulties dealing with (past) assessment of exposure were obvious, which motivated occupational hygienists, epidemiologists, and also biostatisticians to find adapted solutions.

For biomechanical exposures, the (false) idea that it was enough to use the tools provided by ergonomists and specialists in biomechanics was probably widespread. However, in the last 10 years, there has been a remarkable development of research on methodological aspects of assessment of exposure to biomechanical risk factors. Several subjects have been widely studied and discussed: the limits of using job titles to assess exposure<sup>7</sup> and

\*Heinrich J, Blatter BM, Bongers PM. A comparison of methods for the assessment of postural load and duration of computer use. *Occup Environ Med* 2004;**61**:1027–31.

\*\*Svendsen SW, Mathiassen SE, Bonde JP. Task based exposure assessment in ergonomic epidemiology: a study of upper arm elevation in the jobs of machinists, car mechanics, and house painters. *Occup Environ Med* 2005;**62**:18–26.

various other aspects of variability of exposure measures,<sup>1</sup> the validity of questionnaires versus observation or measures,<sup>2</sup> the use of aggregated measures,<sup>8</sup> the retrospective assessment of exposure,<sup>9</sup> and others such as bias due to the presence of pain.

The two articles published recently in *OEM* bring interesting results in this field of research: Heinrich and colleagues<sup>2</sup> compared, in a study population of 87 computer workers, questionnaire data about exposure to postural load and duration of computer use with an observation of the workstation design and posture by a trained observer. They found a low agreement between the two approaches. However, they raise the question of whether the “gold standard” for postural load is the observation technique, since observation is based on short periods of time, and being observed might modify the posture.

Svendsen and colleagues<sup>1</sup> explored an attractive approach to the assessment of upper arm elevation above 90° in machinists, car mechanics, and painters. Workers filled diaries with approximately 10 preprint tasks, and exposure was evaluated using a TEM (task exposure matrix). The diary worked well, and could be used to take into account the variability of tasks between subjects in the same job, which is a positive result. However, there was considerable within-task variability; in addition the exposure contrast between tasks was relatively small in these jobs. In this situation, spending resources on obtaining task information does not seem to be the optimal strategy.

These two studies suggest important directions for future research. The first is the development of simple techniques for assessing biomechanical exposures. There is still no consensus about the validity of questionnaire data to assess postural load.<sup>2</sup> Improving the questionnaires by adding pictures to the questions is probably a good suggestion. It is also necessary to think about the qualities expected from a tool (such as a questionnaire) in the context of epidemiological studies. The most important qualities are not necessarily the same in epidemiology, ergonomics, or biomechanics.

The second direction deals with indirect measures such as JEMs (job exposure matrices) or TEMs (task exposure matrices). The main limitations in the development of matrices for biomechanical exposures is the variability of exposures within a job and within a task. However, this is probably the case more for some risk factors than for others. In large scale studies, JEMs might be useful to evaluate levels of exposure among all the subjects in a first step, possibly in

combination with more precise assessments in specific subsamples.

In addition, adapted statistical methods should be used and developed. Multilevel models should be more widely used.<sup>10</sup> They are adapted to situations where some data were collected at a group level, and also for repeated measures.<sup>11</sup> Methods have also been proposed for combining expert rating and exposure measurement.<sup>12</sup>

For these three aspects, the research needs are partly specific to the field of “ergonomic epidemiology”, but are also common with other fields in occupational epidemiology.

## ACKNOWLEDGEMENTS

I thank Maria Melchior for her editorial help.

*Occup Environ Med* 2005;**62**:143–144.

doi: 10.1136/oem.2004.017889

Correspondence to: Dr A Leclerc, INSERM U88, HNSM, 14 rue du val d’Osne, 94410 Saint-Maurice, France; Annette.leclerc@st-maurice.inserm.fr

Competing interests: none declared

## REFERENCES

- 1 Svendsen SW, Mathiassen SE, Bonde JP. Task based exposure assessment in ergonomic epidemiology: a study of upper arm elevation in the jobs of machinists, car mechanics, and house painters. *Occup Environ Med* 2005;**62**:18–26.
- 2 Heinrich J, Blatter BM, Bongers PM. A comparison of methods for the assessment of postural load and duration of computer use. *Occup Environ Med* 2004;**61**:1027–31.
- 3 Winkel J, Dallner M, Ericson M, and the Stockholm-MUSIC I Study Group, et al. Evaluation of a questionnaire for the estimation of physical load in epidemiologic studies. International Scientific Conference on Prevention of Work-related Musculoskeletal disorders PREMUS. *Arbete och halsa* 1992;**17**:343–5.
- 4 Hoar SK, Morrisson AS, Cole P, et al. An occupation and exposure linkage system for the study of occupational carcinogenesis. *J Occup Med* 1980;**22**:722–6.
- 5 Kromhout H, Symanski E, Rappaport SM. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. *Ann Occup Hyg* 1993;**37**:253–70.
- 6 Goldberg M, Hémon D. Occupational epidemiology and assessment of exposure. *Int J Epidemiol* 1993;**22**(suppl 2):S5–9.
- 7 Gardner LI, Landsittel DP, Nelson NA, et al. Misclassification of physical work exposures as a design issue for musculoskeletal intervention studies. *Scand J Work Environ Health* 2000;**26**:406–13.
- 8 Punnett L, van der Beek AJ. A Comparison of approaches to modeling the relationship between ergonomic exposures and upper extremity disorders. *Am J Ind Med* 2000;**37**:645–55.
- 9 Koster M, Alfredsson L, Michelsen H, et al. Retrospective versus original information on physical and psychosocial exposure at work. *Scand J Work Environ Health* 1999;**25**:410–14.
- 10 Diez-Roux AV. A glossary for multilevel analysis. *J Epidemiol Community Health* 2002;**56**:588–94.
- 11 Jansen JP, Burdorf A, Steyerberg E. A novel approach for evaluating level, frequency and duration of lumbar posture simultaneously during work. *Scand J Work Environ Health* 2001;**27**:373–80.
- 12 Wild P, Sauleau EA, Bourgard E, et al. Combining expert ratings and exposure measurements: a random effect paradigm. *Ann Occup Hyg* 2002;**46**:479–87.