dehydration and increased cardiovascular output – particularly important in the elderly or those with pre-existing disease. A range of antipsychotic drugs are known to inhibit sweating and therefore thermoregulation. Recent work has shown that deaths from respiratory and external causes are particularly increased at high temperatures (Hajat *et al*, 2007). Further research is needed on the pathophysiology of heat, but it is clear that persons with mental illness remain a high-risk group for heatwave mortality (Kovats & Ebi, 2006).

Bouchama, A. & Knochel, J. (2002) Heat stroke. *New England Journal of Medicine*, **346**, 1978–1988.

Hajat, S., Kovats, R. S. & Lachowycz, K. (2007) Heat-related and cold-related deaths in England and Wales: who is at risk? *Occupational and Environmental Medicine*, **64**, 93–100.

Kovats, R. S. & Ebi, K. L. (2006) Heatwaves and public health in Europe. *European Journal of Public Health*, **16**, 592–599.

L. A. Page Department of Psychological Medicine, Institute of Psychiatry, King's College London, Weston Education Centre, London SE5 8RJ, UK. Email: l.page@iop.kcl.ac.uk

S. Hajat, R. S. Kovats Public and Environmental Health Research Unit, London School of Hygiene and Tropical Medicine, London, UK

doi: 10.1192/bjp.191.6.560a

Avoiding errors about 'margins of error'

When discussing actuarial risk assessment instruments (ARAIs), Hart *et al* (2007) acknowledge that 'prediction' may refer to probabilistic statements (e.g. a 'prediction' that an individual 'falls in a category for which the estimated risk of violence was 52%': p. s60). For unclear reasons, however, the authors seem to value only predictions with right or wrong outcomes. They therefore regard statements about future behaviour of large groups (where one can be almost certain that the fraction of persons who act a certain way will fall within a narrow range of proportions) as potentially 'credible', but predictions for individuals as meaningless.

If the purpose of risk assessment is to make choices, then well-grounded probabilistic predictions about single events help us. Suppose we conclude that it is legally and ethically acceptable to impose preventive confinement upon individuals in ARAI categories with estimated recidivism rates above a specified threshold. This policy entails making 'false-negative' and 'false-positive' decision errors. We recognise, however, that unless we are omniscient perfection is not an option and ARAIs simply help us make better decisions than we otherwise could.

How do 'margins of error' in estimated recidivism rates affect our decision process? Hart *et al* believe their 'group risk' and 'individual risk' 95% confidence intervals speak to this problem. Their group intervals are standard confidence intervals for estimated population proportions based on random samples. If the threshold lies outside the group risk confidence interval for a category, then we can be reasonably certain that a decision we make concerning someone in that category is the same decision we would make if we knew the true recidivism rate for that category. If the threshold falls within a category's group risk confidence interval, then our estimate quite possibly might lead to the 'wrong' decision. Statistical decision theory (Berger, 1985) shows, however, that it is still a sensible strategy to choose whether to confine a member of a category based on which side of the threshold our estimated risk falls.

Hart *et al* talk about 'individual risk' as though it is something different from category (or 'group') risk. Yet if all one knows about an individual is his membership of a risk group, what can 'individual risk' mean? The authors do not say. If 'individual risk' refers to believed-to-exist-but-unspecified differences between individuals within a category, such differences should not affect choices by a rational decision-maker. The 95% CIs for 'individual risk' pile nonsense on top of meaninglessness. Hart *et al* describe the replacement of 'n' by '1' in the Wilson (1927) formulae as 'ad hoc', but this substitution makes no sense when the basis for the estimated proportion is an n-member sample. With '1' in place of 'n', the formulae just don't mean anything.

Using ARAIs raises serious moral problems as well as the valid scientific questions that Hart *et al* mention. But in faulting the capacity of ARAIs to address an unspecified quantity called 'individual risk', and in dressing up this notion with misapplied formulae for confidence intervals, Hart *et al* ultimately create a muddle.

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis* (2nd edn). Springer-Verlag.

Hart, S. D., Michie, C. & Cooke, D. J. (2007) Precision of actuarial risk assessment instruments. Evaluating the 'margins of error' of group *v.* individual predictions of violence. *British Journal of Psychiatry*, **190** (suppl. 49), s60–s65.

Wilson, E. B. (1927) Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, **22**, 209–212.

D. Mossman Department of Psychiatry, Wright State University Boonshoft School of Medicine, 627 S. Edwin C. Moses Boulevard, Dayton, Ohio 45408-1461, USA. Email: douglas.mossman@wright.edu

T. Sellke Department of Statistics, Purdue University, West Lafayette, Indiana, USA

doi: 10.1192/bjp.191.6.561

Authors' reply: Actuarial risk assessment instruments (ARAIs), constructed using data from known groups, are used to make life-and-death decisions about individuals. How precisely do they estimate risk in individual cases? The 95% CI for proportions, which evaluates the precision of risk estimates for ARAI groups, cannot be used for individual risk estimates unless one makes a very strong assumption of heterogeneity – that ARAIs carve nature at its joints, separating people with perfect accuracy into non-overlapping categories. No one, not even those who construct ARAIs, makes this assumption. So, we ask again, what is the precision of individual risk estimates made using ARAIs?

Mossman & Sellke criticise us for inadequately defining 'individual risk' and for using an ad hoc procedure to estimate the margin of error for individual risk estimates, which they opine served only to 'pile nonsense on top of meaninglessness'.

We must plead guilty to some of the charges levelled by Mossman & Sellke – indeed, we did so in our paper, acknowledging the conceptual and statistical problems with the approach we used. In our defence, we claimed duress: because developers used inappropriate statistical methods to construct ARAIs, we could not use appropriate methods to evaluate them. Violent recidivism was measured in the ARAI development samples as a dichotomous, time-dependent outcome, and so the developers ought to have used logistic regression or survival analysis to build models; if they had, one could directly calculate logistic regression or survival scores for individuals and their associated 95% CIs.

But we also plead that these charges are irrelevant to our conclusion. As we discussed, to reject our findings that the margins of error for individual risk estimates are large is to acknowledge that they are either unknown or incalculable. Regardless, the current state of affairs is unacceptable for those who seek to use these tests in a professionally responsible manner or argue in favour of their legal admissibility. We urge ARAI developers to recalibrate their statistical models in a way that permits direct calculation of individual risk estimates and their precision or to make their data publicly available so others may do so.

S. D. Hart Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A IS6. Email: hart@sfu.ca

C. Michie, D. J. Cooke Department of Psychology, Glasgow Caledonian University, Glasgow, UK
doi: 10.1192/bjp.191.6.561a

Austrian firearms: data require cautious approach

We note with interest Kapusta *et al*'s (2007) report on firearm suicide and homicide following legislative reform in Austria. However, a note of caution must be applied to statements concerning apparent consistency between Austrian and Australian experiences with firearm legislation.

Recent work demonstrates that Australia's 1996 gun laws had no significant impact on firearm homicide but that the pre-existing decline in firearm suicide accelerated post-reforms (Chapman *et al*, 2006; Baker & McPhedran, 2007). There has been an accompanying decline in non-firearm suicides beginning in the late 1990s.

Unfortunately, these findings may require re-evaluation owing to issues of data quality. The Australian Bureau of Statistics (ABS), a primary data source for researchers in the field, appear to be 'over-counting' unintentional deaths and 'under-counting' suicides. De Leo (2007) showed that ABS data 'under-counted' the total number of suicides (all methods) in one Australian State (Queensland) by 127 cases in 2004 alone. Re-analysis of the updated data reduced the apparent downward trend in suicides that had emerged from previous analyses. This finding has significant implications for assessment of suicide prevention initiatives in Australia, given that most assessments are based on ABS data.

Consequently, it has been suggested that the incidence of firearm suicide in Australia may be higher than thought, and, if so, then studies using ABS suicide figures merit re-evaluation (McPhedran & Baker, 2007). In addition, the National Injury Surveillance Unit has questioned the accuracy of homicide data, which suggests that firearm homicides may also be higher than ABS data show. There are growing calls for ABS data to be cross-checked against coronial records and for ABS records to be updated where discrepancies are found.

Although this situation does not bear directly upon the findings of the Austrian study, other than reinforcing the importance of quality control, it demonstrates that drawing conclusions about the impact or otherwise of restrictive firearm legislation in Australia may be premature.

Effective public health initiatives need to be built on accurate information. We therefore caution researchers against citing Australian figures during wider discussions of the possible role of firearm legislation in public health strategies, until and unless full data accuracy can be guaranteed.

Baker, J. & McPhedran, S. (2007) Gun laws and sudden death: did the Australian firearms legislation of 1996 make a difference? *British Journal of Criminology*, **47**, 455–469.

Chapman, S., Alpers, P., Agho, K., *et al* (2006) Australia's 1996 gun law reforms: faster falls in firearm deaths, firearm suicides, and a decade without mass shootings. *Injury Prevention*, **12**, 365–372.

De Leo, D. (2007) Suicide mortality data needs revision. *Medical Journal of Australia*, **186**, 157.

Kapusta, N. D., Etzersdorfer, E., Krall, C., *et al* (2007) Firearm legislation reform in the European Union: impact on firearm availability, firearm suicide and homicide rates in Austria. *British Journal of Psychiatry*, **191**, 253–257.

McPhedran, S. & Baker, J. (2007) Australian firearms legislation and unintentional firearm deaths: a theoretical explanation for the absence of decline following the 1996 gun laws. *Public Health*, in press.

S. McPhedran School of Psychology, University of Sydney, New South Wales, Australia, 2006. Email: samaram@psych.usyd.edu.au

S. McPhedran, J. Baker International Coalition for Women in Shooting and Hunting, Australia
doi: 10.1192/bjp.191.6.562

In Australia, the 1996 National Firearms Agreement (NFA) was introduced following the Port Arthur massacre, in which 35 people were killed. The NFA introduced access restrictions (particularly of assault weapons), storage regulations and a gun buy-back scheme to reduce firearms in the community. The recent killings at Virginia Tech have refuelled the debate on the causal impact of the NFA, with rates of homicides virtually unchanged but substantial reductions in numbers and rates of firearm suicide (Chapman *et al*, 2006). However, the dramatic decrease in suicide deaths by firearms in Australia began prior to 1996.

In Queensland, on the basis of the Queensland Suicide Register (QSR), rates of firearm suicide in 1994 were more than 30% less than those recorded in 1990 (approximately 10 in 100 000). In addition, in 1994 there was a crossing-over between declining rates of firearm suicide and increasing rates of hanging suicide. Both trends between 1990 (year of constitution of the QSR) and 2004 showed statistically significant variations (R^2=0.88 for firearms and R^2=0.70 for hanging), with firearm suicide being more than 5 times less frequent than hanging suicide in 2004 (it was 2 times more frequent in 1990). Most firearm suicides involved hunting rifles, the use of which started to appear strongly reduced by early 1990s. Minor declines were recorded in the use of other weapons.

Kapusta *et al* (2007) underline the successful effect of the Austrian reform on firearm use on both homicide and suicide rates; moreover, they did not witness any increase in suicide with other methods. We believe this has not happened in Queensland, where the current legislation has not restricted firearms within the community (around 500 000 in four million inhabitants) and there has not been a reduction in male suicide rates (De Leo *et al*, 2006). However, a big shift in the choice of suicide methods has occurred, with younger males increasingly choosing hanging. As pointed out by Kapusta *et al*, causality remains speculative in this type of observation. Although controlling access to means remains of paramount importance in suicide prevention (De Leo, 2002), it seems that a change in societal and cultural views towards firearms has played a bigger role than the NFA. To verify this, we are currently checking if those who died by suicide through other methods were also in possession (and/or had availability) of a firearm at the time of their death.

Chapman, S., Alpers, P., Agho, K., *et al* (2006) Australian's 1996 gun law reforms: faster falls in firearm deaths, firearm suicides, and a decade without mass shootings. *Injury Prevention*, **12**, 365–372.