

CRITIQUE

Lidköping Accident Prevention Programme: what was the impact?

John D Langley, Jonathan C Alsop

Community interventions have been widely embraced by the injury prevention community after the purported success of the safe community project in Falköping.¹ Much has been promised and public agencies have high expectations, yet there have been few published evaluations. The paper by Svanström *et al* in a recent edition of this journal is thus of considerable interest.²

Svanström *et al* described the Lidköping Accident Prevention Programme and its possible effect on injury incidence. A variety of intervention activities of different intensities and duration occurred between 1984 and 1991. Childhood injury rates (per 1000 persons), for girls and boys, were examined for the period 1983–91 for Lidköping and two comparison areas: four border municipalities and Skaraborg County. The authors report 'There were no statistically significant differences between the mean annual differences of the three study areas and the confidence intervals overlap' (p 170). Such an analysis is not informative. For example, one might reasonably expect different mean rates of injury between various communities as not all communities are the same. A fundamental question is: was the intervention in Lidköping associated with a significant change in injury rates over time relative to the comparison groups. The authors sought to answer this using linear regression. They presented 'beta values expressed as per cent changes' (p 171) and concluded that the results suggest that the intervention had an effect.

In their discussion the authors state 'It is assumed that the findings of such evaluations will help decision makers to make wiser choices about future courses of action than they would otherwise' (p 171). Regrettably they failed to translate their findings into the number of injuries prevented.

It is unclear to us how to translate the statistical results they report into the number of injuries prevented. Given the degree to which

community intervention programmes have been promoted and the resources they have attracted, in the absence of significant body of research on their effectiveness, we sought to reanalyse the data using different methods, with view of determining what impact the intervention had made on the incidence of injuries.

Methods and results

We begun by taking a similar approach to that reported by Svanström *et al*. That is we analysed the data within each sex and within the three areas separately. Rather than using linear regression, however, we chose to perform a Poisson regression (under the assumption that injuries per person follow a Poisson distribution, that is data obtained from a count process). The outcome variable was chosen as the actual numbers injured each year (x_i), evaluated by multiplying the injury rate by population size. The logarithm of the population size was then used as an 'offset' variable, so that the actual outcome after the log link is taken into account is the logarithm of the injury rate. The model is thus:

$$\log(x_i) = \log(N_i) + \beta_0 + \beta_1 \text{Year}_i + e_i$$

where x_i denotes the i th observation in an area and for a sex category, N_i is the population size, Year_i the original untransformed year of hospitalisation, and e_i the error term. Results for the six area/sex combinations are presented in table 1.

We can see from table 1 that none of the trends in hospitalisation rates are statistically significant at the 5% level. Large values of deviance divided by degrees of freedom (df) given an indication of lack of fit. Two of the six models appeared not to fit particularly well. These high ratios of deviance to df indicate over dispersion, which may be a result of the lack of explanatory variables included in the model. Also, each of these models had only 7 df so it was thought that a better approach would be to combine the information into an overall model.

We therefore undertook a second stage of modelling. Again a Poisson regression model was used, but with additional terms,

$$\log(x_i) = \log(N_i) + \beta_0 + \beta_1 \text{Year}_i + \beta_2 \text{Area}_i + \beta_3 \text{Sex}_i + \beta_4 \text{Area}_i \times \text{Year}_i + e_i$$

thus we also included sex, area, and an area-year interaction term. This interaction term enabled us to model possible differences in the slopes of rates over time in separate areas. The

Table 1 Poisson regression of outcome over time

Area	Year (β_1)	p Value	Deviance/df
Intervention			
F	-0.0036	0.679	1.04
M	-0.0116	0.093	1.18
Four bordering areas			
F	0.0199	0.374	2.67
M	0.0059	0.730	0.68
Skaraborg County			
F	-0.0261	0.243	1.82
M	-0.0316	0.095	2.11

Injury Prevention
Research Unit,
Department of
Preventive and Social
Medicine, University
of Otago Medical
School, PO Box 913,
Dunedin, New
Zealand
J Langley
J Alsop

Correspondence to:
Dr Langley.

main focus of the paper was to compare the intervention area with the two control areas. For this reason we decided to compare the intervention area first with the four bordering areas, and second with Skaraborg County. For these comparisons the models had 31 df, an improvement upon the original six models. The results are presented in table 2.

One of the important findings was that there were no significant time trends in any of the comparisons. This compares well with the first analysis (table 1). The differences in average injury rates were significant between the intervention area and the other two comparison areas. Also, the differences between sexes were highly significant in all three comparisons (we would expect that a similar sex effect is to be found in all three areas). One crucial finding, however, was that there was a significant difference ($p = 0.041$) in time trends between the intervention and four border areas. We can relate this in terms of a significant difference in slopes (that is non-parallel), even though the slopes themselves were not significantly different from zero. This difference was not significant ($p = 0.174$) when the intervention area was compared against Skaraborg County, though this may have been due to dependence of the areas because of geographical overlap. Note that all models fitted adequately, as the deviance divided by the number of degrees of freedom were not greatly different from 1. As an additional check, the two 'control' areas were compared. No significant difference in injury rate was found between these two areas ($p = 0.224$), and also no difference in the time trends ($p = 0.182$).

Discussion

Svanström *et al* reported a significant effect using linear regression. In our first analyses we failed to show an effect using Poisson regression. We chose to perform Poisson regression over linear regression due to a variety of reasons. The numbers of injuries may not follow a normal distribution, as they can be thought of as a count process within each individual of the population. As they are rare events, they can instead be assumed to follow a Poisson distribution. In addition, the Poisson regression constrains the number of injuries to lie between zero and infinity, while linear regression has no lower bound. Finally, a non-constant variance in the number of injuries may be present also.

Our second analysis also showed no reduction in injury rates over time. However, it showed that the intervention area had a statistically significant reduction in injury rate com-

pared with the four bordering areas. This means that the reduction in the intervention area, when examined in isolation, was not significant, but when offset against the injury rate rise in the four bordering areas, we found to be significant. Somewhat unusually, this effect was not present in the comparison of the intervention area with Skaraborg County.

From our final model we estimate that in 1991 there were 38 female and 58 male hospitalisations for the four border areas. If the intervention had taken place in this area also, we would have expected only 28 female and 42 male hospitalisations. This means that we could expect to save a total of $10 + 16 = 26$ hospitalised injuries. Note that the implicit assumption here, is that the injury rate in the Lidköping area would have risen similarly to the four border areas without the intervention. This is a very modest return on what must have represented a substantial input of resources. What might have been the return if similar resources had been concentrated on one or two very specific issues (for example helmets for cyclists and impact absorbing surfaces under playground equipment)?

The foregoing discussion assumes that the reduction in injury in the intervention area was due to the intervention and not to some other extraneous factors. Regrettably Svanström *et al* fail to provide any quantitative data on the impact of the intervention on injury risk and protective factors. We are thus presented with a 'black box intervention'. In other words we have no process evaluation. For example, although Svanström *et al* briefly describe a campaign on bicycle safety, including helmet promotion, the reader is not provided with any data on changes in bicycle safety behaviour over time. The failure to present this is a serious deterrent to understanding and progress in the field.

A recent evaluation of the Shire of Bulla Safe Living Program, a community intervention programme in Victoria, Australia, failed to show any statistically significant changes in overall injury rates or hospital bed days.³ This, coupled with our findings, suggests greater caution should be exercised in promoting these broad, multifaceted intervention programmes. Building up community expectations but failing to achieve significant reductions in injury could well result in disenchanting communities, concluding that there is, after all, nothing that they can do which can make a difference. Nothing could be further from the truth. As investment in community injury interventions grows so too has the importance of rigorous evaluation. A critical component of such evaluations is determining the impact of specific interventions (for example promotion of helmets) on specific behaviours (for example helmet wearing) and relevant injury outcomes (for example head injuries to cyclists) and seeking to determine if a range of specific interventions have had a more wider impact by influencing injury prevention behaviour. Given the resources safe communities have attracted to date a comprehensive critique of the evaluations published to date is warranted.

Table 2 Poisson regression of factors for various area comparisons

Comparison	Factor (<i>p</i> values)				
	Year	Area	Sex	Area-year	Deviance/df
Intervention <i>v</i> four bordering areas	0.417	0.001	<0.001	0.041	1.79
Intervention <i>v</i> Skaraborg County	0.117	0.002	<0.001	0.174	1.47
Four bordering areas <i>v</i> Skaraborg County	0.116	0.224	<0.001	0.182	1.33

The Injury Prevention Research Unit funded by the Accident Rehabilitation and Compensation Insurance Corporation, and the Health Research Council (HRC) of New Zealand. Dr Langley is a HRC principal research fellow. The authors wish to thank John Eastwood, David Chalmers, and Lesley Day for their useful comments on an earlier version of this paper.

- 1 Schelp L. *Epidemiology as a basis for evaluation of a community intervention programme on accidents*. Sundbyberg, Sweden: Karolinska Institute, Department of Social Medicine, 1987.
- 2 Svanström L, Ekman R, Schelp L, Lindström Å. The Lidköping Accident Prevention Programme — a community approach to preventing childhood injuries in Sweden. *Injury Prevention* 1995; 1: 169–72.
- 3 Ozanne-Smith J, Sherrard J, Brumen IA, Vulcan P. *Community based injury prevention evaluation report: Shire of Bulla Safe Living Program. The first three years 1991–1993*. Melbourne: Monash University Accident Research Centre, 1994.

Commentary: statistical perspectives on the Lidköping papers

We have examined the critique by Langley and Alsop and the original Lidköping paper and re-entered the data in order to replicate both sets of analyses. As we see it, the authors of the original paper (Svanström *et al*) do indeed lose statistical power by breaking the data down too far and performing two separate analyses (one for boys, one for girls). For each of the two genders, they find the difference in the trends over time between the intervention and neighbouring area to be non-significant ($-0.3 - 0.2 = -0.5/\text{year}$ for girls, and $-0.4 - 0.1 = -0.5/\text{year}$ for boys). The *p* values are thus 0.33 and 0.25. Instead of reporting these *p* values, they use words like 'strengthen the impression' and 'support the proposition'.

Before commenting on the method suggested by Langley and Alsop, we suggest two ways in which the original authors could have made more efficient use of their data. The first of these is to aggregate the genders and calculate one rate per year for the two genders combined. When we do this (either by a 'rougher' straight average of the two rates, or in a fancier way by adding numerators and adding denominators), we get somewhat stronger evidence of a difference in the two slopes. One way to test for a difference in slopes is to include an $\text{year} \times \text{area}$ term in the model in addition to area and year — just as Langley and Alsop did. The single *p* value is 0.18, corresponding to a 'difference in slopes' of $-0.47/\text{year}$. In the critique from New Zealand the authors give a *p* value of 0.041 for this interaction in the Poisson model. They fail, however, to tell the reader in which direction this 'difference of slopes' goes. We hope that the journal will set a firm tone by asking for magnitude and direction first, confidence interval (CI) second, and *p* values last (if at all).

The second approach, still staying with rates, is to treat the rates in boys and girls as one data set, but to account for (that is, take out), the variation in the rates between the two genders by including gender as a term in the regression. In PROC GLM in SAS for example, one can model the rates, $\text{rate} = \text{gender} \text{ year} \text{ area} \text{ year} \times \text{area}$.

If one does this with the 36 data points in the first four columns of the table in the paper, and if one codes area as 1 if intervention and 0

otherwise, and gender as 1 if boys and 0 if girls, and if we center the years by subtracting 1987 so that 1983 is -4 and 1991 is $+4$, we get the following fit (SEs in parenthesis):

$$\begin{aligned} \text{rate} &= 8.5 \\ &+ 4.3 (0.78) \text{ gender} \\ &+ 0.12 (0.21) \text{ year} \\ &+ 1.69 (0.78) \text{ area} \\ &- 0.48 (0.30) \text{ year} \times \text{area}. \end{aligned}$$

Notice that the coefficient for area is $+1.69$, indicating that, in general, rates are higher in the intervention area. The key is the -0.48 (very close to the average of the two gender specific differences in slopes in the authors' analysis. The *t* statistic for this is -1.6 so the *p* value (from a *t* with 31 df) is 0.12.

Before dealing in depth with the submission from New Zealand, two preliminary comments. First, we believe that — from a purely data analytic viewpoint — they are wise to treat the 36 observations (intervention *v* neighbouring area) as one data set rather than two (one for each gender). Second, however, we would have liked to see coefficients, not just *p* values.

Now, to the main issue, which is the use of a Poisson model. There is a paradox here in that the authors find evidence of some extra-Poisson variation when the genders are considered separately, yet mysteriously these problems 'go away' when all 36 data points are considered together. The issue of whether the Poisson distribution is — in principle, and in this data set — the 'correct' model to use is an important one. The Poisson model appears to 'bring out the signal' better than the analysis that treats the variations around the lines as Gaussian. If we were dealing with counts that were unlikely to be influenced by factors such as weather, 'local' short term interventions, clustering because of injuries to multiple persons from the same source/cause (for example, an incident with a school bus), or any other such perturbations, then the Poisson distribution would make sense. But *a priori* we would have expected extra-Poisson variation and suspect that the reason the authors don't 'find' it is that it is buried in the large number of degrees of freedom.

Indeed, when we look more closely, the ratio of deviance/df of 1.79 tells us that the deviance is around 55, which is beyond the $p = 0.01$ point of the χ^2 distribution with 31 df. So, in fact, the Poisson model does not fit that well: there is more variation than the Poisson model would predict. Thus, the standard errors obtained for the coefficients for the model in table 2 are too small, and the *p* values too extreme. Indeed, it appears that the more naive model in the original paper, which considers the variance around the line to have a magnitude independent of (bigger than) the mean, is more appropriate in this case. In general, it makes more sense to use a model that allows the variation to be estimated from the data rather than from an assumption that is not fulfilled by the data. The model used in our analysis above is the same as the one used by the original authors, except that we use all the data in one analysis.

In spite of using all the data at once, and