Repeat self-harm: application of hurdle models
Jennifer Bethell, Anne E. Rhodes, Susan J. Bondy, W. Y. Wendy Lou and Astrid Guttmann

Summary
Among those who present to the emergency department for self-harm, many will repeat. Self-harm repetition is an outcome of interest in both observational and intervention studies. However, few such studies analyse the number of repeat self-harm presentations. Here, hurdle models are introduced as a potentially useful statistical method for these analyses. Emergency department data from the Province of Ontario, Canada, are used to illustrate an example of implementing hurdle models and interpreting their results.

Declaration of interest
None.

Method
This is a population-based retrospective cohort study of 12- to 17-year-olds presenting to the emergency department for self-harm in Ontario, Canada. Data are from the National Ambulatory Care Reporting System (NACRS), covering a 7-year period (1 April 2002 to 31 March 2009). The data capture every emergency department visit; all legal residents are insured for acute and primary healthcare services and every hospital submitted NACRS emergency department data. The 2006 Ontario population of 12- to 17-year-olds was about 1 million. Ethical approval was obtained from St Michael’s Hospital.

Results
The cohort included 10,917 individuals (8012 (73.3%) girls and 2925 (26.7%) boys), of whom 3546 (32.4%) were admitted at their index episode. Overall, 1325 (12.1%) made at least one repeat self-harm presentation within 1 year of their index episode (classified as repeaters), and this proportion was almost identical in the two exposure groups (12.2% and 12.1% among admitted and non-admitted respectively).

The binary models, logistic regression and survival analysis found no statistically significant association between admission and repetition (odds ratio (OR) 1.01, P = 0.8309; hazard ratio 1.01, P = 0.8614). The count models’ AIC and BIC (Table 1) suggest substantial improvement in model fit from selecting the negative binomial, Poisson hurdle and negative binomial hurdle models over the Poisson model. Both fit indices favour the negative binomial hurdle model, demonstrating their flexibility in accounting for overdispersion from excess zeroes as well as other sources.

Interpreting the negative binomial hurdle model, similar to the binary analyses, the logit portion shows admission as the exposure variable. The outcome was repeat self-harm within 1 year of the index episode, calculated from the emergency department or in-patient discharge date (as applicable). Individuals with less than 1 year of follow-up data (index episodes after March 2008 and those who died), were excluded. The data were analysed in SAS (version 9.1.3). First, two binary models were fitted: logistic regression, categorising the outcome as repeater or non-repeater; and survival analysis (Cox regression), using time to first repeat presentation as the outcome. Next, four count models were fitted: Poisson, negative binomial, Poisson hurdle and negative binomial hurdle. The outcome was the count of repeat self-harm presentations, incorporating random effects for hospital-level clustering. Model fit was compared using Akaike and Bayesian information criteria (AIC and BIC), where smaller values are better.

Hurdle models combine a binary (e.g. logit) model with a zero-truncated count (e.g. Poisson) model. For the self-harm repetition example, the first part tests factors associated with any repetition (repeaters vs. non-repeaters) and the second part tests factors associated with the number of presentations (among repeaters). Population-based emergency department data are used to illustrate implementing and interpreting hurdle models. Hurdle models are also shown to be more informative than traditional binary analyses, but also adequately fit these data relative to some other count models.

The cohort included 10,917 individuals (8012 (73.3%) girls and 2925 (26.7%) boys), of whom 3546 (32.4%) were admitted at their index episode. Overall, 1325 (12.1%) made at least one repeat self-harm presentation within 1 year of their index episode (classified as repeaters), and this proportion was almost identical in the two exposure groups (12.2% and 12.1% among admitted and non-admitted respectively).

The binary models, logistic regression and survival analysis found no statistically significant association between admission and repetition (odds ratio (OR) 1.01, P = 0.8309; hazard ratio 1.01, P = 0.8614). The count models’ AIC and BIC (Table 1) suggest substantial improvement in model fit from selecting the negative binomial, Poisson hurdle and negative binomial hurdle models over the Poisson model. Both fit indices favour the negative binomial hurdle model, demonstrating their flexibility in accounting for overdispersion from excess zeroes as well as other sources.

Interpreting the negative binomial hurdle model, similar to the binary analyses, the logit portion shows admission as the exposure variable. The outcome was repeat self-harm within 1 year of the index episode, calculated from the emergency department or in-patient discharge date (as applicable). Individuals with less than 1 year of follow-up data (index episodes after March 2008 and those who died), were excluded. The data were analysed in SAS (version 9.1.3). First, two binary models were fitted: logistic regression, categorising the outcome as repeater or non-repeater; and survival analysis (Cox regression), using time to first repeat presentation as the outcome. Next, four count models were fitted: Poisson, negative binomial, Poisson hurdle and negative binomial hurdle. The outcome was the count of repeat self-harm presentations, incorporating random effects for hospital-level clustering. Model fit was compared using Akaike and Bayesian information criteria (AIC and BIC), where smaller values are better.

The binary models, logistic regression and survival analysis found no statistically significant association between admission and repetition (odds ratio (OR) 1.01, P = 0.8309; hazard ratio 1.01, P = 0.8614). The count models’ AIC and BIC (Table 1) suggest substantial improvement in model fit from selecting the negative binomial, Poisson hurdle and negative binomial hurdle models over the Poisson model. Both fit indices favour the negative binomial hurdle model, demonstrating their flexibility in accounting for overdispersion from excess zeroes as well as other sources.

Interpreting the negative binomial hurdle model, similar to the binary analyses, the logit portion shows admission as the exposure variable. The outcome was repeat self-harm within 1 year of the index episode, calculated from the emergency department or in-patient discharge date (as applicable). Individuals with less than 1 year of follow-up data (index episodes after March 2008 and those who died), were excluded. The data were analysed in SAS (version 9.1.3). First, two binary models were fitted: logistic regression, categorising the outcome as repeater or non-repeater; and survival analysis (Cox regression), using time to first repeat presentation as the outcome. Next, four count models were fitted: Poisson, negative binomial, Poisson hurdle and negative binomial hurdle. The outcome was the count of repeat self-harm presentations, incorporating random effects for hospital-level clustering. Model fit was compared using Akaike and Bayesian information criteria (AIC and BIC), where smaller values are better.
multinomial logistic regression. Here, we have shown that hurdle models are also an appropriate and useful statistical method. They are independent from the funding sources. No endorsement is intended or should be inferred.

The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement is intended or should be inferred.

The results highlight the importance of considering the number of repeat presentations when studying self-harm. Others have already acknowledged the tendency for self-harm to be a complex and multifaceted phenomenon, involving a variety of factors that can influence the likelihood of recurrence. These factors include both individual and environmental variables, such as prior history of self-harm, family and peer influences, and access to mental health services.

Our findings suggest that hurdle models can be of value for policy, research and clinical practice. Hurdle models are one way of assessing these patterns.

**Table 1**

<table>
<thead>
<tr>
<th>Count models</th>
<th>Coefficient (standard error)</th>
<th>P</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>−0.0537 (0.0495)</td>
<td>0.2791</td>
<td>12317</td>
<td>12327</td>
</tr>
<tr>
<td>Logt</td>
<td>0.0345 (0.0637)</td>
<td>0.5889</td>
<td>10940</td>
<td>10962</td>
</tr>
<tr>
<td>Negative binomial</td>
<td>−0.0606 (0.0697)</td>
<td>0.3854</td>
<td>10601</td>
<td>10613</td>
</tr>
<tr>
<td>Poisson hurdle</td>
<td>−0.2527 (0.0832)</td>
<td>0.0027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative binomial hurdle</td>
<td>−0.2854 (0.1194)</td>
<td>0.0179</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**AIC**: Akaike information criterion; **BIC**: Bayesian information criterion.

**References**

3. Zhang W, W. Y. Wendy Lou, Astrid Guttmann, MSc, Suicide Studies Unit, St Michael’s Hospital, 30 Bond Street, Toronto, Ontario, M5B 1W8, Canada. Email: bethell@smh.toronto.on.ca

First received 24 May 2009, final revision 4 Nov 2009, accepted 2 Dec 2009

**Funding**

J.B. is supported by a studentship from the Ontario Mental Health Foundation. A.G. is supported through a salary award from the Canadian Institutes of Health Research. The data were accessed through the Institute for Clinical Evaluative Sciences, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care.

**Acknowledgements**

The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement is intended or should be inferred.

**Correspondence**: Jennifer Bethell, Suicide Studies Unit, St Michael’s Hospital, Dalla Lana School of Public Health, University of Toronto; Anne E. Rhodes, PhD, Suicide Studies Unit, St Michael’s Hospital, Dalla Lana School of Public Health and Department of Psychiatry, University of Toronto, and Institute for Clinical Evaluative Sciences; Susan J. Bondy, PhD, W. Y. Wendy Lou, PhD, Dalla Lana School of Public Health, University of Toronto; Astrid Guttmann, MSc, Department of Paediatrics and Department of Health Policy, Management and Evaluation, University of Toronto, Institute for Clinical Evaluative Sciences, and Hospital for Sick children, Toronto, Ontario, Canada.

These results highlight the importance of considering the number of repeat presentations when studying self-harm. Others have already acknowledged the tendency for self-harm to be a complex and multifaceted phenomenon, involving a variety of factors that can influence the likelihood of recurrence. These factors include both individual and environmental variables, such as prior history of self-harm, family and peer influences, and access to mental health services.

Our findings suggest that hurdle models can be of value for policy, research and clinical practice. Hurdle models are one way of assessing these patterns.
Repeat self-harm: application of hurdle models
Jennifer Bethell, Anne E. Rhodes, Susan J. Bondy, W. Y. Wendy Lou and Astrid Guttmann
BJP 2010, 196:243-244.
Access the most recent version at DOI: 10.1192/bjp.bp.109.068809

This article cites 10 articles, 5 of which you can access for free at:
http://bjp.rcpsych.org/content/196/3/243#BIBL

To obtain reprints or permission to reproduce material from this paper, please write to permissions@rcpsych.ac.uk

You can respond to this article at /letters/submit/bjprcpsych;196/3/243

Downloaded from http://bjp.rcpsych.org/ on April 3, 2017
Published by The Royal College of Psychiatrists

To subscribe to The British Journal of Psychiatry go to:
http://bjp.rcpsych.org/site/subscriptions/