METHODOLOGIC ISSUES

# Injury surveillance in accident and emergency departments: to sample or not to sample?

A Morrison, D H Stone

**Abstract**

***Objectives*—To establish whether injury surveillance based on sampling strategies is as valid as total patient surveillance.**

***Methods*—Canadian Hospitals Injury Reporting and Prevention Program (CHIRPP) data for 1996 were retrospectively analysed using five sampling frames. Proportions for key variables were calculated for each sample, then compared with the proportions for the total population of patients.**

***Results*—Two of the five sampling frames produced statistically significant differences from the total population, which can be explained by seasonal variations. However, no significant differences were observed between the remaining three samples and the total population.**

***Conclusions*—A well planned and executed sampling strategy can generate as valid data as total patient surveillance, obviating the need for data collection on every patient presenting with an injury or poisoning. In practice, however, systematic sampling can be difficult to implement and sustain, counterbalancing the economic advantages.**

(*Injury Prevention* 1998;**4**:50–53)

Keywords: surveillance; sampling

**Paediatric Epidemiology and Community Health (PEACH) Unit, Department of Child Health, University of Glasgow, Royal Hospital for Sick Children, Yorkhill, Glasgow G3 8SJ**
A Morrison
D H Stone

Correspondence to:
Anita Morrison.

Evidence suggests that injury surveillance in accident and emergency departments is a valuable and achievable objective.[1–3] Such systems provide a proactive mechanism for routinely monitoring injury incidence, identifying risk factors, stimulating preventive programmes, and evaluating their effectiveness.[3] However, the implementation and operation of such a system involves additional, often scarce, resources. In addition to the costs of system installation and maintenance, a considerable investment in personnel is usually required to collect, computerise, and analyse information.[1 2]

Although hospitals may be sampled to provide a representative data set across regions or nations, the majority of injury surveillance systems collect information on all patients presenting to the designated accident and emergency departments,[1 4–7] making these systems relatively expensive. This paper examines whether it is necessary to include all patients in injury surveillance, or whether a sample offers a valid, and cheaper, alternative. Data collected during 1996 by the Canadian Hospitals Injury Reporting and Prevention Program (CHIRPP) system at the Royal Hospital for Sick Children in Glasgow were analysed retrospectively to examine the extent to which the use of various sampling frames would have resulted in representative samples, as reflected by the frequency distribution of selected key variables.

## Method

CHIRPP is a computerised information system that collects data on all patients presenting with injuries or poisoning to accident and emergency departments. This system has been in operation at the Accident and Emergency Department, Royal Hospital for Sick Children, Yorkhill in Glasgow since 1993.[1] During the calendar year 1996, 7940 CHIRPP forms were completed (over 90% of eligible patients). Several sampling frames were applied to these data. The sampling frames were largely arbitrary, although they had been identified as logistically possible: sample 1: every 10th attendance (n=794); sample 2: every eighth day (n=1032); sample 3: once a week (week day) (n=1104); sample 4: once a week (weekend) (n=1123); and sample 5: four months out of 12 (n=2245).

Proportions for the following key variables were calculated: age, sex, injury type, injury location, activity at the time of injury, and the body part injured. For each sampling frame, the proportions for each sample were compared with the proportions for the total population of injured children recorded on CHIRPP. Confidence intervals were calculated, and differences between proportions examined using the common proportion and the standard error.[7] The null hypothesis assumed no differences in the proportions between the samples.[8]

## Results

Three statistically significant differences were found between the proportions generated by the sampling frames and the events observed in the total population of injured children (table 1). Significantly more children sustained

*Table 1 CHIRPP data 1996: sampling frames compared; values are per cent (95% confidence interval)*

| *Variable* | *Total population of injured children (n=7940)* | *Sample 1: every 10th attendance (n=794)* | *Sample 2: every eighth day (n=1032)* | *Sample 3: once a week (weekday) (n=1104)* | *Sample 4: once a week (weekend) (n=1123)* | *Sample 5: four months (Jan Apr, Jul, Oct) (n=2245)* |
|---|---|---|---|---|---|---|
| Age (years) | | | | | | |
| 0–4 | 37 (36 to 38) | 39 (36 to 42) | 37 (34 to 40) | 37 (34 to 40) | 39 (36 to 42) | 39 (37 to 41) |
| 5–9 | 34 (33 to 35) | 31 (28 to 34) | 33 (30 to 36) | 33 (30 to 36) | 34 (31 to 37) | 33 (31 to 35) |
| 10–14 | 29 (29 to 30) | 30 (27 to 33) | 30 (27 to 33) | 30 (27 to 33) | 27 (24 to 30) | 28 (26 to 30) |
| Sex | | | | | | |
| Male | 58 (57 to 59) | 60 (57 to 63) | 56 (53 to 59) | 59 (56 to 62) | 59 (56 to 62) | 57 (55 to 59) |
| Female | 42 (41 to 43) | 40 (37 to 43) | 44 (41 to 47) | 41 (38 to 44) | 41 (28 to 44) | 43 (41 to 45) |
| Injury type | | | | | | |
| Cut/laceration | 21 (20 to 22) | 22 (19 to 25) | 19 (17 to 21) | 19 (17 to 21) | 20 (18 to 22) | 23 (21 to 25) |
| Haematoma/bruise | 19 (18 to 20) | 16 (14 to 19) | 20 (18 to 22) | 19 (17 to 21) | 19 (17 to 21) | 17 (15 to 19) |
| Fracture | 13 (12 to 14) | 11 (10 to 12) | 11 (9 to 13) | 13 (11 to 15) | 11 (9 to 13) | 13 (12 to 14) |
| Inflamation/oedema | 12 (11 to 13) | 14 (12 to 16) | 13 (11 to 15) | 14 (12 to 16) | 13 (11 to 15) | 13 (12 to 14) |
| Sprain/strain | 12 (11 to 13) | 11 (10 to 12) | 11 (9 to 13) | 10 (8 to 12) | 12 (10 to 12) | 12 (11 to 13) |
| Injury location | | | | | | |
| Own home | 42 (41 to 43) | 42 (38 to 46) | 42 (39 to 45) | 44 (41 to 47) | 46 (43 to 49) | 45 (43 to 47) |
| Public footpath | 18 (17 to 19) | 17 (14 to 20) | 15 (13 to 17) | 16 (14 to 18) | 20 (18 to 22) | 20 (18 to 22) |
| School playground | 8 (7 to 9) | 7 (5 to 9) | 9 (7 to 11) | 9 (7 to 11) | 0* | 3* (2 to 4) |
| Other home | 6 (5 to 7) | 6 (4 to 8) | 6 (5 to 7) | 5 (4 to 6) | 6 (5 to 7) | 6 (5 to 7) |
| Public playground | 6 (5 to 7) | 6 (4 to 8) | 6 (5 to 7) | 5 (4 to 6) | 8 (6 to 10) | 7 (6 to 8) |
| Activity | | | | | | |
| Playing | 67 (66 to 68) | 68 (65 to 71) | 64 (61 to 67) | 66 (63 to 69) | 70 (67 to 73) | 73* (71 to 75) |
| Walk/run/crawl | 16 (15 to 17) | 14 (12 to 16) | 16 (14 to 18) | 16 (14 to 18) | 14 (12 to 16) | 15 (13 to 17) |
| Informal sport | 5 (5 to 5) | 5 (3 to 7) | 6 (5 to 7) | 5 (4 to 6) | 4 (3 to 5) | 5 (4 to 6) |
| Body part injured | | | | | | |
| Head | 39 (38 to 40) | 40 (36 to 44) | 39 (36 to 42) | 40 (38 to 42) | 39 (36 to 42) | 38 (36 to 40) |
| Upper extremities | 35 (34 to 36) | 38 (35 to 41) | 35 (32 to 38) | 36 (34 to 38) | 35 (32 to 38) | 35 (33 to 37) |
| Lower extremities | 21 (20 to 22) | 19 (16 to 22) | 21 (19 to 23) | 21 (19 to 23) | 21 (19 to 23) | 23 (21 to 25) |

*Statistically significant difference from total population.

an injury while playing and significantly fewer sustained injuries in the school playground in the four month sample (sample 5) than in the total population. The third difference was observed between the weekend sample (sample 4) and the total population: the proportion occurring in the school playground was significantly higher in the total population. No significant differences were found between the other sampling frames (samples 1, 2, and 3) and the total population.

## Discussion

The significant differences observed in the proportions between two of the systematic sampling frames and the total population of injured children may be explained by seasonal variation. July (included in sample 5) encompasses the school holidays, when it is likely that there will be an increase in the number of injuries sustained while playing in and around the home and a decrease in the number of injuries sustained in the school playground. This is probably also true for the weekend sample (sample 4) with respect to injuries sustained in the school playground.

No other significant differences were found, suggesting that a well planned and executed sampling strategy, such as sampling every 10th attendance or collecting data every eighth day, can generate data of equal quality to surveillance including all patients. The use of a systematic sampling strategy may relieve the pressure on both personnel and other resources at accident and emergency departments by obviating the need for data collection on every patient presenting with an injury or poisoning. Such sampling strategies would not affect the overall quality of information available on a medical record because the CHIRPP data collection sheet is completed separately from routine record keeping.

Systematic sampling is also economically attractive. Adopting a sampling strategy for the CHIRPP system in Glasgow would result in cost reductions of approximately 30%. The costs of installing CHIRPP, staff training, and data analysis would remain the same. However, considerable savings would be made in the data collection, coding, and computerisation phases. Such a reduction in the costs of running an injury surveillance system may entice other public health organisations and agencies to investigate the possibilities of introducing such a system in their locality.

Nevertheless, in practice, systematic sampling may present logistical problems that counterbalance the economic attractions. Staff may forget to include every patient, or more seriously, select cases in a biased manner (according to perceived severity, for example). Other disadvantages include the inability of a sample to provide a comprehensive profile of injuries, especially unusual or densely clustered events, and the unsuitability of a sample based surveillance system for clinical purposes, such as audit or medicolegal review.

Of course, surveillance based on data collected from just one hospital (whether sampled or not) is a relatively poor alternative to a population based system. However, if as this study suggests, a systematic sampling strategy is not detrimental to the quality of data collected at any one hospital, the adoption of a systematic sampling strategy in a representative sample of hospitals in a locality may result in better quality epidemiological information than a single hospital collecting information on all injury events.

This paper has demonstrated that a systematic sampling strategy can generate as valid data on the pattern of injuries presenting to accident and emergency departments as total population surveillance. However, there are

both advantages and disadvantages to adopting sampling strategies. Each department must assess whether sampling is appropriate depending on their available resources and information requirements.

We would like to thank Mr Doraiswamy, consultant at the Accident and Emergency Department, Royal Hospital for Sick Children (RHSC), Yorkhill and the department staff for their assistance in implementing and operating the CHIRPP system. Thanks also go to the clinical audit team at RHSC for inputting and managing the data.

1 Stone DH, Doraiswamy NV. The Canadian Hospitals Injury Reporting and Prevention Program (CHIRPP) in the UK: a pilot study. *Inj Prev* 1996;**2**:47–51.
2 Beattie TF. An accident and emergency based child accident surveillance system: is it possible? *J Accid Emerg Med* 1996;**13**:116–8.
3 Graitcer PL. The development of state and local injury surveillance systems. *J Safety Res* 1987;**18**:191–8.
4 Harrison J, Tyson D. Injury surveillance in Australia. *Acta Paediatr Jpn* 1993;**35**:171–8.
5 Rogmans WHJ, Mulder S. *European home and leisure accident surveillance system: evaluation of activities undertaken in the frame of the EC-demonstration project*. Amsterdam: Consumer Safety Unit, 1990.
6 Department of Trade and Industry. *HASS / LASS reports*. London: Consumer Safety Unit, Department of Trade and Industry, 1995.
7 Canadian Hospitals Injury Reporting and Prevention Program. *CHIRPP News* (monthly). Ottawa: Bureau of Chronic Disease Epidemiology (Health Canada), 1996.
8 Bland M. *An introduction to medical statistics*. Oxford: Oxford Medical Publications, 1996.

# ANOVA, *t* tests, and linear regression

Robert W Platt

In the last issue, I discussed logistic regression and the structure of linear models when the response or outcome is binary. Binary outcomes can take on only two values, like dead/alive or boy/girl, as compared with continuous outcomes which can take on any value on a numeric scale, like blood pressure or weight. Now, let's take a step back and consider the various models and tests for continuous outcomes. The common theme in these methods is *explaining variability* in the response variable, and dividing the total variance of a statistic into variation that can be explained and random variation that cannot be explained.

The *t* test is probably the simplest commonly used statistical procedure. To compare the mean of a continuous variable in two different populations, the difference between the two means divided by its standard deviation has a special distribution, known in this case as the "*t* distribution". This relationship also allows construction of confidence intervals for the difference in means, and these provide information about the mean difference and its variability. When the difference between the two means (the between groups variability) is large relative to its standard deviation (the random variability) the *t* test will be statistically significant.

What happens when we want to test if there is a difference in means among three or more groups? Analysis of variance, or ANOVA, generalizes the *t* test to several groups. Since there are more than two groups being compared, we have to look at more than just mean differences. The method for testing the whether the mean level in all of the groups is the same follows a general pattern similar to that for the *t* test. The variance between groups summarizes the part of the total variability in the measures that can be explained by the assumption that the measurements come from different populations. The ratio between this "between groups variance" and the total variance in the dataset is high when there is a significant difference. This will occur when the means of the groups are far apart and the variability within the groups is small. The appropriate test of statistical significance here is the F test, which compares the ratio of the two variances to values found in F distribution tables.

The general test in the ANOVA model tests the null hypothesis that all of the group means are equal. Rejecting this hypothesis means that we believe that at least one difference of two means is not zero; often, we are interested in a specific difference, or in finding out which of the differences is significantly different from zero. To do this requires a second step—one that compares individual means using a modified version of the *t* test which can be done with a variety of common procedures.

Finally, consider the situation where, rather than dividing the population into groups, we wish to examine the association between a continuous outcome and a continuous variable (this can be thought of as an ANOVA where we have many different groups and these groups are ordered by the values of the continuous covariate). Here, we use linear regression, which associates the two variables through a $\beta$ coefficient.[1] This can easily be generalized to multiple regression, where we consider several covariates at the same time to try to understand their joint relationship to the outcome.

The *t* test can be thought of as a simple regression model with the covariate taking on only two values, and the ANOVA can also be viewed as a regression model with multiple covariates. More complicated ANOVA models can also be thought of in regression frameworks. The regression approach requires more work but it allows us to consider all these models in one unified framework and thus allows complete control of the comparisons made. Further, the calculation of the $\beta$ coefficients and standard errors for these coefficients allows us to use confidence intervals rather than relying on hypothesis tests as in the ANOVA.