LETTER TO THE EDITOR

James Harrison, Jerry Moller, John Dolinis
Australian Institute of Health & Welfare National Injury Surveillance Unit

Suicide mortality in NSW: geographic variations

Stewart et al raise issues of methodology, the significance of which extends well beyond the instance of suicide mortality[1]. The recent emergence of an approach to managing public health in Australia (at both State and national levels) through defining, setting and monitoring quantitative targets carries implicit technical challenges, for which most public health practitioners are ill-prepared.

A key challenge is to provide advice that enables policy-makers and other users of the information to make good comparisons of values of indicators, over time and between places. Frequently the indicators are derived from complete counts of deaths or hospital separations (and hence have no sampling error) and are expressed as population-based rates. The methods for analysing these data properly are not trivial and are not (in our experience) given much attention in the training that public health practitioners receive in epidemiology and biostatistics.

The article by Stewart et al is at the forefront of attempts to meet this challenge and can also help to describe it. We note that they have calculated confidence intervals on the basis of an assumption that the underlying distribution is a Poisson distribution. An assumption normally required for valid use of this distribution is that the data do not have marked trends. We have not seen time series of suicide for the areas studied and we do not know whether they show marked trends. The point we wish to make applies in any case: questions may arise in the course of routine analysis of these routine data which few public health practitioners have been trained to solve. For example, how much trend would constitute a violation of the Poisson assumption, would such a violation materially affect the findings and what alternative method might be more appropriate?

Selection of an appropriate distributional assumption is a problem that arises in other ways. For example, at the geographic level of analysis used by Stewart et al (Health Areas and Districts in NSW), the Poisson distribution may be the most appropriate in a study of suicide. Is it also the most appropriate in a study at State or national level, where case numbers are larger? On another tack, suicide cases are sometimes found to cluster in time and place. Should the negative binomial distribution be assumed?

Selection of an appropriate distribution is only one aspect of the task. Stewart et al raise the issue of analysing change where case numbers (and sometimes also populations) are small. How might one go about evaluating the impact of a suicide prevention program directed at a population such as that of the Far West Health District?

Most textbooks of biostatistics do not address such questions directly (the best we have seen is the latest in the Statistical Methods in Cancer Research series, only recently published in English[2]). As practitioners, we seem to deal with them in one of three ways. We seek advice from a biostatistician (if we have access to one), we "have a go"

Research and development

► *Continued from page 91*

- propose that Area and District Health Services develop local R&D policies and channel funding to support infrastructure for local groups which attract peer-reviewed grant funding;
- relate R&D funding to accountability requirements under the NSW Health Department's new program reporting structure;
- emphasise the importance of effective reporting and communication of research results, to promote the dissemination and application of research-based knowledge and facilitate the monitoring of research outputs and outcomes;
- seek to foster effective working relationships among different types of research organisations in different localities throughout the State;
- propose a commitment to R&D investment in improving
  – health system planning,
  – the organisation of health services, and
  – clinical, public health and managerial decision-making,
  with a particular emphasis on health informatics;
- identify specific initiatives to facilitate R&D, including support for institutional ethics committees;
- support the development of an R&D workforce, with emphasis on career opportunities in the biomedical sciences and training in health economics, public health and applied epidemiology, clinical epidemiology and health informatics; and
- propose approaches to monitoring the effectiveness of health and medical research in NSW.

HEALTH INDUSTRY DEVELOPMENT

Parallel to this strategy, the NSW Health Industry Forum is evolving a plan for health industry development. Issues papers emanating from the Forum expand on points raised in this paper, especially the commercial development of research.

NEXT STEPS

This discussion paper is being circulated for comment. When comments have been received a workshop will be convened to discuss options and make recommendations on the NSW Health Department's R&D policy. A final paper will be issued. It will incorporate these recommendations and the Department's response and will set a firm agenda for the implementation of policies and plans.

Copies of the discussion paper can be obtained from Amanda Lees, R&D Policy Branch, Centre for Research & Development, NSW Health Department (telephone 02 391-9204).

(based sometimes on limited knowledge), or we avoid the issue by limiting analysis to simple inspection of data. Of these, only the first option is satisfactory and it depends on a limited and expensive resource. Moreover, advice seems to vary.

Given the emerging significance of quantitative targets for public health practice, we suggest that a more systematic approach should be taken to the problem. One approach would be to fund one or more statisticians to prepare a paper or short handbook, designed for use by practitioners who have some relevant training (say, at Master of Public Health level), presenting methods and worked examples of the analytic tasks commonly involved in monitoring targets. This would not replace the need for proper biostatistical advice concerning special studies and unusual circumstances. It would be intended as an aid to the increasingly routine tasks imposed by the move towards target-based public health practice. We would welcome other views on this.

---

1. Stewart G, Chipps J, Sayer G. Suicide mortality in NSW: geographic variations. *NSW Public Health Bulletin* 1995; 6(6):49-52.
2. Esteve J, Benhamou E, Raymond L. Statistical methods in cancer research volume IV: Descriptive Epidemiology (IARC Publications; 128) Lyon: IARC, 1994 (ISBN 92 832 2128 1).

ON THE ACHIEVEMENT OF REAL TARGETS: REPLY TO HARRISON ET AL

Gavin Stewart, Jennifer Chipps, Geoffrey Sayer

We welcome the opportunity to expand on some of the issues mentioned in our article, and on others raised by the letter from Dr Harrison and his colleagues. The main purpose of our paper was to **illustrate** something that can be **proven** quite easily by a calculation: that the estimate of the rate of a rare event in a small population is not at all precise, that suicide is a rare event and that District Health Services in NSW have small populations. The problem is that calculations of this kind are not very helpful to those with the responsibility for addressing health problems as important as suicide. We felt it would be valuable to present **a conventional** analysis of geographic variation in suicide rates and indicate the problems that arise.

The many challenges in defining and monitoring progress towards quantitative targets are often not appreciated. For example, the introduction of the Commonwealth publication on the national goals and targets, Better Health Outcomes for Australians[1], contains a very good discussion of the general issues, but includes some unsatisfactory definitions. The term "descriptive target" is used to refer to targets for which *"... there are not adequate trend data, nor sufficient understanding of the potential impact of interventions to state with any certainty that the targets are attainable"*.

The introduction also states that *"... targets are a method of assessing progress towards desired health outcomes. They should not be seen as 'magic' numbers that must be precisely achieved by the year 2000"*.

It would indeed be very hard to make an absolute commitment to reducing the youth suicide rate by 15 per cent over a 10-year period, especially if progress were monitored and there were sanctions for failing to stay on track towards the target. One would be very worried by the lack of proven interventions. All sorts of difficult and expensive things would need to be set in train, quantified and costed. There is no "magic" in numbers one must achieve precisely. The magical thinking lies in supposing that one can have the benefits of numerical targets without solving the problems of taking them seriously. Those were the sorts of issues we wished to illustrate.

The technical concerns about the effect of assuming independent Poisson variates in pooling data over time are quite correct, if one assumes that apparent trends within strata are real. If so, the resulting distribution would have a larger variance than we assumed, our confidence intervals would be wider than we estimated on the assumption of independence and some of the "detectable" differences might become "undetectable". The boundary between "detectable" and "undetectable" differences would also change if we had adopted a different but statistically simpler decision rule – for example, insisting on 99 per cent confidence for ***any*** "detected" difference across all comparisons.

We agree that it would be valuable to have a handbook covering techniques that are going to be commonly needed in dealing with the setting of realistic goals and targets, and monitoring change. Although we are less in agreement that it should be written only by statisticians, we strongly recommend the first two chapters of Data Analysis and Regression: A Second Course in Statistics by Mosteller and Tukey[2]. Their definition of "indicator" includes any "summary" of data, including a graphical one, and needs to satisfy only two criteria:

- "It must differ from an anecdote by allowing each of the observations to contribute to it. (Anecdotes usually involve one or a few observations).
- It must be expressed in such a way that at least some of those who are interested in the subject can think about its interpretation."

It is from this perspective that we approach the task of presenting "good enough" rather than technically innovative analyses. To fit a complex regression model to the within-strata trends over time (which exist, at least for males, in some age groups) and formally model the data simply to draw the same conclusion would not have been warranted. Where possible changes over time are of substantive interest, as in our latest report on suicide in mental health clients[3], we have used the SAS procedure GENMOD to test appropriate models.

Our general approach, following Mosteller and Tukey but emphasising the second criterion, is to present the most straightforward analysis and leave the technical apparatus as much as possible in the background. It is not the technical accuracy of a single analysis, but rather the convergence of the conclusions from a variety of possibly accurate analyses, which is important. Conclusions should be robust over a range of reasonable guesses about the true nature of the data, which is rarely – if ever – determinable from the actual data. The existence of an apparent trend over time is not, in and of itself, evidence that the rate of suicide in one year is correlated with the rate of suicide in

Continued on page 94 ►

► Continued from page 93

another. Thus pooling (which simply forms a weighted sum of Poisson variates) is one of the possibly legitimate analyses, and our results are entirely accurate under the assumption that the 14-year pooled rate within a stratum is the weighted sum of independent Poisson variates[4].

No doubt it would be possible to model the 14-year suicide data in more complex and technically preferable ways. However, the overall conclusion remains robust against any likely violation of the assumptions of our analysis. Suicide rates **cannot** be used as indicators to monitor the short-term achievements of local programs. That alone is a useful result, because it relieves a great deal of concern in the minds of those who feel responsible for taking action but have an intuitive feeling for the uncertainty of the local suicide rate as an indicator of achievement in small populations, where one or two cases can make an enormous difference to the observed rate[5]. To quote the conclusion of an excellent paper on small area analyses[6]: *"In the absence of a prior hypothesis, small area analysis of epidemiological data for periods of less than 10 years will almost always give misleading results for all but the most common diseases."* The recommendation in that paper was for case-control studies, and that was one of the specific issues we addressed in our follow-up article on clinical audit of suicides[7].

In reality, the only acceptable suicide rate is zero and any other observed number must represent some degree of failure. The existing approach to target achievement reminds one of primitive ballistics. We choose the target, pour in some funding, light the enthusiasm and trust that all will be well. That is simply not good enough, or, if it is, then epidemiologists have little role to play in the process. If we are to learn from the United States experience[8,9], we should make a strong distinction between **real** targets versus wishes, and in the former case we should quantify the attributable benefit of programs, and fund them and evaluate them, on an ongoing basis. We should avoid "descriptive targets", or targets that need not be precisely achieved, and the processes that tend to flow from aiming at such vague things.

To say that one does not know how to achieve a desirable change is the first step in acquiring knowledge. It generates the right kind of activity – either to invest in finding out if others know things that we do not, or in conducting investigations ourselves, to bridge the gap between where we are and where we want to be, even if neither of those things can be quantified very well, as in the case of many aspects of mental illness and the precursors of suicide. It means that evaluation must be a major part of any program which **is** funded and not just whatever is easiest to evaluate, or most conventional. The evaluation must focus on the provable connection between the program and some outcome closer to the desired end-state.

This means that epidemiologists will have to acquire program evaluation skills and learn about psychosocial research, as well as improving their analyses of mortality data. A handbook would be useful, but it would need to cover more than statistical methods.

The Mental Health Epidemiology Group (MHEG)[10] has been in operation for only a few months. Most mental health data in NSW have never been analysed, even in a conventional way, and each new analysis presents unexpected pitfalls. In these circumstances we prefer to stay on the safe ground of conventional analyses so that any curious aspects of the data will not be confounded with analytical novelty. We appreciate the comments on our first paper on suicide mortality. We would be delighted, however, if others would carry the methodological work forward.

1. Better health outcomes for Australians: national goals, targets and strategies for better health outcomes into the next century. Commonwealth of Australia, 1994.
2. Mosteller F, Tukey JW. Data analysis and regression: a second course in statistics. Reading, Massachusetts: Addison-Wesley, 1977.
3. Chipps J, Stewart G, Sayer G. Suicide mortality in NSW: clients of mental health services. *NSW Public Health Bulletin* 1995; 6(8):75-81 (in press).
4. Dobson AJ, Kuulsmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine* 1991; 10:457-462. (We thank Dr Tim Churches for this and the following reference.)
5. Stevenson JM, Olson D. Methods for analysing county-level mortality rates. *Statistics in Medicine* 1993; 12:393-401.
6. Hole DJ, Lamont DW. Problems in the interpretation of small area analysis of epidemiological data: the case of cancer incidence in the West of Scotland. *J Epidemiology and Community Health* 1992; 46:305-310. (We thank Dr Peter Sainsbury for providing this reference.)
7. Chipps J, Stewart G, Sayer G. Suicide mortality in NSW: an introduction to clinical audits. *NSW Public Health Bulletin* 1995; 6(7):68-70.
8. National Center for Health Statistics. *Health, United States,* 1994. Hyattsville, Maryland: Public Health Service, 1995.
9. In our previous paper we reported that in 1980 the US set a youth suicide target of 11/100,000 for 1990, against a baseline of 12.4/100,000 in 1978, and at the mid-point in 1985 the rate had actually increased to 12.9/100,000. We have since obtained more recent data. The rates per 100,000 people aged 15-24 were 12.9, 13.0, 13.2, 13.1 and 13.0 in the years 1988-1992 respectively. Perhaps they would have been much higher but for the target being set. But clearly, there was no magic in the number chosen by the United States in 1980 and it was not "precisely" achieved in 1990.
10. Membership of MHEG is open to people with a professional interest and expertise in mental health epidemiology who are willing to contribute to the planning and production of a series of publications and reports on important mental health topics. The policy of MHEG is joint publication by the group as a whole in which authors are listed in order of their contribution to the particular report. The contact address for MHEG on matters concerning this report is: Mental Health Epidemiology Group, Centre for Research & Development, Public Health Division, NSW Health Department, Locked Bag 961, PO North Sydney 2059 (Fax: 391-9232, Internet e-mail gstew@gwsm.doh.health.nsw.gov.au).