

## METHODOLOGIC ISSUES

# Injury outcome indicators: the development of a validation tool

C Cryer, J D Langley, S N Jarvis, S G Mackenzie, S C R Stephenson, P Heywood, on behalf of the International Collaborative Effort on Injury Statistics Injury Indicators Group

*Injury Prevention* 2005;11:53–57. doi: 10.1136/ip.2003.004143

See end of article for authors' affiliations

Correspondence to: Associate Professor Colin Cryer, Injury Prevention Research Unit, Department of Preventive and Social Medicine, Dunedin School of Medicine, University of Otago, PO Box 913, Dunedin, New Zealand; colin.cryer@ipru.otago.ac.nz

**Background:** Researchers have previously expressed concern about some national indicators of injury incidence and have argued that indicators should be validated before their introduction.

**Aims:** To develop a tool to assess the validity of indicators of injury incidence and to carry out initial testing of the tool to explore consistency on application.

**Methods:** Previously proposed criteria were shared for comment with members of the International Collaborative Effort on Injury Statistics (ICE) Injury Indicators Group over a period of six months. Immediately after, at a meeting of Injury ICE in Washington, DC in April 2001, revised criteria were agreed over two days of meetings. The criteria were applied, by three raters, to six non-fatal indicators that underpin the national road safety targets for Canada, New Zealand, and the United Kingdom. Consistency of ratings were judged.

**Consensus outcome:** The development process resulted in a validation tool that comprised criteria relating to: (1) case definition, (2) a focus on serious injury, (3) unbiased case ascertainment, (4) source data for the indicator being representative of the target population, (5) availability of data to generate the indicator, and (6) the existence of a full written specification for the indicator. On application of these criteria to the six road safety indicators, some problems of agreement between raters were identified.

**Conclusion:** This paper has presented an early step in the development of a tool for validating injury indicators, as well as some directions that can be taken in its further development.

Sim and Mackie have described the importance of indicators in their recent editorial.<sup>1</sup>

“Public health systems across the world are being encouraged ... to show evidence of health gain. Defining accurate indicators has become a major area of work for academics and professionals. ... Getting such indicators right is essential since the effectiveness of [national and] local healthcare systems may be judged using such indicators. Perhaps more importantly, financial resources may flow—or be withheld—on the basis of such indicators”.

How do we ascertain whether we are getting an indicator right? This is a question of validity. Validity in measurement addresses the degree to which the concept under study is accurately represented by the particular measuring device.<sup>2</sup> “In the crudest terms, we say that [an indicator] is valid when it measures what it is presumed to measure”.<sup>3</sup>

The choice of indicator does matter. Work by Langley and colleagues has illustrated this in regard to motor vehicle traffic crash indicators in New Zealand.<sup>4</sup> Different trends over time are observed for different indicators. Whereas some show major reductions in injury incidence (rate) over time, others show little or no reduction. The influence of indicators can be great, and the trends demonstrated by these various indicators could affect policy, programmes, and expenditure in different ways. So the choice of indicator is important; and it is important to choose valid indicators. Although this may be obvious to many, it seems that groups who have devised indicators have often proposed and implemented international, national, and state indicators without any apparent systematic consideration of validity issues. Validation should be an essential element to guide the activities of these groups.

The aim of the work described in this paper was to further develop a tool for assessing the validity of injury outcome indicators and then to carry out initial testing of the tool to explore consistency on application.

## METHODS

We and others have expressed concern about some national indicators<sup>5–6</sup> and argued that current international, national, and state indicators should be validated before their introduction.<sup>6–7</sup> Our previous work developed an early version of a validation tool that was based on identifying those characteristics that we believe an ideal injury indicator should possess.<sup>5–9</sup> Namely: the indicator would focus attention on important events<sup>8</sup>; reflect the underlying events it is trying to measure; be based on an epidemiological definition of an event; and be derived from routinely or easily collected data. We developed this work further in the following manner.

Background and justification for the four original validation criteria were shared, with a description of how the work of this group would progress, via email with members of the International Collaborative Effort on Injury Statistics (ICE) Injury Indicators Group, six months before a planned meeting of the Injury ICE. Members of the group were invited to comment, and the documentation was modified and enhanced on the basis of feedback received. There were three such iterations over the six months.

These criteria were then presented to and discussed at the Injury ICE meeting held in Washington, DC in April 2001. Forty six government and academic statisticians, epidemiologists, nosologists, and others involved with injury data, from

\*Important events, in this context, are those that result in injury that are associated with significant threat-to-life, threat-of-disablement, loss of quality of life, or increased cost.

10 countries, the European Commission, and the World Health Organisation, attended this meeting. The list of participants can be found at: [http://www.cdc.gov/nchs/data/ice/part4\\_2\\_01.pdf](http://www.cdc.gov/nchs/data/ice/part4_2_01.pdf).

At the meeting, the results of the six month consultation were initially presented at the plenary session. This was followed by two breakout group meetings during which the strengths and limitations of the proposed criteria were discussed, and modifications, enhancements, and new criteria proposed. These criteria were presented back to the full plenary where they were discussed further, and agreed.

Initial testing of the validation criteria was subsequently carried out. Three raters applied the criteria to six injury indicators that underpin national road safety targets. A desirable property of a measurement instrument is reliability—for example, the existence of limited variation between raters. In a similar manner, the hoped for outcome of this initial testing was limited variability between raters. Totally consistent results between three raters could occur by chance. Consequently little emphasis was placed on the consistent results in this testing. More interesting was where the scores provided by the raters were quite different from one another. Large discrepancies in scores, even in this small test, indicated problems of consistency.

The criteria were applied to the non-fatal† injury indicators that underpin the national road safety targets for Canada,<sup>10</sup> New Zealand,<sup>11</sup> and the United Kingdom,<sup>12</sup> listed in box 1. Written specifications of the road safety indicators were not available within any of the three countries and so specifications were derived for the purposes of this project. Information was requested from representatives of the national road safety departments in Canada, New Zealand and the United Kingdom, and one of the authors (CC) drafted specifications of the selected indicators. The specifications included: definitions of terms used in the specification, a statement of the indicator, sources of data, the method of calculation of the indicator, and the entity the indicator aimed to reflect. These were shared with the relevant road safety department representatives for their comment, to facilitate the completion of accurate specifications, before finalising.

These standardised specifications of each of the indicators were sent to each of the authors who acted as raters (SGM, JDL, and SNJ) with instructions to assess each indicator against each of the criteria included in the validation tool. The raters were asked to assign a score ranging from 0 (criterion not satisfied at all) to 10 (criterion completely satisfied) for each possible indicator-criterion combination.

### CONSENSUS OUTCOME

The outcome of this process was that the wording of the original four criteria was changed, and the criteria were expanded from four to six by the addition of the requirements for “representativeness” and “indicator specification”. The criteria, with brief justification for their inclusion, are described below.

#### (1) Case definition

The indicator should reflect the occurrence of injury satisfying some case definition of anatomical or physiological damage:

Case definition should be based on diagnosis, defects, pathology, etc rather than use of services, since use of service is heavily influenced by supply of and access to service, professional decisions, and the behaviour of the various population groups.

† Some of the “non-fatal” injury indicators include fatal and non-fatal injuries. They are described as non-fatal injury indicators since the vast majority of the injuries captured by these indicators are non-fatal.

### Box 1: Non-fatal indicators underpinning road safety targets in Canada, New Zealand, and the United Kingdom

#### Canada (C)

- C1: Number of road users killed and seriously\* injured in motor vehicle traffic crashes (*police-serious*)†
- C1.8: Number of vulnerable (pedestrians, motorcyclists and cyclists) road users killed and seriously injured in motor vehicle traffic crashes (*police-serious-vulnerable*)

#### New Zealand (NZ)

- NZ1: Reported injuries resulting from motor vehicle (MV) accidents per 100 000 people (*police-all*)
- NZ2: Number hospitalised (discharges) people for reported injuries resulting from motor vehicle accidents (*hospitalisations*)

#### United Kingdom (UK)

- UK1: Number of road users killed and seriously‡ injured in road accidents (*police-serious*)
- UK2: Number of people slightly§ injured in road accidents per 100 million vehicle kilometers (*police-slight*)

#### Footnotes:

\*Defined as comprising people involved in a traffic crash who suffer non-fatal injuries that result in hospitalization, including for observation only, for a period of at least 24 hours. (Based on police registered crashes. Police make a judgment regarding which injuries result in a hospital stay of at least 24 hours.)

†In parenthesis are the short form names for the indicators that are used in the text.

‡Serious injury includes fracture, internal injury, severe cuts, crushing, burns (excluding friction burns), concussion, severe general shock requiring hospital treatment, detention in hospital as an inpatient, either immediately or later, injuries to casualties who die 30 or more days after the accident from injuries sustained in that accident. (Based on police attendances or notifications of a crash, not on hospital records.)

§Slight injury include sprains, including neck and whiplash injury, not necessarily requiring medical treatment, bruises, slight cuts, slight shock requiring roadside attention. Persons who are merely shaken and who have no other injury should not be included unless they receive or appear to need medical treatment.

#### (2) Serious injury

The indicator should be based on events that are associated with significantly increased risk of impairment, functional limitation, disability or death, decreased quality of life, or increased cost (that is, serious injury):

Serious injury is associated with substantially higher average burden of injury, compared with minor injury, in terms of cost, functional capacity, impairment, disability, quality of life, and survival. Greatest priority should be given to the prevention of the most serious injuries.

#### (3) Case ascertainment

The probability of a case being ascertained should be independent of social, economic, and demographic factors, as well as service supply and access factors:

We want indicators to measure, with minimal bias, the occurrence of injury rather than the use of services.

#### (4) Representativeness

The indicator should be derived from data that are inclusive or representative of the target population that the indicator aims to reflect:

We want the indicator to measure the occurrence of events relating to all subpopulations equally well.

#### (5) Data availability

It should be possible to use existing data systems, or it should be practical to develop new systems, to provide data for computing the indicator:

Given the cost associated with collecting and analysing data on a regular basis we need to capitalise on existing systems. Typically, we also have a need for information now.

#### (6) Indicator specification

The indicator should be fully specified to allow calculation to be consistent at any place and at any time:

In order to be able replicate the indicator consistently across populations, places and over time, a comprehensive written specification is required that includes definitions, specifications of data sources, and methods of calculating the indicator.

It should be noted that the above criteria (and the proposed additional criteria in the discussion to follow) are aspirational in the sense that we do not expect indicators to completely satisfy these criteria. However, they are goals to aim at when developing new indicators (or assessing existing ones). We would expect indicators for which there is little or no threat to validity to satisfy these criteria as far as is practically possible.

### PRELIMINARY TESTING

The results of the assessments made by the three raters are shown in table 1. Criterion 5 was not relevant for inclusion in the assessments since it only applies to the development of new indicators, and is inevitably satisfied for these existing indicators. Additionally, criterion 6 was not satisfied since written specifications from the national road safety agencies did not exist. Only the results for criterion 1 to 4 are presented here.

The ratings indicated the following:

- For all the indicators, there was reasonable consensus between the three raters that criterion 1 (case definition) was poorly satisfied, with the exception of the United Kingdom indicators, where ratings were inconsistent.
- There was reasonable consensus that criterion 2 (serious injury) was poorly satisfied for the first New Zealand (*police-all*) and the second United Kingdom indicator (*police-slight*), but that for the remainder the ratings were inconsistent.
- There was reasonable consensus across all the indicators that criterion 3 (case ascertainment) was satisfied to a moderate degree, with the exception of the first United Kingdom indicator (*police-serious*) for which the results were inconsistent.
- For criterion 4 (representativeness), there was a reasonable degree of consistency of rating across all indicators.

### DISCUSSION

Work of this nature is central to the better measurement of injury incidence, which is a necessary part of effective injury prevention. This paper reports on an early step towards formalisation of the development and validation of injury outcome indicators in a way that should improve their utility. Subsequent steps in this work to develop validation criteria will include refining the tool according to the findings of this exploratory study.

Consensus was reached on an expanded validation tool for injury indicators; however, some of the criteria may not be relevant in some applications. For example, some participants at the Injury ICE meeting expressed the view that in some instances it would be appropriate to replace criterion (2)—serious injury—with (2a) “The indicator should reflect a well-defined information objective”.

One approach to improve consistency between raters is to develop stronger guidelines for raters. For example, for criterion 1 the following scoring system could be developed further:

- 0: None or grossly inconsistent case definition
- 1–3: Largely consistent case definition used, but not based on anatomical or physiological damage
- 4–6: Largely consistent case definition based on anatomical/physiological damage but poorly measured

**Table 1** Ratings\* of each road safety indicator against the validation criteria†

Indicators‡	Validation criteria‡			
	1. Case definition	2. Serious injury	3. Case ascertainment	4. Representativeness
C1: Police-serious	<b>0</b> ,0,4	2,6,7	6,6,8	7,8,8
C1.8: Police-serious-vulnerable	<b>0,4</b> **	2,7,**	5,8,**	5,8,**
NZ1: Police-all	0,0,2	0,3,4	4,4,5	5,6,7
NZ2: Hospitalisations	0,0,3	2,5,6	5,5,8	8,8,10
UK1: Police-serious	0,6,10	4,5,8	4,7,9	7,9,9
UK2: Police-slight	0,5,10	0,1,4	4,4,5	6,7,9

\*10 point rating system: 0 = validation criterion not satisfied; 10 = validation criterion completely satisfied; 1–9 = validation criterion incompletely satisfied—the magnitude of the score indicating the degree to which the criterion is satisfied.

\*\* = One rater was unable to assess this indicator.

†See box 1 for a statement of the non-fatal indicators.

‡See “Consensus outcome” for a full statement of the criteria included in the validation tool for indicators of injury incidence.

§Ratings shown in bold were made by the assessor living in that country (for example, the assessment of C1 shown in bold text was made by the Canadian assessor).

### Key points

- We and others have previously expressed concern about some national indicators of injury incidence and have argued that current international, national, and state indicators should be validated using a suite of methods before their introduction.
- Our previous work developed an early version of a validation tool based on identifying those characteristics that an ideal injury indicator should possess.
- In this current work, consensus among experts and professionals who work with injury data was reached on an expanded validation tool for injury indicators.
- We tested this tool and identified a number of ways in which both it, and the testing procedures, could be improved.
- This paper has presented an early but important step in the development of a tool for validating injury indicators, as well as some directions that can be taken in its further development.

- 7–9: Largely consistent case definition based on anatomical/physiological damage and measured well
- 10: Consistent case definition and assessment based on anatomical/physiological damage

Similar scoring systems could be developed for each criterion.

There is some similarity between criterion 3 and criterion 4, in that they both deal with the ascertainment of cases that satisfy the chosen definition of a case. This caused difficulties for at least one of the raters. The aims of criteria 3 and 4 differ, however. Criterion 3 aims to identify technical problems due to variations in case ascertainment resulting from the limitations of the source data (for example, given a reportable crash has occurred, the likelihood of reports to the police of a road traffic crash depends on age of the injured person). On the other hand, criterion 4 aims to identify problems due to a mismatch between the scope of the source data for the indicator and the scope of the indicator. For example, if the aim is to have an indicator that focuses attention on vehicle crashes among cyclists, then use of police reports for this purpose is problematic in some countries since, in some circumstances, they exclude these injury events.<sup>13</sup>

A limitation of the validation criteria is that they solely address the characteristics of the numerator of the indicator, and make no judgment about the denominator. For example, these criteria would not distinguish between indicators based on cyclist injuries per 1000 population, or per 1000 cyclists, or per 10 000 cycled kilometers. The following further validation criterion is therefore suggested:

- The denominator(s), in instances where the indicator is a risk, rate or ratio estimate, should be specified to reflect the exposure of the population to relevant injury hazards.

This is related to criterion 2a above, namely: “The indicator should reflect a well-defined information objective”.

Crucial to the usefulness and validity of indicators are data related issues. For example:

- (1) In the United Kingdom, although the situation has been improving, there is still incomplete external cause of injury coding for hospital inpatient data;

- (2) In New Zealand, there have been significant changes in the use of some diagnosis codes—use of some multiple injury codes has declined substantially, since coders are tending to code each injury that make up the multiple injury and to record this information in the first, second, third, etc, diagnosis fields in the hospital inpatient record.

Additionally, in all three countries, some of these indicators are based on data collected by the police. Key fields in these data sets based on police reports (for example, severity of injury) have been found to include inaccurate data.<sup>14 15</sup> Substantial inaccuracies and changes to data sources will obviously affect the validity of indicators. Some additional criteria that should be considered are, therefore:

- The indicator should be derived from data that reach acceptable levels of completeness and accuracy.
- The indicator should be robust to potential or known changes/differences in coding frames or in coding practice between places or over time.

The discussion above has suggested an expansion in the number of validation criteria. On the other hand, some may desire that the number of validation criteria be kept to a minimum. Work in Australia has used an abridged and modified version of our six criteria to guide their national indicator assessment and development<sup>16</sup>:

- (A) Case definition should be in terms of specified anatomical or physiological damage.
- (B) Cases included should be all of those that the indicator aims to reflect, or a well defined sample of them.
- (C) Probability of case ascertainment should be independent of extraneous factors.

The work reported here is an incremental step in a program of work in which we, and others in the International Collaborative Effort on Injury Statistics, have been engaged for several years. The preliminary testing of the validation tool, based on comparison of assessments of road safety indicators by several raters, is novel in this area. Subsequent steps include: refine the tool according to the findings of this exploratory study (for example, greater clarity on how to score validity on each criterion); use consensus methods involving a representative group of injury professionals to revise the criteria; apply the criteria to a variety of indicators across several countries and target areas (for example, transport, work) using a number of independent evaluators with different disciplinary perspectives to assess both the performance of the criteria and the theoretical validity of these indicators.

Application of our current (or further developed) validation criteria to assess an indicator would need to be followed up with consideration of other aspects of validity, particularly those that require empirical investigation, for example, criterion validity in which the indicator is judged against a statistical “gold standard”.<sup>6 17</sup> If at any stage of the validation an indicator is found wanting, this should trigger a reconsideration of the indicator, its possible rejection, and the development/search for alternatives. Validation of indicators is necessary if we wish to minimise the risk of deluding ourselves about apparent gains (or losses) we have made in reducing injury incidence.

### ACKNOWLEDGEMENT

The Centre for Health Services Studies is funded by the English Department of Health as a Research and Development Support Unit. The New Zealand Injury Prevention Research Unit is funded by the Health Research Council of New Zealand and the Accident

Compensation Corporation. The views and/or conclusions in this article are those of the authors and do not necessarily reflect those of the funders. We would like to thank the International Collaborative Effort on Injury Statistics and the participants at the meeting in Washington, DC in 2001. Our thanks also to Paul Gutoskie, Transport Canada, Wayne Jones, Land Transport Safety Authority, New Zealand, and Valerie Davies, United Kingdom Department for Transport for their help when developing the specifications of the injury indicators. We thank the referees of an earlier draft of this paper for their extremely helpful comments.

#### Authors' affiliations

**C Cryer**, Centre for Health Services Studies, University of Kent, Canterbury, UK and Injury Prevention Research Unit, University of Otago, Dunedin, New Zealand

**J D Langley**, **S C R Stephenson**, Injury Prevention Research Unit, University of Otago, Dunedin, New Zealand

**S N Jarvis**, Community Child Health, University of Newcastle, Newcastle upon Tyne, UK

**S G Mackenzie**, Health Surveillance and Epidemiology Division, Health Canada, Ottawa, Canada

**P Heywood**, School of Population and Health Sciences, University of Newcastle, Newcastle upon Tyne, UK

#### REFERENCES

- 1 **Sim F**, Mackie P. Editorial. It were to be wished the flaws were fewer... but the main thing is, does it hold good measure? *Public Health* 2002;**116**:251.
- 2 **Green LW**, Lewis FM. *Validity of measurement. Measurement and evaluation in health education and health promotion*. Palo Alto, CA: Mayfield, 1986:101–15.
- 3 **Guilford JP**, Fruchter D. *Fundamental statistics in psychology and education*. 5th Ed. New York: McGraw-Hill, 1973.
- 4 **Langley J**, Stephenson S, Cryer C. Measuring road traffic safety performance: monitoring trends in non-fatal injury. *Traffic Injury Prevention* 2003;**4**:291–6.
- 5 **Cryer PC**, Jarvis SN, Edwards P, et al. Why the government was right to change the "Our Healthier Nation" accidental injury target. *Public Health* 2000;**114**:232–7.
- 6 **Cryer C**, Langley JD, Stephenson SCR, et al. Measure for measure. The quest for valid indicators of non-fatal injury. *Public Health* 2002;**116**:257–62.
- 7 **Campbell SM**, Braspenning J, Hutchinson A, et al. Improving the quality of health care. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003;**326**:816–9.
- 8 **Cryer PC**, Jarvis SN, Edwards P, et al. How can we reliably measure the occurrence of non-fatal injury? *International Journal of Consumer Product Safety* 1999;**6**:183–91.
- 9 **Langley J**, Cryer C. Indicators for injury surveillance. *Australasian Epidemiologist* 2000;**7**:1, 5–9.
- 10 **Gutoskie P**. *Canada's road safety targets to 2010*. Ottawa: Transport Canada, Minister of Public Works and Government Services, 2001.
- 11 **Land Transport Safety Authority**. *New Zealand road safety programme 2001–2002*. Wellington: LTSA, 2001.
- 12 **Department of the Environment Transport and the Regions**. *Tomorrow's roads: safer for everyone. The government's road safety strategy and casualty reduction targets for 2010*. London: DETR, 2000.
- 13 **Langley JD**, Dow N, Stephenson SCR, et al. Missing cyclists. *Inj Prev* 2003;**9**:376–9.
- 14 **Haigney D**. STAS19. *Journal of the Institute of Road Safety Officers* 1995;October:11–17.
- 15 **Cryer C**, Westrup S, Cook A. *Injury prevention through data linkage. Phase 3: the linkage of hospital inpatient data to police road traffic accident reports in East and West Sussex*. Canterbury: CHSS, University of Kent, 2000.
- 16 **Harrison J**, Steenkamp M. *Technical review and documentation of current NHPA injury indicators and data sources*. Canberra: Australian Institute of Health and Welfare, 2002.
- 17 **McClure RJ**, Peel N, Kassulke D, et al. Appropriate indicators for injury control? *Public Health* 2002;**116**:252–6.

## LACUNAE

### Australian injury collaboration

A collaboration of leading injury researchers from Sydney, Australia, has been awarded a \$2.35 million Capacity Building Grant in Population Health Research from the National Health Medical Research Council (of Australia) for *Addressing injury in a population health framework—an integrated approach to prevention, acute care, and rehabilitation*. This partnership involves both established researchers and new postdoctoral researchers at a number of research sites across both the University of New South Wales (UNSW) and the University of Sydney: NSW Injury Risk Management Research Centre, UNSW, The George Institute, University of Sydney, The Prince of Wales Medical Research Institute, UNSW; School of Public Health and Community Medicine, UNSW and the Rehabilitation Studies Unit, University of Sydney. This partnership brings together a range of researchers from multiple disciplines to provide training in the skills suitable to respond to the growing challenges and complexity posed by *injury prevention, acute care, and rehabilitation*. A major strength is the integration of both clinical and biomedical researchers with public health researchers. The program of work addresses one of the Australian National Health Priority Areas: *injury prevention and control* and a number of its current nominated national priorities: injury prevention in indigenous communities; workforce development; the national injury prevention plan; as well as addressing other areas of national concern, such as falls, sports, and road safety issues. The program has been built around clearly defined injury domains and across five thematic areas. The domains comprise specific population subgroups (indigenous, young adults, and older persons), as well as "context of injury" settings (road traffic, falls, sport, and work). The five themes reflect the areas of greatest need for capacity building in population health research applied to *injury prevention, acute care, and rehabilitation*, and for the application and translation of such research:

- (1) Integration and interrogation of data systems to underpin population health research in injury prevention, acute care, and rehabilitation.
- (2) Understanding the nature of injury risk.
- (3) Systematic development and evaluation of injury prevention interventions.
- (4) Quality provision of trauma care systems and rehabilitation services.
- (5) Translating injury research into policy and practice.

Further information can be obtained from Caroline Finch (NSW Injury Risk Management Research Centre), [c.finch@unsw.edu.au](mailto:c.finch@unsw.edu.au), or Mark Stevenson (The George Institute), [mstevenson@thegeorgeinstitute.org](mailto:mstevenson@thegeorgeinstitute.org).