Chapter 14
Validity of classification of scenarios

14.1 INTRODUCTION

As part of the initial and on-going training of research officers and assessment of their coding accuracy during the data collection phase of the study, brief scenarios describing various death circumstances were developed. These were coded by the research officers. This section describes this aspect of the study.

The aim of the work presented here was to:

- determine the level of accuracy of the classification of scenarios by research officers;
- identify circumstances that were easy or difficult to classify validly; and
- determine if there was any change in the level of accuracy over time.

14.2 METHODS

Three hundred and thirty three brief scenarios describing various circumstances in which people died were prepared. Most, but not all, of these described incidents in which one or more people were fatally injured. Although fictitious, the scenarios covered the type of circumstances commonly described in coronial files. The scenarios were developed to assist in the training and monitoring of the classification reliability of the research officers. Some were based on actual events that occurred before or after the study period, and some were fictitious. They were designed to cover most conceivable circumstances in which someone might be fatally injured in relation to work of some sort ("Cases"), circumstances that were not related to work ("Non-cases"), and other circumstances that were ambiguous or lacking detail as to their

relationship to work ("Indeterminates"). Within the Case and Indeterminate groups, a number of specific categories of circumstances, which were of interest to the study, were covered. These were "Worker", "Commuter", "Volunteer", "Student", "Home Duties", "Workplace Bystander", "Road Bystander" and "Other Farm". In brief, the intent of the definitions behind each of these groups were as described in Chapter 3:

Worker:	working for pay, profit or kind;
Commuter:	worker traveling to or from the place of work;
Volunteer:	performing unpaid duties under some sort of formal arrangement;
Student:	undergoing unpaid instruction in connection with formal primary, secondary, tertiary or directly vocational education;
Home Duties:	performing unpaid domestic duties in one's own home (or the home of someone else);
Workplace Bystander:	non-worker fatally injured as a result of workplace activities, generally not associated with public roads or public transport;
Road Bystander:	non-worker fatally injured in a motor vehicle incident on a public road (or on public transport) as a result of other people's work, where the working vehicle was primarily 'at fault' in the incident; and
Other Farm:	person injured on a farm but who didn't meet the definition of any of the other categories.

The definition of each of these groups is covered in more detail in Chapter 3.

For each of these categories, and the "Non-case" category, a number of scenarios were developed, some which clearly met the study definitions, and some of which were ambiguous for one or more reasons (sometimes just through lack of information) and which were designed to be coded as "Indeterminate". Within some categories, several scenarios were used to describe each of a number of specific sub-categories (eg deaths on farms, child deaths, bystander deaths in motor vehicle incidents, natural cause deaths). Examples of the scenarios included:

"A man stops at a petrol station on his way home from work. As he goes to the shop to pay, he is struck by a car leaving the premises.";

"A garbage truck runs over the pay manager at a waste disposal centre."; and

"A woman drowns whilst surfing with friends at a popular public beach".

Using a stratified randomisation approach, the 333 scenarios were separated into three sets of 111 scenarios in random order, thereby producing three comparable sets of scenarios that were used at different stages of the data collection process for the study. The categories used for the coding were:

1 Worker

2 Commuter

3 Volunteer

4 Student

5 Home Duties

6 Workplace Bystander

7 Road Bystander

8 Other on-farm death ("Other Farm")

9 Uncertain

0 Not in any of the above categories (ie "Non-case").

For each scenario, two decisions were required and a corresponding two-digit code recorded. The first decision required the coder to choose the most appropriate category (from those shown above) to which the scenario should be assigned. If a scenario equally satisfied more than one category, coders were instructed to choose the higher category on the list (ie the one with the smallest number). This determined the second digit of the code. The second decision concerned the coder's certainty that, on the basis of the information provided, the scenario belonged in that category (coded as "definite" (= "1") or "indeterminate" (= "2")). This determined the first digit of the code. Non-cases were only assigned a zero code, rather than a "10" code.

The allocated code was recorded next to each scenario on specifically prepared sheets. The codes were then checked against the correct codes allocated by the author. Discrepancies were discussed with the research officers at the time, in order to identify (and correct) any differences in interpretation, or misunderstanding, of the definitions. The primary aim of this was to maintain (and improve if required) the accuracy of coding during the data collection, and to identify specific problem areas that needed to be checked once the initial data collection procedure was completed. The allocated codes were later entered electronically for analysis.

Of the 333 scenarios, 231 (69.4%) were Cases, 30 (9.0%) were Indeterminates and 72 (21.6%) were Non-cases. Within the group of Cases, the *Worker* (78: 23.4% of all files) and *Workplace Bystander* (33: 9.9%) categories were the largest. The stratified randomisation of scenarios between trials was satisfactory (Table 14.1).

Research officers were instructed to complete the three different sets of scenarios at three times during the data collection process – at the beginning, half way through, and

at the end. Most of the research officers completed the sets in this way, although the second set was not always completed exactly half way through. Different collection periods were required for different jurisdictions. Also, some research officers were recruited after the collection had commenced. Therefore, the period between completion of scenarios differed between research officers in different jurisdictions, and between some research officers in the same jurisdiction.

All 21 research officers were included in the comparison (one additional officer who left the study after two weeks was not included). For various reasons, not all research officers completed all trials. Twenty research officers completed Trial One, 18 completed Trial Two and 18 completed Trial Three. Thirteen research officers completed all three trials and all but one completed at least two trials. In addition, four coders mistakenly did not allocate codes to one or more scenarios. Eleven scenarios consequently had missing codes. This left a total of 6,094 scenarios (comprising 333 different scenarios that were each coded by up to 21 persons) for which research officer codes were allocated. These 6,094 scenarios are the basis of the presented analysis. The non-completed scenarios did not bias the randomisation of scenario types between trials (see Table 14.1).

Table 14.1 Comparison of codes assigned to scenarios – as initially devised and as coded – by category

Code	Case category	Scenarios as initially devised: Trial 1	Trial 2	Trial 3	Total[1] n	Total[1] %	Actual total[2] %
Case							
11	Worker	25	25	28	78	*23.4*	*23.4*
12	Commuter	7	7	7	21	*6.3*	*6.3*
13	Volunteer	7	7	8	22	*6.6*	*6.6*
14	Student	8	6	6	20	*6.0*	*6.0*
15	Home duties	7	6	5	18	*5.4*	*5.4*
16	Workplace bystander	11	9	13	33	*9.9*	*9.9*
17	Road bystander	5	6	7	18	*5.4*	*5.4*
18	Other on-farm death	4	4	6	14	*4.2*	*4.2*
19	Uncertain	3	2	2	7	*2.1*	*2.1*
	All cases	77	72	82	231	*69.4*	*69.4*
Indeterminate							
21	Worker	0	3	0	3	*0.9*	*0.9*
22	Commuter	0	0	0	0	-	-
23	Volunteer	2	0	0	2	*0.6*	*0.6*
24	Student	1	1	0	2	*0.6*	*0.6*
25	Home duties	0	0	0	0	-	-
26	Workplace bystander	2	1	1	4	*1.2*	*1.2*
27	Road bystander	2	3	1	6	*1.8*	*1.8*
28	Other on-farm death	0	0	0	0	-	-
29	Uncertain	3	7	3	13	*3.9*	*3.9*
	All indeterminates	10	15	5	30	*9.0*	*9.0*
Non-case							
0	Non-cases	24	24	24	72	*21.6*	*21.6*
Total							
	All-scenarios	111	111	111	333	*100.0*	*100.0*
Specific features[3]							
	Child involvement	18	16	15	49	*14.7*	*14.8*
	Farm involvement	15	16	16	47	*14.1*	*14.1*
	MV involvement[4]	28	34	27	89	*26.7*	*26.7*
	Bystander[5]	35	42	43	120	*36.0*	*35.9*

1: All scenarios as initially devised and given to each research officer for coding.
2: Total scenarios that were coded by the research officers (n=6094).
3: Specific aspects of the scenarios.
4: Scenarios involving motor vehicle incidents.
5: Scenarios where exposure to another's person's work activity led to fatal injury (the correct code might not have been a bystander code because a higher category might have been more appropriate).

For the analysis, comparison was made with the exact code, of which there were 19 possible choices (nine Case codes, nine Indeterminate codes and a Non-case code), and with the three general codes of Case, Indeterminate and Non-case. Reported codes for the scenarios are those assigned by the author (the gold standard), unless otherwise stated. Since each scenario was allocated a code by up to 20 of the 21 research officers, any one scenario could contribute information up to 20 times. Unless otherwise stated, the percentages are based on all 6,094 coded scenarios.

Sensitivity, specificity and accuracy were the main measures of interest. They are presented as proportions. The definitions for these terms were:

sensitivity: the proportion of all scenarios in a category that were correctly assigned to that category;

specificity: the proportion of all scenarios not in a category that were correctly assigned not to that category;

accuracy: the proportion of all scenarios assigned to the correct category.

14.3 RESULTS

OVERALL COMPARISON

Comparing the broad categories, 5,392 (88.5%) scenarios were classified correctly, and 164 (2.7%) were Cases coded as Non-cases or vice versa. The remaining discordant scenarios were Cases coded as Indeterminate (and vice versa – 326: 5.3%) and Indeterminates coded as Non-cases (and vice versa – 212: 3.5%). The sensitivity and specificity were, respectively, 0.92 and 0.93 for Cases, 0.64 and 0.94 for Indeterminates, and 0.87 and 0.95 for Non-cases (Table 14.2). Taking into account all 19 possible codes, 4,779 (78.4%) scenarios were classified correctly.

Table 14.2 Comparison of codes assigned to scenarios: broad categories

		Research officer code			
		Case	**Indeterminate**	**Non-case**	**Total**
Truth	**Case**	3,877	247	103	**4,227**
	Indeterminate	79	350	119	**548**
	Non-case	61	93	1,165	**1,319**
	Total	**4,017**	**690**	**1,387**	**6,094**

On an individual research officer basis, accuracy at the broad category level ranged from 0.80 to 0.95, and at the individual category level from 0.64 to 0.92. For the jurisdictions, accuracy at the broad level ranged from 0.87 to 0.95, and at the individual category level from 0.74 to 0.91.

WORKING CASES ONLY

Since the primary focus of the overall study was on the work-related deaths of workers, mis-classification of the *Worker* category was the main area of interest and concern in the comparison. For the 1,429 scenarios that described *Worker* deaths, the sensitivity was 0.85 and the specificity 0.99. The overall accuracy for the classification of deaths as working or not was 0.96 (Table 14.3).

Table 14.3 Comparison of codes assigned to scenarios: working cases only

		Research officer code		
		Case	**Non-case**	**Total**
Truth	**Case**	1,217	212	**1,429**
	Non-case	39	4,626	**4,665**
	Total	1,256	4,838	**6,094**

OTHER CASE CATEGORIES

Of the other specific Case categories, the sensitivity ranged from 0.62 (for *Workplace Bystanders*) to 0.91 (for *Commuters*), the specificity was 0.97 or higher for all categories, and the accuracy 0.94 or higher for all categories (Table 14.4).

Table 14.4 Comparison of codes assigned to scenarios: sensitivity, specificity and accuracy for cases

Case category	Number	Sensitivity	Specificity	Accuracy
Worker	1,429	0.85	0.99	0.96
Commuter	385	0.91	1.0	0.99
Volunteer	401	0.85	0.98	0.99
Student	367	0.86	0.99	0.98
Home Duties	331	0.83	0.99	0.98
Workplace Bystander	602	0.62	0.98	0.94
Road Bystander	327	0.81	0.99	0.98
Other Farm	256	0.86	0.97	0.96
Uncertain	129	0.26	0.99	0.97

There was no evidence of a consistent trend during the data collection period in the sensitivity of classification for any of the categories. For most Case categories, the sensitivities varied within a range of 0.82 to 0.93 (Table 14.5).

The sensitivity for road bystander scenarios was much less in the third trial. Inspection of the involved scenarios showed that nearly all of the difference was due to four scenarios, which described incidents in which emergency service vehicles struck and killed non-working persons, being coded by some research officers as *Workplace Bystander*. The correct code was *Road Bystander*, but during the data collection there was discussion as to whether this sort of circumstance would be better coded as *Workplace Bystander*. The research officers in the larger states, who were still involved in data collection when this discussion took place, were instructed to change their coding approach accordingly (although later, during coding by the study team, this

decision was reversed). It was these research officers who allocated the *Workplace Bystander* codes to these four scenarios, although not entirely consistently. This change in coding approach was not taken into account in the analysis presented here, as it is not clear exactly when the research officers were instructed to change their coding approach. Therefore, the apparent decrease in the sensitivity is partly due to the uncertainty caused by the changed coding instructions, and partly an artefact of the analysis not being able to appropriately take the change into account.

In contrast, the sensitivity for Cases of uncertain category was particularly poor for the first trial. This was due to three scenarios where nearly all research officers applied a specific code or indeterminate code to scenarios with circumstances that clearly qualified as a Case, but that were ambiguous enough to make it difficult to determine the correct category. Learning when to accept a given amount of information as sufficient to identify a particular category (or to rule out other categories) is a difficult aspect of coding. So, it is not surprising that this was not done well when the research officers were first becoming familiar with the definitions.

Table 14.5 Sensitivity of classification of case scenarios by trial

Case category	Trial 1	Trial 2	Trial 3	Total
Worker	0.86	0.82	0.88	0.85
Commuter	0.85	0.92	0.95	0.91
Volunteer	0.83	0.93	0.80	0.85
Student	0.82	0.89	0.87	0.86
Home Duties	0.86	0.81	0.81	0.83
Workplace Bystander	0.60	0.52	0.71	0.62
Road Bystander	0.92	0.92	0.64	0.81
Other Farm	0.84	0.86	0.86	0.86
Uncertain	0.05	0.47	0.36	0.26

SOURCES OF DISAGREEMENT

Since the scenarios were brief, the research officers did not have to search hard for the key information. The main sources of disagreement for wrongly coded scenarios seemed to be difficulty deciding where the boundaries lay between Case and Indeterminate, and between different categories.

For working scenarios, there were 52 different scenarios for which at least one coder allocated a different code (212 wrongly coded scenarios in total). The most common different categories assigned by the coders were *Other Farm* (19.3%), *Indeterminate Worker* (19.3%), *Uncertain Case* (12.3%), *Uncertain Indeterminate* (12.3%) and *Student* (11.8%). The main types of circumstances which caused problems were business trips, hobby farms, elderly farmers, apprentices at TAFE classes and other workers training in their occupational area, murdered workers and farmers exposed to farm hazards (such as weed-killer or snakes) where no obvious working activity was mentioned. The scenarios for which five or more coders assigned a non-working code, the codes they assigned, and comments on the mis-classifications, are shown in Table 14.6.

There were 27 different non-working scenarios coded by at least one research officer as *Worker*, but only six non-working scenarios coded by two or more research officers as *Worker*. These involved domestic disputes, a non-injury death, unpaid training not arising out of the person's current vocation, and ad-hoc/informal work activity that had possibly stopped by the time of the injury. The non-working scenarios coded as *Worker* by two or more research officers are shown and discussed in Table 14.7.

Most of the other categories had a few specific areas of difficulty. For *Commuters*, persons traveling to an official "after-hours" work function (coded by the research officers as *Worker*), persons who had a minor break or deviation in the journey to or from work (coded as *Indeterminate Commuter*), and persons who would otherwise have been classified as workplace or road bystanders (coded as *Workplace Bystander* or *Road Bystander*), all lead to some difficulties. For *Volunteers*, problems arose with circumstances where the official status of the volunteer work was not clear (coded as *Indeterminate Volunteer*). The biggest problem for *Workplace Bystanders* was with farm exposures, such as children drowning in dams and persons injured by farm equipment or machinery (coded as *Other Farm*), persons injured as bystanders to Home Duties (coded as *Non-cases, Home Duties* or *Indeterminate Home Duties*), and persons performing unpaid training not arising out of their current vocation who died as a result of the actions of their paid instructor (coded as *Student*). The main problem with the *Road Bystander* category, as discussed above, was the coding of incidents involving emergency vehicles as *Workplace Bystanders,* following some changed coding instructions towards the latter part of data collection. The *Uncertain Case* category scenarios were not well coded, but there was no clear pattern in the mis-coding. The *Student, Home Duties* and *Other Farm* categories were generally well coded and had no clear pattern in the mis-coding.

14.4 CONCLUSIONS

The comparison of classification of scenarios showed a high level of agreement between the true classification (as determined by the author) and the classification assigned by the research officers. This was the case for all categories, except for the *Work Bystander* and *Uncertain Case* categories.

The comparison provides good evidence that deaths were being appropriately classified and that there was no important drift in the approach to coding by the research officers during the data collection phase of the study. The process of coding the scenarios and discussing those causing disagreement helped to identify and correct problems areas, which was the main function of this aspect of the study.

The main problem areas were found to be determining the strength of evidence required before a definitive code could be assigned; and the allocation of deaths to a case category when the circumstances had elements of more than one category.

Table 14.6 Scenarios describing working circumstances that were coded by five or more research officers as non-working: description, research officer codes and comment

	Description	ROs' Codes[1]	Comment
1 – 42	A business executive is murdered walking back from a restaurant to her hotel during an interstate business trip.	"21" * 4; "22" * 3; "0" * 5	This is a *Worker* case on the basis that virtually all exposures on a business trip are accepted as arising from work.
1 – 48	A farmer is fatally poisoned after accidentally drinking weed killer from a lemonade bottle during a break in shearing.	"18" * 7	A farmer, on his/her farm, is considered to always be working unless there is specific information to conclude otherwise. This includes work breaks. Also, the weed killer is almost certainly a hazard arising from farm work.
1 – 49	A 75 year old man dies when his clothing becomes caught in the power take-off of a tractor he is operating.	"19" * 3; "18" * 3; "21"; "28" ; "29" * 2	The scenario implies that the man was working, but he could also have been performing home duties. Therefore, "*19: Uncertain Case*" is a reasonable code, but "*11: Worker*" was deemed slightly more appropriate.
1 – 63	An oyster farmer dies when his dinghy capsizes in rough weather as he returns from tending his oyster beds.	"12" * 5	The oyster farmer should be considered to be at work until he returns to his base of work. The water and rough weather are clearly hazards that arise directly as part of his work.
1 – 92	A doctor dies in a tractor rollover incident whilst working on his hobby farm.	"18" * 9	Hobby farmers were specifically included as workers by the study definitions.
2 – 6	A roofer dies when he falls whilst replacing asbestos-cement roof sheeting.	"19" * 4; "21" * 2	The scenario strongly implies that the person was working, although this is not specifically stated. Therefore, "*19: Uncertain Case*" is a reasonable code, but "*11: Worker*" was deemed more appropriate.
2 – 17	A 50 year old man dies when his jumper catches in the power take-off of an auger he is operating.	"19" * 3 "29" * 3	As for 1 – 49.
2 – 48	A footballer dies as a result of a head injury received during a game. He is paid $100- per week in "petrol money" by the Club.	"21" * 3; "0" * 3	The injury clearly resulted from an activity for which the man was being paid. The informal nature of the payment doesn't change the fact that he was receiving payment for his services and therefore working.

Table 14.6 continued

	Description	ROs' Codes[1]	Comment
2 – 55	A 20 year old apprentice fitter and turner is fatally injured when struck by a steel rod that flies from a lathe while working on a class project at TAFE.	"14" * 8	TAFE classes are part of an apprentice's formal training for which payment is received. Apprentices were specifically included as workers by the study definitions.
2 – 57	A tourist from the United States picking cherries dies when he is run over by a tractor hauling collecting bins.	*18" * 4; "19" * 2	As for 2 – 6. The fact that the person was an overseas national is not relevant if they were injured in Australia.
2 – 59	A 36 year old self-employed pilot is killed while retraining for a crop dusting certificate.	"14" * 8; "19"; "0"	This training arises directly from the pilot's current vocation, rather than being an attempt to start a new occupation which has not yet generated income. So, it is best seen as resulting from work exposures.
2 – 63	An apprentice carpenter dies when he is struck in the chest by a piece of wood while working on a lathe during a TAFE class.	"14" * 8	As for 2 – 55.
2 – 87	A motor cycle police officer dies when he loses control of his bike at high speed and collides with a tree.	"21" * 3: "29" * 3	As for 2 – 6.
2 – 88	A man dies as he returns from his back paddock when he steps from his truck and is bitten by a tiger snake.	"18" * 5; "19 * 2; "29" * 2	As for 1 – 48.
3 – 20	The body of a taxi driver is found bound and gagged in the boot of his burnt out taxi.	"21" * 8; "29" * 2	As for 2 – 6. Involvement in criminal activity is not ruled out, so "*21: Indeterminate Worker*" would be a reasonable code, but the connection to work is strongly implied.
3 – 23	A man sitting in the rear of a utility at the back of his farm is struck in the head by the side of a semi-trailer delivering sheep to his property.	"18" * 2; "19" "21" * 2; "29"	As for 2 – 6.

	Description	ROs' Codes[1]	Comment
3 – 49	On a dinner organised by the project manager, a consultant research officer is killed as she crosses a busy road.	"12" * 5	A social function directly connected to work is considered work activity. There is some ambiguity in the scenario, but the information strongly suggests that the person was walking as part of the work meeting. If she was travelling from home to the meeting when the incident occurred, she would be properly classified as a commuter.

Table 14.6 continued

	Description	ROs' Codes[1]	Comment
3 – 50	A man is murdered by an employee when the victim refuses to grant the employee the pay rise he has been promising him.	"21" * 3; "0" * 2	There is a clear and direct connection between the murdered man's work activities/responsibilities and the assault by the employee. It is not necessary that the incident occur at work for it to be properly considered to be work-related.
3 – 86	A nun dies when she falls from a ladder while cleaning the windows in the cloisters.	"13"; "15" * 5	Nuns were specifically identified in the study definitions as being considered to be working when involved in activities related to their status as a nun.

1: Identifies the wrong codes applied and the number of times each code was applied.

Table 14.7 Scenarios describing non-working circumstances that were coded by two or more research officers as worker: description, correct code, number of research officers assigning a working code and comment

	Description	Correct code	RO number[1]	Comment
1 – 15	A man dies of a stroke while working at his computer in his office.	0	2	The death was not due to external factors. Therefore, the person is not a Case, regardless of whether there is a connection to work.
1 – 109	A medical student doing a rotation in an Emergency Department is killed when run over by an ambulance in the hospital driveway.	14	6	The student is clearly not being paid and so should be classified as a *Student*. He would also qualify as a *Workplace Bystander*, but the *Student* code is higher and so should be selected.
2 – 84	A man is brutally murdered by his wife because, she says, he spends too much time at his computer working on his "damned Work-Related Fatalities Study".	0	2	This is a domestic dispute and not related to work exposures as they are generally understood. Therefore, the person is not a Case.
2 – 94	A man selling very cheap videos in a pub is assaulted and robbed as he leaves the pub.	21	3	This is a difficult situation to classify. The connection to his informal work is suggested, but unclear enough to warrant an *Indeterminate* code.
2 – 97	A public servant engaged in a torrid affair at work is shot in her office by her jealous husband.	0	2	As for 2 – 84.
3 – 62	An unpaid trainee geologist gaining work-experience, who has only been in Australia for a few weeks, is crushed by a large rock while taking samples on a mining reconnaissance expedition.	16	3	This is a difficult situation to classify. There is no suggestion of payment of any sort. Therefore, the person is not working. There is also no strong evidence that he is part of a formal training program. Therefore, he is not a student. The injuries were definitely connected to the work exposures of he reconnaissance team. Therefore, Workplace Bystander is the most appropriate classification.

1: Number of research officers who coded the scenario as *Working*.