Pitfalls in the Analyses of Accident Records: The Meaning of Missing Values

Bob Uttl[1,*] and Kelly Kisinger[2]
[1] Mount Royal College, Calgary, AB, Canada
[2] University of Calgary, Calgary, AB, Canada

ABSTRACT: Accident records are frequently used to learn about conditions and behavior before, during, and after accidents. However, accident records frequently do not mention the conditions and behavior of interest (e.g., thaw or melting snow) leading to missing values and to uncertainty about their presence or absence. In the development of the Avaluator Avalanche Accident Prevention Card, Haegeli & McCammon (2006) analyzed avalanche accident records for the presence or absence of obvious clues to avalanche danger (e.g., avalanches, thaw) to determine the prevention values of limiting one's travel to slopes with a certain number of clues. However, they were unable to determine the presence or absence of all clues in 1,148 accidents out of 1,400 and deleted them from their data set. Thus, the Avaluator's Obvious Clues are based on only 252 selected accidents. The critical question arises: What is the meaning of these missing values? If the missing values occurred because eyewitnesses did not mention absence of absent clues, the missing values mean absence and the exclusion of accidents was inappropriate. To investigate the meaning of missing values, study participants were shown video clips of accidents and later asked to recall as much information as they could remember about the accident. The results show that eyewitnesses are far more likely to report the presence of present conditions and behaviors (e.g., snow) than the absence of absent conditions and behaviors (e.g., absence of snow). The results highlight that the exlusion of 82% of all accidents by Haegeli and McCammon (2006) prior to the computation of the Obvious Clues prevention values was inappropriate and resulted in inflated prevention values.

KEYWORDS: Avaluator, Obvious Clues, accident records, missing values, avalanche prevention.

1 INTRODUCTION

Accident records are frequently used to learn about conditions and behavior before, during, and after accidents. However, accident records frequently do not mention the conditions and behavior of interest (e.g., thaw) leading to missing values and to uncertainty about their presence or absence. The development of the Avaluator Avalanche Accident Prevention Card (Haegeli & McCammon, 2006) highlights this problem.

Haegeli and McCammon (2006) analyzed over 1,400 avalanche accidents and, based on their analyses of these accident records, developed the Avaluator Avalanche Accident Prevention Card to reduce the number of avalanche accidents in Canada. Specifically, they reviewed records of avalanche accidents and for each record they determined whether each so-called Obvious Clue (avalanches, loading, path, terrain trap, rating, unstable snow, thaw instability) was present, absent, or indeterminate from the accident record. Next, they summed up the number of Obvious Clues to obtain the frequency of avalanche accidents that occurred when 0 to 7 Obvious Clues were present. Finally, using the frequency distribution of the obvious clues they calculated the percentage of accidents prevented if users limited themselves to a certain number of clues (i.e, relative risk reduction) (see Uttl, Uttl, & Henry, 2008, for the review of the Avaluator and the Obvious Clues method). Thus, to evaluate local conditions, the user of the Avaluator adds up the number of Obvious Clues present and the Avaluator informs the user what proportion of historical accidents would have been avoided if people had limited themselves to slopes with a given number of clues. For example, the Avaluator claims that the “4 or less (sic)” clues limit would have prevented 77% of historical accidents (i.e., 77% relative risk reduction).

However, as discovered by Uttl, Uttl, & Henry (2008a,b; see also Uttl, Henry, & Uttl, 2008) and later confirmed by Haegeli as well as McCammon (personal communication), Haegeli and McCammon (2006) could not determine the presence or absence of each of the seven clues in over 82% of the accidents and simply deleted the accidents with missing values from their data set. Thus, the resulting clue distribution and the associated prevention values published in the Avaluator are based on only 252 accidents out of > 1,400. Haegeli and McCammon (2006) never disclosed to Avaluator users that the prevention values of the obvious clues are based on

Corresponding author address: Bob Uttl, Psychology Department, Mount Royal College, 4825 Mt. Royal Gate SW, Calgary, AB, T3E 6K6, Canada; email: buttl@mtroyal.ca; web: www.docbob.ca.

only a small fraction of the >1,400 accidents they claimed were used for their derivation in the Avaluator booklet. At this time, over 20,000 of the Avaluators have either been purchased off the shelf or distributed through the Canadian Avalanche Association safety training courses to unsuspecting winter backcountry recreational users in Canada.

Several questions arise: What is the effect of eliminating 82% of accident records on the resulting statistics and prevention values? What is the meaning of these missing values? Do they mean absence of the relevant condition (e.g., avalanches, loading, thaw) or are they merely a by-product of some random process? If the missing values occurred due to some purely random process, the deletion of even 82% of all accidents may not necessarily bias the resulting distribution of the obvious clues and the associated prevention values. However, if the missing values occurred because, for example, accident victims and eyewitnesss are more likely to report the presence than absence of the obvious clues, the deletion of 82% of the accidents would shift the distribution of the obvious clues towards more clues and result in inflated prevention values (Uttl, Uttl, & Henry, 2008). In turn, the inflated prevention values would give Avaluator users a false sense of confidence in the stability of slopes they are about to cross and lead to more rather than fewer avalanche accidents, injuries, and deaths (see Uttl, Uttl, & Henry, 2008; Uttl, Henry, & Uttl, 2008a,b). Unfortunately, Haegeli and McCammon neither considered the meaning of missing values in the accident records nor described their distribution (see Uttl, Uttl, & Henry, 2008, for the review of the Avaluator and the Obvious Clues method).

The question of the meaning of missing values is highlighted by a severe discrepancy between the distributions of the obvious clues reported on different occasions by McCammon and Haegeli (McCammon, 2002; McCammon, 2004; Haegeli and McCammon, 2006; see Uttl, Uttl, & Henry 2008 for details). McCammon (2004) did not exclude any accidents due to missing values and reported that the average number of obvious clues was 3.3 for the 715 accidents he analyzed. In contrast, for the Avaluator, Haegeli and McCammon (2006) excluded > 1,148 accidents due to missing values and reported that the average number of the obvious clues was much higher – 5.2 – in the remaining 252 accidents. Figure 1 highlights the discrepancy between the obvious clues distributions reported by McCammon (2004) and Haegeli and McCammon (2006) and between the associated prevention values. In response to Uttl, Uttl, and Henry's inquiries into these discrepancies, Haegeli and McCammon confirmed that the Avaluator is indeed based on only 252 accidents rather than 1,400 but they have repeatedly refused to provide details about their methods, details on the 252 accidents that remained in their analyses for the Avaluator, and access to their data for the limited purpose of verifying their claims in the Avaluator (see Uttl, Uttl, & Henry, 2008).

Unable to obtain access to the Avaluator data, Uttl, Henry, and Uttl (2008a,b) attempted to replicate the Avaluator's prevention values and to establish the meaning of the missing values from external weather and avalanche bulletin data. They found the Avaluator's prevention values to be grossly inflated; moreover, external weather data confirmed that missing values for the thaw and unstable snow clues meant the clues were actually absent. In response to Uttl et al.'s findings, the Canadian Avalanche Center (CAC) also attempted to obtain access to the Avaluator data from Haegeli and McCammon, and, when they also failed, they commissioned another attempt to replicate the Avaluator's prevention values. Working for the CAC, Floyer (2008) was also unable to replicate the Avaluator's prevention values and found them to be inflated even though, following Haegeli and McCammon, he also inappropriately deleted 71% of his accident reports due to missing values.

Figure 1 shows the distribution of the obvious clues in US avalanche accidents reported by the above investigators when accidents with missing values are either included or excluded and the corresponding percentage of accidents prevented if recreational backcountry users limited their travel to conditions with the specific number or fewer clues. Three studies (McCammon, 2004; Uttl, Henry, & Uttl, 2008; Floyer, 2008) that have reported the distribution of obvious clues based on all accidents in their sample found nearly identical distributions of the obvious clues. In contrast, the two studies that excluded accidents due to missing values – the Avaluator (82% of accidents excluded) and Floyer (71% of accidents excluded) – reported obvious clue distributions that are markedly shifted towards a higher number of clues and prevention values that are much higher relative to the obvious clues distributions and prevention values reported by the studies that did not exclude records with missing values. Moreover, the Avaluator's prevention values are inflated, even relative to Floyer data, after exclusion of all records with the missing values.

The substantial differences between distributions when accidents with missing values are included vs. excluded strongly suggest that the missing values did not occur due to a random process but rather were caused by victims, rescuers, and eyewitnesses not reporting the absence of the obvious clues (Uttl, Henry, & Uttl,

2008; Uttl, Uttl, & Henry, 2008). For at least two of the obvious clues, this interpretation is supported by Uttl, Henry, and Uttl's (2008) findings based on weather and avalanche bulletin data showing that when avalanche accident records did not mention anything about the presence or absence of the thaw and unstable snow clues, the clues were actually absent. In combination, these findings suggest that eyewitnesses are far more likely to report the presence of obvious clues and are unlikely to report the absence of the obvious clues.

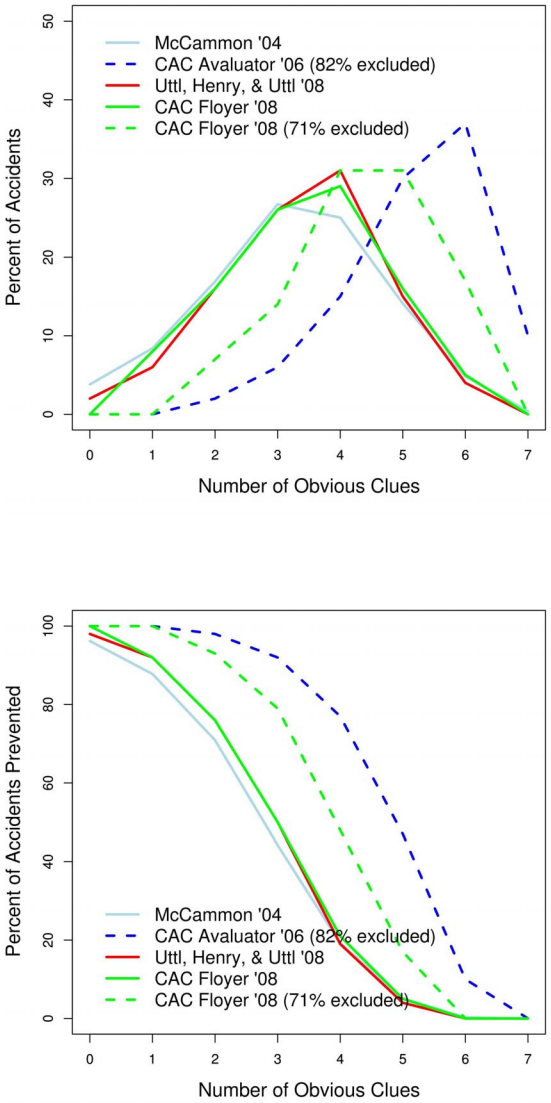

Figure 1. Distributions of Obvious Clues (top panel) and associated prevention values (bottom panel) when accidents with missing values are either included (solid lines) or excluded (dashed lines)

In the present study, we use car accidents as a model to examine the probability that eyewitnesses report the presence versus absence of obvious clue to accident danger. We chose to study car accidents for two reasons: First, high quality videos of car accidents are relatively easy to obtain from various movies. Second, at least some of the obvious clues to car accident danger are similar to the obvious clues for avalanche accident danger (e.g., snow, rain).

Consistent with Uttl, Henry, and Uttl (2008a,b) findings that eyewitnesses were likely to report the presence but unlikely to report the absence of obvious clues, we expected to find that eyewitnesses are far more likely to report the presence versus the absence of obvious clues to accident danger.

2 METHOD

Participants were 240 undergraduate students.

Eight car accidents were selected from movies: four with no rain and no snow clues (no clues), two with rain clue but no snow clue (rain clue), and two with snow clue but no rain clue (snow clue). All movie clips were presented in their original DVD quality with sound turned on.

In total, five obvious clues were considered: snow and rain clues (each present in two accidents) and three additional clues that appeared in at least two of the eight movie clips. The additional clues were: reduced visibility (visibility clue), failure to yield (yield clue), and being distracted while driving (distraction clue). The visibility clue was present in four accidents, the yield clue was present in five accidents, and the distraction clue was present in three accidents. One accident had none of the five clues present.

As part of a larger study lasting 1.5 to 2 hours, each participant watched a randomly assigned accident movie clip and immediately thereafter his or her memory for accident details was queried using undirected recall followed later by an accident questionnaire. For undirected recall, participants were instructed “to write down as much as you can remember about what happened in the movie clip as best as you can. Please be specific and provide as much detail as you can remember.” Later, the accident questionnaire examined participants' memory for various aspects of the accidents including driver, driver conditions, road condition and visibility, using checklists. Participants were asked to mark all items on checklists that applied to the accident they saw.

Accident records (recall protocols) were coded for the presence or absence of these “obvious clues” to accident danger (snow/ice, rain, poor visibility, failure to yield, distraction) using the scale developed by Uttl, Henry, and Uttl (2008): Yes = the clue was present, Weak Yes = the clue was probably present, DNK/Don't Know

= presence or absence of the clue is unclear from the record, Weak No = the clue was probably absent, No = the clue was absent.

3 RESULTS

Figure 2 shows the proportion of participants reporting the presence vs. absence of the five clues using the strict (Yes and No, top panel) and liberal (Yes+Weak Yes and No+Weak No, botton panel) criteria with error bars indicating 95% Confidence Intervals for the means. Participants were very likely to report the presence of the obvious clues and were very unlikely to report their absence.

Importantly, the responses on multiple choice questionnaire revealed that over 90% of participants noticed that the snow and rain clues were absent when they were absent. Accordingly, they knew the rain and snow clues were absent but chose not to report their absence.

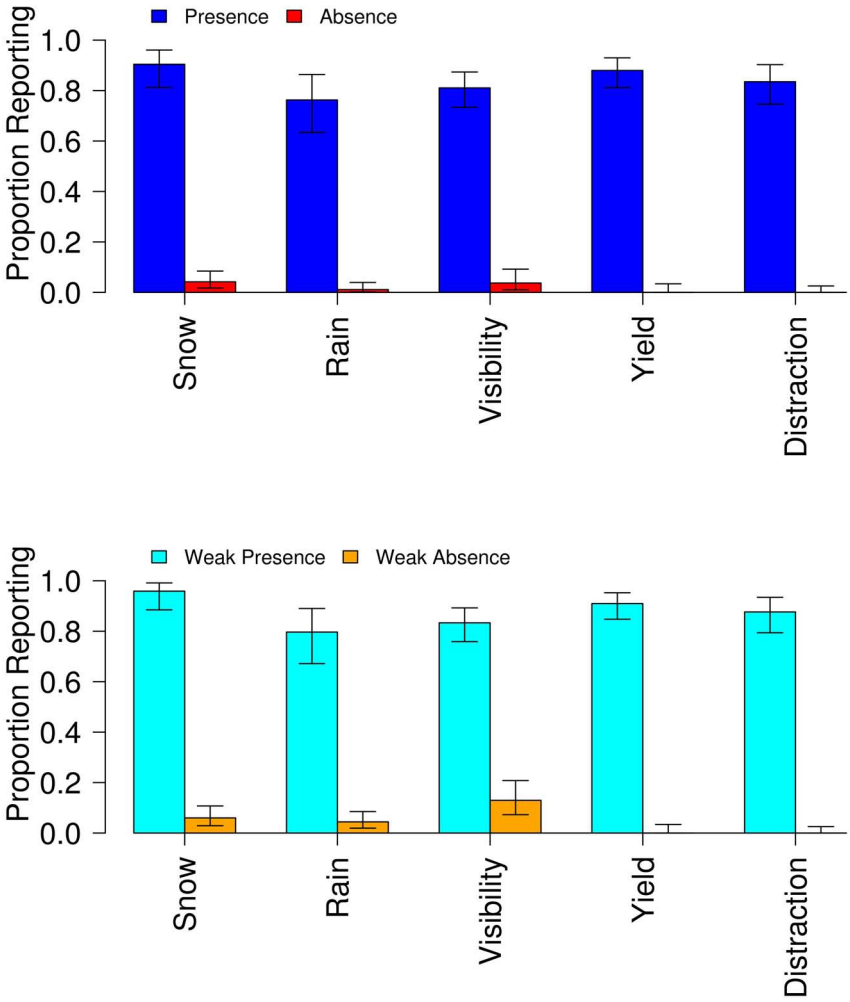

Figure 2. Reporting the presence vs. absence of the five clues. The figure shows the proportion of participants reporting the presence vs. absence of the five clues (snow, rain, visibility, yield, distraction) using the strict (Yes and No; top panel) and liberal (Yes+Weak Yes and No+Weak No; bottom panel) criteria. Error bars indicated 95% Confidence Intervals for the means.

4 DISCUSSION

This study revealed several important findings. First, eyewitnesses reported the presence of present obvious clues and only rarely reported the absence of absent obvious clues. In turn, the accident records themselves do not allow researchers to determine whether clues were present or absent in the vast majority of cases, resulting in many missing values. Second, the multiple choice test results revealed that the eyewitnesses were fully aware that the obvious

clues to accident danger were absent but chose not to report their absence. Thus, the failure to report the absence of the absent clues is not due to eyewitnesses not noticing their absence. Rather, one of the principal mechanisms causing missing values is eyewitnesses' tendency not to talk about the absence of obvious clues. Third, the vast majority of missing values occurred because the obvious clues were actually absent. In turn, the estimated prevalence of the obvious clues, under the assumption that the missing values mean the absence of the clues, was very close to the actual prevalence for all of the obvious clues.

The current results are consistent with the findings by Uttl, Henry, and Uttl (2008). Using external weather and avalanche bulletin data, they found that for at least the two clues – Unstable Snow and Thaw – the missing values meant that the clues were actually absent.

In combination, these findings indicate that the listwise deletion of 82% of accident records by Haegeli and McCammon (2006) in development of the Avaluator was inappropriate because the missing values were caused by the victims, eyewitnesses, and rescuers not reporting the absence of absent clues. In turn, the deletion of accidents with missing values grossly inflated the prevention values of the Obvious Clues published in the Avaluator. As a result, the Avaluator falsely tells users that they are much safer than their predecessors, encourages them to cross unsafe slopes, and, with all likelihood, facilitates recreational avalanche accidents.

Consistent with the grossly inflated prevention values published in the Avaluator, yearly numbers of recreational avalanche accidents in Canada have increased to the highest levels in at least the last 15 years (Uttl, Kibreab, Kisinger, & Uttl, 2009).

5 REFERENCES

Floyer, J., 2008. Review of the Obvious Clue Method. Report for Canadian Avalanche Centre, Revelstoke, BC, Canada: Applied Avalanche Research.

Haegeli, P., & McCammon, I., 2006. Avaluator Avalanche Accident Prevention Card. Revelstoke, BC: Canadian Avalanche Association.

McCammon, I., 2002. Evidence of heuristic traps in recreational avalanche accidents. Proceedings of the International Snow Science Workshop, Penticton, BC, Canada.

McCammon, I., 2004. Heuristic traps in recreational avalanche accidents: Evidence and implications. Avalanche News, 68, 1-11.

Uttl, B., Henry, M., Uttl, J., 2008a. Human Factors in Avalanche Avoidance and Survival. Canadian Society for Brain Behavior and Cognitive Science, Victoria, BC, Canada.

Uttl, B., Henry, M., Uttl, J.. 2008b. Avaluator's Obvious Clues Prevention Values: Are They Replicable? Proceedings of the International Snow Science Workshop, Whistler, BC, Canada.

Uttl, B., Uttl, J., & Henry, M., 2008. The Avaluator Avalanche Accident Prevention Card: Facts, Fictions, and Controversies. Proceedings of the International Snow Science Workshop, Whistler, BC, Canada.

Uttl, B., Kibreab, M., Kisinger, K, & Uttl, J., 2009. Trend Analysis of Canadian Avalanche Accidents: The Avaluator Avalanche Accident Prevention Card Has Not Reduced the Number of Accidents. Proceedings of the International Snow Science Workship, Davos, Switzerland.