ABSTRACT: We present a new avalanche accident prevention tool for recreational backcountry users: the 7CF. We evaluated the new tool's effectiveness on hundreds of avalanche accident records in the USA and Canada and found that the new 7CF tool has the equivalent risk reduction (sometimes called “prevention value”) to the Avaluator Accident Prevention Card (Haegeli & McCammon, 2006) used in Canada (see Uttl et al., 2008a,b; Uttl et al., 2009a,b,c,d). However, the new tool requires significantly less user knowledge and training and relies only on the easily recognized clues. Thus, in comparison to the Avaluator’s Obvious Clues Method (Haegeli & McCammon, 2006), the most significant advantages of the new 7CF tool include the reliability of clue detection and the ease of use. To illustrate, our research demonstrates that even users with no prior avalanche terrain experience are able to correctly recognize the presence and absence of all the 7CF clues with 99.9% accuracy. The 7CF tool is available for any interested users for free and is released under GNU General Public License, meaning that anyone is permitted to copy, change, and distribute the new tool. However, the 7CF tool is subject to some of the same limitations that plague all of the avalanche accident prevention tools developed and evaluated using the avalanche accident records and the risk reduction strategy. We discuss some of these limitations and illustrate them using the 7CF, the Avaluator’s Obvious Clues Methods, and other tools.

KEYWORDS: Avaluator Avalanche Accident Prevention Card, Obvious Clues, risk reduction, avalanche accidents

1. INTRODUCTION

Every year, hundreds of people around the world die in avalanches, often in pursuit of recreational experiences. In an attempt to reduce the number of accidents, injuries and deaths, governments and other agencies have funded the development of a variety of accident prevention programs and risk reduction tools such as Munter’s 3x3 (Munter, 2002) and Nivo (Bolognesi, 2007). In Canada, Haegeli and McCammon (2006) developed a “science-based” risk reduction tool called the Avaluator Avalanche Accident Prevention Card.

Avalanche risk reduction tools typically require users to input clues to avalanche danger such as slope angle, slope aspect, snow loading, presence/absence of weak layers, etc., and the tool returns a recommendation about avalanche risk in the form of “Go” vs. “No Go” advice or perhaps “Caution” and “Extra Caution” vs. “Not Recommended” (Avaluator).

The tools should be easy to use but should also provide valid diagnostic information regarding the risk that the slopes would avalanche.

Evaluation of a tool’s ease-of-use is simple; one can use the usual assortment of research methods for that purpose, including surveys, observations, and experiments. What is harder to measure is the diagnosticity of the risk reduction tools.

One approach to assessing the quality of an avalanche risk reduction tool is to ask what effect the use of the tool would have had on the number of accidents that are known to have occurred in its absence. Thus, a researcher examines historical accident records, applies a risk reduction tool, and determines what percentage of the historical accidents would have been prevented if the historical victims had used the tool.

The best example of this relative risk reduction approach is the Avaluator Avalanche Accident Prevention Card (Haegeli & McCammon, 2006) that provides users with very precise relative risk reduction values called “prevention values” by the Avaluator authors. To illustrate, Haegeli and McCammon claim that if historical victims had limited their travel across avalanche slopes when a maximum of four so called “obvious clues” were present, 77% of the accidents would not have happened.

Using the newly developed relative risk reduction tool called 7CF and the Avaluator example, we will demonstrate that the risk reduction approach by itself is not sufficient for...
evaluating the effectiveness of the avalanche risk reduction tools.

2. AVALUATOR: THE “BEST” DECISION TOOL

The Avaluator Avalanche Accident Prevention Card, developed by Haegeli and McCammon (2006), has been marketed by the Canadian Avalanche Centre and endorsed by Parks Canada as a “science-based decision-making tool” for helping amateur recreationists (e.g., skiers, snowmobilers, hikers) avoid avalanche accidents, and thus, possible injuries and deaths (Parks Canada Avalanche Awareness Initiatives, retrieved on September 5, 2010, www.ec.gc.ca). The Avaluator was first printed in September 2006, and an experiment with human participants evaluating the Avaluator’s effectiveness in preventing accidents was announced on October 3, 2006, by Dr. Ian McCammon on behalf of Haegeli et al. (2006). As of November 15, 2006, the Canadian Avalanche Centre (2006) made the Avaluator the primary decision making process taught to all students of avalanche safety training level 1 courses (AST 1) in Canada. In turn, AST 1 students form the bulk of participants in this ongoing experiment, without their knowledge or consent (Uttl et al., 2010).

The Avaluator consists of two components: the Trip Planner and Obvious Clues. Although the same problems plague the Trip Planner, here we focus on the Obvious Clues only. The Obvious Clues tool advises users about a specific slope’s stability; it helps users to “determine whether a slope is safe enough to cross (Haegeli & McCammon, 2006, p.14). The Obvious Clues is a checklist of seven clues to avalanche danger: avalanches (within 48 hours), loading (within 48 hours), path, terrain trap, rating, unstable snow, and thaw instability.

The user adds up the number of Obvious Clues applicable to a specific slope and the Avaluator makes one of three recommendations: proceed with “caution” (2 or fewer clues), proceed with “extra caution” (3 or 4 clues), or “not recommended” (5 or more clues). The Avaluator also includes “scientifically” derived accident prevention values (relative risk reduction values), that is, the percentage of historical accidents prevented if users had limited their travel to slopes with no more than a given number of Obvious Clues. For example, the Avaluator states that 77% of historical accidents would have been prevented if users had limited their travel to slopes with 4 or fewer clues. The Avaluator authors reviewed over 1,400 accidents (Haegeli & McCammon, 2006, p. 1), deleted over 1,148 (82%) records which did not have enough information to determine the status of all of the Obvious Clues, and derived the prevention values from the remaining 252 accidents (Uttl et al., 2008a,b; Uttl & Kisinger, 2010).

Subsequent research revealed a number of methodological flaws with the Avaluator; here we mention only the three most important flaws that render the tool unscientific. First, most critically, Haegeli and McCammon inappropriately deleted 82% of accidents due to missing data (Uttl et al., 2008a,b; Uttl & Kisinger, 2010).

Second, several attempts to replicate the Obvious Clues prevention values published in the Avaluator have shown them to be hugely inflated. As shown in Figure 1, the true prevention value of limiting one’s travel to 4 or fewer clues is only about 20% (Uttl et al., 2008b; Floyer, 2008; McCammon, 2004).

Third, Haegeli and McCammon have refused to allow access to their methods and data to others for inspection (Uttl et al., 2008a; Floyer, 2008).

Nonetheless, the CAC continues to publicly claim that the “[Avaluator] is the best decision guidance tool available for amateur backcountry users” (M. Clayton, in Risky Business, Red Deer Advocate, Jan. 10, 2009).
there a problem with the Avaluator?, Calgary Herald, April 19, 2009). Most recently, the CAC added two new fine print liability disclaimers to the third printing of the Avaluator (2009, January) stating that the Avaluator Avalanche Accident Prevention Card is suitable for no particular purpose and that neither the developers nor the publisher are responsible for any damages, including injuries and deaths arising from its use.

3. 7CF: A NEW AVALANCHE RISK REDUCTION TOOL

Apart from being neither scientific nor suitable “for any particular purpose, use, or application”, the Avaluator's Obvious Clues method has several other undesirable features. First, the reliability with which the obvious clues can be assessed in typical field conditions is far from perfect or even acceptable (Haegeli & Haider, 2008; Uttl et al., 2008a,b; Uttl et al., 2009d). As suggested by the “seven obvious clues” moniker, we expect some of the clues, e.g. the presence/absence of an avalanche path, to be easy to assess in a reliable fashion. In contrast, the determination of the presence vs. absence of other clues is far more problematic. For example, the determination of whether the last 48 hours has resulted in the accumulation of 20 cm of new snow can be challenging. How many back-country recreationists carry rulers in their equipment and sample the depth of new snow? Moreover, how does one determine how much snow fell within the last 48 hours as opposed to the last 49, 72, or 96 hours? How does a snowmobiler assess snow instability such as cracking or whumphing while roaring across the snow at 70 km/hr with a helmet on?

Second, from its conception, the Avaluator was to be a “simple” decision tool. Albi Sole, the coordinator of avalanche safety programs at University of Calgary, stated: “I say keep it simple. Seven clues is plenty.” (A. Sole, in Is there a problem with the Avaluator, Calgary Herald, April 19, 2009). However, the Avaluator may still be too challenging in part because of the difficulty with reliably detecting the presence/absence of the clues.

In response to these challenges that severely limit the Avaluator’s usefulness as well as confidence in its behavioral recommendations, we elected to develop a new relative risk reduction tool, the 7CF, with the following criteria in mind: (1) the tool should achieve accident prevention values at least comparable to the Avaluator; (2) the reliability of clue assessment must be very high, preferably close to 1.00; (3) the tool should be simple to use even for Canadian backcountry users, even simpler than the Avaluator. The new tool also needed to be fully transparent, open, and prevention values replicable by any interested person, public, or scientist.

After several false starts, we developed a new tool, called the 7CF, that satisfies all of the above requirements. As with the Avaluator, our avalanche risk reduction tool requires that the user observe the presence/absence of seven obvious clues. However, in contrast to the Avaluator, the clues are very simple and completely unambiguous to all but the most inept users.

Figure 2 shows the clue distributions and true prevention values for the Avaluator (based on McCammon, 2004; Uttl et al., 2008b; Floyer, 2008) as well as the distribution and prevention values for the 7CF tool. The 7CF has achieved or even exceeded our design objectives and expectations:

- Figure 2 highlights that the 7CF prevention values are nearly identical to those of the Avaluator.
- The reliability of assessing the presence or absence of the 7 clues is extremely high. Our preliminary results indicate that the correlation between the number of clues detected and the actual number of clues present is over 0.999.
- The 7CF is extremely easy to use, in part because of the ease with which individual clues can be reliably assessed. Our preliminary research shows that the users require at most 2 minutes of training to be fully proficient with the 7CF.
- The 7CF’s prevention values are fully transparent and easy to replicate by anyone interested in verifying our claims. The full replication can be completed in a matter of hours or even minutes by any skillful researcher.
- Surprisingly, the 7CF is also very inexpensive; it can be assembled for mere pennies. Although it was not part of the 7CF’s original design specification, the 7CF is also eco-friendly and time saving. It makes no difference whether you assess the stability of the slope just before you cross it or in the comfort of your home before you even embark on the drive to your destination. Thus, you will reduce your carbon footprint by eliminating the drive to slopes that have been deemed unsafe to cross.

The reason that the 7CF can be produced for pennies is that the clues are, in fact, coins. The 7CF (or 7 Coin Flips) requires the user to flip seven coins. The number of resulting “Heads” is the number of clues present for the specific slope...
(the number of “Tails” is the number of clues not present). The 7CF’s distribution is the binomial distribution with 7 trials and a 50% chance of success for each trial.

4. LESSONS FROM THE 7CF

The 7CF is a tongue-in-cheek proposal. What are not tongue-in-cheek are the implications of its prevention values for the evaluation of avalanche risk reduction tools. The equivalence of the 7CF’s prevention values to those of the Avaluator demonstrates that there are limitations to the usefulness of evaluating risk reduction tools based on prevention values only. Prevention values derived from accident data alone are simply not sufficient to conclude that a tool would be effective in reducing avalanche accidents.

Why? First, the clues used to make decisions whether slopes are safe to cross and their combinations should be diagnostic of avalanches occurring (preferably the most diagnostic). Second, the recreational user pondering whether a slope is safe to cross wants to know the positive predictive value, or probability, of the slope avalanching given that the tool indicated the slope is not safe. Third, a risk reduction tool should be more effective than the users’ native strategies.

4.1 Diagnosticity of Clues/Tools

In general, an event is diagnostic of an outcome to the extent to which the probability of the event is higher when the outcome occurs vs. not occurs. Formally, Bayesian diagnosticity of an event (E) for the occurrence of an outcome (O) is:

\[ \text{Diagnosticity of } E = \frac{p(E|O)}{p(E|-O)} \]

where \( p \) indicates probability and -O indicates “not outcome” or outcome absence. To obtain the diagnosticity of a clue or clue combination, one merely needs to replace an event (E) with a clue or clues (C) and an outcome (O) with an avalanche (A) in the above formula.

The clues, individually and in combination, should be diagnostic of an avalanche occurring. Clearly, the 7CF fails dramatically when evaluated on the diagnosticity of its clues since the occurrence of avalanches is uncorrelated with outcomes of coin flipping.

So how can the 7CF have positive prevention values when its clues lack diagnosticity? The reason is that prevention values are not influenced by how prevention is achieved. Anything that would keep people (in this case former avalanche victims) off a slope will result in positive prevention values, even a rule as trivial as turning around if someone in a group shows up for the trip in a red jacket. It doesn’t matter whether the criterion for staying off the slope is in any way associated with the likelihood of an avalanche occurring. Having a person flip a coin (or seven) as a way of determining when not to ski a particular slope is no different from having them look for 20 cm of new snow within 48 hours, at least as far as prevention values are concerned.
One challenge in marketing the 7CF would be that potential users would object to its random nature. It would be patently obvious that the prevention of accidents was occurring simply by rendering back-country recreational opportunities unavailable without any rational basis.

But do we have any good evidence that more accepted avalanche risk assessment tools are any better? We don’t. We do not have non-accident data, p(C|-A), and therefore, we do not know whether the Avaluator’s (or any other tool’s) clues are diagnostic or not, or to what extent. Figure 3, panels A to C, illustrates the problem using a hypothetical risk reduction tool that has a much finer gradation of danger points (similar to the Nivo scale) than the Avaluator’s severely limited 7-point scale (see Uttl et al., 2009d). Panel A shows the distribution of the danger points in accident data only together with a hypothetical caution/not recommended advice. If the user accumulates more than 24 points, the advice is “not recommended” whereas 24 or fewer points means proceed with “caution” or “extra caution”. The prevention value of avoiding slopes with more than 24 point is about 84% (i.e., the percent of area under the curve above 24 points). Panel B shows the distribution of danger points in both accident data and trips with no accidents, assuming that for every trip ending in an accident there is one trip without an accident (Accident Ratio AR = 1:1) and that the distribution of clues does not differ between accident trips and accident-free trips. It is clear that the tool maintains its 84% prevention value but at the cost of eliminating 84% of accident-free trips. Its diagnosticity is 1, that is, the tool is not diagnostic, and thus, completely useless. Panel C shows a far more desirable situation. The accident ratio is still 1:1 but the accident vs. non-accident distributions are widely separated. The prevention value of the tool remains at 84% but at only a minimal cost of eliminating a few accident free trips. The diagnosticity is high.

Unfortunately, having no data on the distribution of the Obvious Clues in accident free trips, we have no way of knowing whether the Avaluator’s Obvious Clues method has any diagnosticity at all.

4.2 Positive Predictive Value

What recreational users want to know is not the prevention value but the so-called positive predictive value (PV+). That is, given that the decision tool indicated the slope is not safe to cross (i.e., an avalanche is likely to happen), what is the probability of the slope actually avalanching? This probability depends not only on the separation between the clue distributions for accident trips vs. accident-free trips (diagnosticity) but also on the ratio of accident trips to accident-free trips.

Figure 3 illustrates this point. Panel B shows no separation between the clue distributions for accident trips vs. accident-free trips. Half of the “not recommended” crossings will end up in an accident, and thus, PV+ is 50%. Panel C shows substantial separation between the two distributions, with nearly all “not recommended” crossings ending up in an accident (PV+ = 98%). Panel D shows a smaller separation between the two distributions and “only” 62% of the “not recommended” trips will result in an accident. Panel E shows the same smaller separation between the two distributions but the accident ratio (the number of trips resulting in an accident over the number of accident free trips) is 1:10. The PV+ dropped to 14%. Panel F shows the same data for a more realistic AR 1:1000; the PV+ is only 0.2%. Out of 1,000 “not recommended” crossings, only 2 will result in an accident.

In all of the above scenarios with widely different outcomes, the prevention value remained at an impressive 84%. Thus, the lack of relevant data on the presence vs. absence of the Obvious Clues in accident-free trips as well as the lack of data on the ratio between accident and accident free trips suggest that recommending the use of the Avaluator for accident prevention is akin to selling snake oil for improving one’s reaction time. It may be effective, yet again, there is no evidence that it is effective. It may also be dangerous if snake oil users start driving faster because they believe they can react faster. Moreover, if the tool indicates to the user that slopes are not safe to cross but he or she repeatedly observes that nothing bad happens to people crossing the slopes, the user is likely to toss the tool and consider it useless.

4.3 Native strategies

When evaluating an avalanche risk reduction tool via prevention values, the analysis is focused on a particular group – historical victims only. The victims were faced with an immediate choice about whether or not to cross potentially dangerous slopes, they chose to go, and the consequences were catastrophic. The purpose of avalanche risk reduction tools is to prevent accidents by making the potential victims change their decision from “Go” to “Don’t Go”.

2010 International Snow Science Workshop
Figure 3. Hypothetical distributions of danger points (e.g., sum of weighted clues to avalanche danger) in accident records (red) vs. accident free trips, for different accident ratios and effect sizes (i.e., separation between accident vs. accident free distributions of danger points). In all cases, prevention value is 84%.

The effectiveness of a decision-making tool can be easily evaluated for historical victims. If the victims had limited their travel to 4 or fewer of the Avaluator’s Obvious Clues, the number of accidents would be reduced by about 20% (Uttl et al., 2008b). Thus, out of 1,000 historical trips, only 800 would have ended in an accident and 200 would have been accident-free.

However, the purpose of avalanche accident reduction tools is to minimize the total number of accidents rather than the number of accidents among the historical victims. Many more people were faced with the same choice to cross a potentially unsafe slope but chose not to. These people were not represented in the data sets used to design the risk reduction tools.

Let’s consider a hypothetical world in which people have a native strategy that keeps them off 50% of avalanche slopes when faced with such decisions. Table 1 shows the frequency of “Go” (“caution” or “extra caution”) vs. “No Go” (“Not recommended”) decisions when 1000 parties use native strategies vs. a 20% prevention strategy such as the Avaluator with 4 or fewer clues. We
can see that by replacing people’s native strategy with a less restrictive strategy, the overall number of people putting themselves onto slopes ready to avalanche has gone up, not down. This will occur regardless of the relative diagnosticity of the strategies involved. Not surprisingly, with a greater number of people on the slopes, our hypothetical world would also see more avalanche accidents, injuries, and deaths.

Moreover, even the simple possession of an “avalanche accident prevention card” is likely to convince more people that back-country risks are manageable and lure more people into avalanche terrain, and again, result in a greater number of avalanche accidents.

Table 1. Decisions under native vs. 20% prevention strategy.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>“Go”</th>
<th>“No Go”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>20% Prevention</td>
<td>800</td>
<td>200</td>
</tr>
</tbody>
</table>

One may argue that risk reduction tools such as the Avaluator are not meant to replace people’s native strategies, merely to supplement them. However, one must recognize that there will be times when the recommendations of the “official tool” will conflict with untutored native wit. In such cases, an appeal to the tool developed by “experts” and endorsed by government agencies is often likely to win out. In studies of conformity, individuals are more likely to conform to the group’s standard if the group members are perceived as having greater expertise (e.g., Crano, 1970). Similarly, people are more likely to adopt recommendations of perceived experts vs. non-experts (e.g., Chaiken & Maheswaran, 1994).

In summary, the net impact of having people make decisions based on an avalanche risk reduction tool cannot be deduced from its effects on the decisions of former victims alone (DiGiacomo, 2006). One must also consider the effect on the decisions made by people who did not become victims, but who could have. If the influence of the tool is such that more of the second group are killed than are saved from the first group, then net fatalities will increase. In fact, the introduction of the Avaluator to the Canadian market and as a primary decision making strategy taught in AST 1 courses approved by the Canadian Avalanche Association was followed by a doubling of the number of avalanche accidents in Canada rather than in the promised reduction (Uttl et al., 2008b; Uttl et al., 2009a). It is possible that the rise in Canadian avalanche accidents is coincidental. However, the increase in the number of accidents is expected from the grossly inflated prevention values published in the Avaluator. The inflated prevention values likely give users a false sense of security, encourage them to cross unsafe slopes, and the subsequent increase in the number of accidents is expected.

5. CONCLUSIONS

In this paper we put forward a tongue-in-cheek proposal for an avalanche risk reduction tool, the 7CF or 7 Coin Flips. The 7CF is obviously grossly flawed, but examining it reveals some interesting truths about how such tools are evaluated. The evaluation of avalanche risk assessment tools must go beyond its current exclusive reliance on prevention values. The prevention values are valuable but they are not sufficient for evaluating the effectiveness of risk reduction tools in preventing accidents and their general usefulness. Developers wishing to validate risk assessment tools should place far more effort into establishing the diagnosticity of the individual clues, their combinations, and accident prevalence. They also need to consider the consequences of replacing people’s native strategies for safe travel in avalanche terrain.

At this time, evidence for the usefulness of the Avaluator for preventing avalanche accidents is lacking. If anything, the Avaluator has been facilitating rather than preventing avalanche accidents (Uttl et al., 2008a,b).

Not surprisingly, the Avaluator’s developers (Drs. Haegeli & McCammon) have recently arrived to the conclusion that the Avaluator Avalanche Accident Prevention Card’s Obvious Clues Method is not suitable “for any particular purpose” and both the developers and the CAC now disclaim liability for any damages including “injury or death” arising from the Avaluator’s use (Haegeli & McCammon, 2006; 3rd printing January 2009).

We concluded that the Avaluator was not suitable for avalanche accident prevention two years ago and we are pleased that the developers and the publisher now concur with our original recommendation regarding the suitability of Avaluator for avalanche accident prevention: Do not use it for that particular purpose.

What we find surprising and unethical, however, is that key messages about the Avaluator’s lack of suitability for avalanche accident prevention are provided only in fine print. We believe the Avaluator should be recalled or at
the very least everyone who purchased the Avaluator should be sent a sticker to be placed over the Avaluator Avalanche Accident Prevention Card stating, in big bold print, “this tool is suitable for no particular purpose.”

In fact, we are not alone. Over 100 participants in our study on people’s perception of the Avaluator's developers’ and publisher's actions agree with us: they believe the developers (Drs. Haegeli and McCammon), Canadian Avalanche Association, and Canadian Avalanche Centre should tell the truth and recall the Avaluator (Uttl et al., 2010).

6. ACKNOWLEDGMENTS

We thank Amy L. Siegenthaler for careful reading and comments on the manuscript.

7. REFERENCES


