Further comments on the correction denominator *Rp* and double-counting
By Julie Yee, March 19, 2008

After having a discussion with Bill Warren-Hicks, the biometrician working for FPLE, I decided to further consider the *Rp* factor in the denominator of the adjusted mortality estimate (Equation 1 of the monitoring team's report, document M21). This is mainly to address his concerns documented in item 2b of the FPLE memo (P77). Although that particular memo does not discuss joint and marginal probabilities of scavenging and observer detection, it did come up in my discussion with him in relation to that topic. So, I have laid out my understanding of the correction factors in terms of joint and marginal probabilities. As a result, I still don't believe that the application of these correction factors is contributing to upward bias in the mortality estimates, as long as the correction factors are applied appropriately (more about this in my comments in P80). However, I have started to more deeply examine the potential for large biases due to double-counting, which has raised questions and concerns for me.

Joint and Marginal Probabilities of Scavenging and Observer Detection. Here is what we currently know or have defined. We know scavengers are roaming around that may or may not pick up specific carcass. We also know that on a specified date, observers will search the plot and may or may not detect specific carcasses. We have defined the probability that a specific carcass disappears prior to the next survey date is $(1-R)$ and the probability that the carcass continues to remain just prior to the survey is R. Obviously, if the carcass has disappeared prior to the survey date, then the observers will absolutely not detect it on the search plot. If the carcass remains on the plot at the time of the survey, then we have defined the observers will detect it with a probability of p and fail to detect it with a probability of $(1-p)$. Note that p is a conditional probability of detection, with the condition that the specific carcass is still on the plot.

Suppose we have a single carcass that has fallen on a search plot on a particular day sometime between surveys.

Define:
Z = # of carcasses from that particular day remaining at the end of search period. Since we are starting with 1 carcass, then Z is either 0 (if carcass was removed by scavengers) or 1 (if not removed by scavengers).
Y = # of carcasses from that particular day that is detected at the next survey period. Again, Y can only be 0 or 1.

Here is the joint probability distribution for (Z, Y):

Z	Y	
	1	0
1	Rp	$R(1-p)$
0	0	$(1-R)$

In other words, there are three possible outcomes of (Z, Y):

1. (Z, Y) = (1,1), where the carcass has remained and is detected. This occurs with probability $R \times p$.
2. (Z, Y) = (1,0), where the carcass has remained and is not detected. This occurs with probability $R \times (1-p)$.
3. (Z, Y) = (0,0), where the carcass disappears and thus would not be detected. This occurs with probability $(1-R)$.

Note that the (0,1) fate has zero probability because it is impossible that a carcass would disappear and be detected where it no longer is.
Also note that outcome 2 is a failure to detect due to observer bias. Conversely, outcome 3 is a failure to detect due to scavenger removal. Outcomes 2 and 3 cannot both occur for the same carcass in this interval.

The marginal probability that Y is 1, or in other words the probability that the carcass is observed, is Pr{Y=1} = Pr{(1,1)} + Pr{(0,1)} = Rp + 0 = Rp.
The marginal probability that Y is 0, or in other words the probability that the carcass is missed either due to observer bias or scavenger removal is Pr{Y=0} = Pr{(1,0)} + Pr{(0,0)} = $R(1-p) + (1-R) = 1-Rp$.

The marginal distribution of Y is Bernoulli with probability Rp. When we are no longer considering the fate of just one carcass, but rather the fate of an arbitrary number of carcasses, X, then we might reasonably assume independence of fates and extend the distribution of Y to a Binomial distribution with X trials and probability Rp. The likelihood function is

$$L = \frac{X!}{Y!(X-Y)!}(Rp)^Y(1-Rp)^{X-Y}$$

On average, we expect Y to roughly equal XRp, or X to roughly equal $Y/(Rp)$, i.e. actual mortalities to roughly equal unadjusted mortalities divided by Rp.

Example: Suppose 100 fatalities occur on a given day. If R=0.2, then we expect roughly 20 out of 100 of those fatalities will remain by the time of the survey and the other 80 will have disappeared. If additionally p=0.2, then we expect roughly 4 of those approximately 20 fatalities that remained will be detected and counted while the other 16 will have been missed. On average, if R=0.2 and p=0.2, then 80% of carcasses will have been missed due to scavenger removal and another 16% will have been missed due to observer bias, for a total of 80%+16%=96% of carcasses being missed due to one or the other failures. By taking the roughly 4 observed carcasses and dividing by (0.2)(0.2)=0.04, then we obtain the estimate that roughly 100 fatalities actually occurred.

In general, for a given R or p, one can calculate the expected proportion of carcasses missed due to scavenger removal and the expected proportion of carcasses missed due to observer bias. The proportion of carcasses missed due to scavenger removal is, on average, the probability for outcome (0,0) in the above table, or $(1-R)$ (which is 1-0.2=80% in the previous example). The proportion of carcasses missed due to observer bias is, on average, the probability for outcome (1,0) in the above table, or $R(1-p)$ (which is 0.2(1-0.2)=16% in the previous example).

In the monitoring team's analysis, the correction factor for scavenger removal, R, is modified to correct for cumulative numbers of carcasses over the span of the search interval and is known as R_c instead of R. The correction is applied the same way, by dividing by $R_c p$ instead of Rp (I've satisfied this for myself using reasoning described in document P80).

Double Counting. There is a potential concern with double counting which is worth discussing. As an example, consider the scenario in which a carcass falls onto a survey plot in one search period, is not scavenged during that period, is missed by observers on the first subsequent survey, is not scavenged during the second period, and is then detected by observers on the second subsequent survey. In this scenario, the carcass has been double-counted so-to-speak. It was "counted" in the first survey by being accounted for in the extrapolation from unadjusted to adjusted numbers, and then it was counted again in the second survey as part of the unadjusted count. Applying an adjustment on that second count extrapolates and inflates the adjusted estimate even further. The probability of persisting beyond the search interval when the carcass fell on the search plot is the probability that $(Z,Y)=(1,0)$, or $R(1-p)$. I would be particularly concerned with double-counting if R was large and p was small to moderate, because high R and small to moderate p increases the chance that a carcass persists on a plot and is detected on a second or later survey. I would not be concerned if p was very large, such as p=100% for large raptors, because this would mean that the carcasses are unlikely to persist beyond their search interval to the point of causing large biases.

I'm less certain about the small raptors. In the monitoring team's analysis, the correction factors p and R vary between small and large raptors. For small raptors, the probability of persisting on a search plot is estimated as p =74% and R_c is estimated between 22% and 26% depending on the search interval. Then we expect the probability of a carcass persisting beyond the interval when it actually occurred is 0.22(1-0.74) to 0.26(1-0.74) (i.e. 5.7-6.8%). So, roughly 5.7-6.8% of small raptor carcasses are unscavenged but fail to be detected due to observer bias, and they are automatically accounted for as part of the adjusted count M_a. And if these roughly 5.7-6.8% of small raptor carcasses remain on the search plot for another search rotation and are detected on the next survey, then they are counted in again as part of unadjusted count M_u. Since p=0.74 and R_c=0.24 (roughly), then each double-counting event inflates the overall count not just by 1, but by $1/(Rp)$ = $1/((0.74)(0.24)) = 5.6$.

So, consider this scenario:
1000 small raptor carcasses fall on the Altamont in a year. 500 of these fall on the roughly ½ portion of the Altamont that is searched. Assuming p=74% and R_c=24%, then 0.24(1-0.74)=6.24% of those 500 carcasses, or (0.0624)(500)=31 carcasses, fail to be detected at the survey corresponding to the interval in which they occurred. As long as those carcasses remain undetected, then they are appropriately counted as part of the adjusted mortality. However if any of those carcasses should be later detected and included into the unadjusted count, then they will have been double-counted. The monitoring team evaluates carcasses and sometimes makes a decision that a carcass is old

enough that it probably occurred in the previous search interval. I'm unsure whether the monitoring team includes those carcasses into the unadjusted count for the previous search interval. If they do, then I think that would contribute to double-counting.

In the worst case scenario of double-counting, those 31 carcasses are eventually found and included into the unadjusted count. This would assume that carcasses are not eventually scavenged once they have persisted beyond their first interval on the plot (maybe they have gotten too old?) and that the search team returns often enough to eventually detect them. If the 31 carcasses are included as part of the unadjusted count (whether or not these carcasses were backdated into the correct interval), then the adjusted count increases by $31/((0.74)(0.24))=175$. After extrapolating to the entire Altamont, then the adjusted count increases by 350. The adjusted count could end up being 1350 when truth is actually 1000.

Questions to consider:

1. Is this type of double-counting occurring?
2. When the MT finds an old carcass that appears to have been missed from a prior survey, are they put into the unadjusted estimate?
3. Can the potential for bias be eliminated by excluding old carcasses from the unadjusted count?
4. Can the MT reliably determine whether a carcass is old enough to have occurred in a previous search interval?
5. Is this level of double-counting occurring?
6. Will carcasses, after some time, fail to ever disappear?
7. Is there a way to include all carcasses into the unadjusted count, and then apply an additional correction to eliminate the effect of double-counting?
8. How well do we know the probability of scavenger removal for carcasses that are in the second or later interval?
9. How well do we really know scavenger removal?
10. How well do we really know observer bias?