28 September 2010
Steven L. Cain
Grand Teton National Park
P.O. Drawer 170
Moose, WY 83012
307-739-3485, Fax 307-739-3490
Steve_cain@nps.gov

RH: Cain et al. · Bison Pregnancy Prediction

Using Fecal Progestagens and Logistic Regression to Predict Individual Bison Pregnancy Status

STEVEN L. CAIN,[1] *National Park Service, Grand Teton National Park, P. O. Box 170, Moose, WY 83012, USA*

MEGAN D. HIGGS,[2] *Montana State University, Department of Mathematical Sciences, 2-242 Wilson Hall, Montana State University, Bozeman, MT 59717-2400, USA*

THOMAS J. ROFFE,[3] *U.S. Fish and Wildlife Service, 1400 S. 19th Ave, Bozeman, MT 59718, USA*

STEVEN L. MONFORT,[4] *Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA*

JOEL BERGER,[5] *University of Montana, Division of Biological Sciences / Wildlife Conservation Society, Northern Rockies Field Office, Missoula, MT 59812, USA*

[1] Email: steve_cain@nps.gov
[2]Email: higgs@math.montana.edu
[3]Email: troffe@montana.edu
[4] Email: monforts@si.edu
[5] Email: jberger@wcs.org

ABSTRACT Ungulate ecological studies often include components of reproduction because of its demographic importance and the ecological factors affecting it. Pregnancy status, in particular, is key because it represents a starting point for succeeding measurements of vital rates. As part of a broader study on free-ranging bison (*Bison bison*) reproductive ecology in the southern Greater Yellowstone Ecosystem, we sought to evaluate fecundity and the prevalence of fetal and neonatal loss in a marked sample of reproductive aged animals. Our goal was to develop a non-invasive method for assessing pregnancy by establishing the empirical probability of pregnancy in each female, through evaluation of fecal progestagens in a single bison scat collected in winter. Using a binary indicator of pregnancy status as the response variable, we developed a logistic regression model that predicted pregnancy status through estimation of the probability of pregnancy, thereby providing the ability to gauge our confidence in the prediction using the quantified uncertainty in the probability estimates. For 155 observations of 42 marked bison, we used combinations of transrectal uterine palpation and calf status as independent measures of pregnancy, recognizing greater potential for error using either by itself. We evaluated models by comparing mis-prediction rates derived with leave-one-out cross validation, and by comparing percentage of 95% confidence intervals that crossed a pregnant-not pregnant threshold. Correct predictions with high confidence were made with a model that used year-centered, log transformed progestagen concentration from a single mid-late gestation scat sample, resulting in an overall successful pregnancy prediction rate of 93.5%. By providing a measure of confidence for each prediction, our approach allows practitioners to further evaluate estimates for individual animals, thereby improving prediction interpretations. Our work underscores the importance of independently validating fecal hormones with other objective measures of pregnancy. Detailed knowledge of endocrine excretion dynamics and more

comprehensive longitudinal assessments of fecal metabolites in pregnant and non pregnant bison would help identify ideal sampling periods and increase the ability to estimate pregnancy using fecal progestagens. Our approach may be appealing to practitioners because fecal samples necessary for evaluating progestagens are easily collected and preserved, laboratory procedures are well documented, and logistic regression is readily available in statistical computer software. Furthermore, samples can be obtained non-invasively, which reduces cost and potential bias, and increases animal safety, human safety, and social acceptability.

KEY WORDS bison, fecal, hormones, logistic, non-invasive, pregnancy, pspb, regression, Wyoming, Yellowstone.

Ungulate ecological studies often include components of reproduction because of its importance to demography (e.g., Testa 2004, Coulson et al. 2005, Fuller et al. 2007) and the ecological factors affecting it (Berger et al. 1999). Among parameters commonly used to assess reproductive performance, pregnancy status is key because, following ovulation, it represents a starting point for succeeding measurements of vital rates, including birth rates, neonatal survival, and recruitment of breeding aged individuals into a population (Ramsay and Sadleir 1979, Messier et al. 1990, Kirkpatrick et al. 1993, Russell et al. 1998, Stoops et al. 1999). Knowledge of pregnancy can be useful at the population level, where it may be expressed as an overall rate (Messier et al. 1990), or at the individual level, where the status of a marked individual and its subsequent reproductive outcome is of interest (Messier et al. 1990, White et al. 1995, Garrott et al. 1998).

Determining pregnancy of wild ungulates is inherently difficult and presents a serious challenge to researchers interested in reproductive ecology. Handling animals, either through capture and physical restraint or chemical immobilization, is required for direct evaluations of pregnancy status, including transrectal palpation (Hein et al. 1991) and ultra-sound (Barrett 1981, Stephenson et al. 1995), as well as for indirect assessments using serum pregnancy specific protein B (PSBP) (Houston et al. 1986, Sasser et al. 1986, Wood et al. 1986, Haigh et al. 1988). An obvious disadvantage with these methods is that there are hazards associated with handling animals, both to the research subjects (e.g., Ballard and Tobey 1981, Valkenburg et al. 1983, Larsen and Gauthier 1989, Peterson et al. 2003) and-or to the handlers (Jessup et al. 1988, Haigh 1990, Dein et al. 2005, Roffe et al. 2005 and references therein), some of which could bias the results of the intended research (Welsh and Johnson 1981). The potential for capture related mortality is of biological concern as well, particularly when handling rare or endangered species (Berger et al. 2010, Association for the Study of Animal Behaviour 2003). Likewise, sociological concerns need to be considered given that contemporary wildlife science audiences are highly informed, influential, and do not universally accept the need to manipulate animals or use invasive, potentially harmful techniques considered conventional by the scientific community (Farnsworth and Rosovsky 1993, see Bekoff and Jamieson 1996, Peterson et al. 2003).

In general, non-invasive methods for determining ungulate pregnancy status (Lasley and Kirkpatrick 1991, Monfort 2003) are more limited in inferential capability but have several inherent advantages, including: 1) safety, because capture and handling are not required; 2) reduced bias from potential effects on behavior, activity, distribution, or physiology; 3) lower

expense because immobilizing drugs or specially trained personnel to administer them or to perform uterine transrectal palpation are not necessary; and 4) social acceptability. To date in ungulates, these techniques have focused on the measurement of reproductive steroid hormone metabolites excreted in urine (Poole et al. 1984, Kirkpatrick et al. 1990, Monfort et al. 1990,) and-or feces (Kirkpatrick et al. 1990, Monfort et al. 1993, Garrott et al. 1998, Berger et al. 1999, Schoenecker et al. 2004). Application of these techniques to accurately estimate pregnancy status for individual animals from single fecal samples has been proposed only for elk (*Cervus elaphus*, Garrott et al. 1998) and bighorn sheep (*Ovis canadensis*, Schoenecker et al. 2004).

As part of a broader study on bison (*Bison bison*) reproductive ecology, we sought to evaluate fecundity and the prevalence of fetal and neonatal loss in a marked sample of reproductive-aged animals. We directly assessed pregnancy status in a subset of our population handled each year to determine the incidence of brucellosis and its effect on reproduction. For the remainder of our study animals, our goal was to develop a non-invasive method for assessing pregnancy that established the empirical probability of pregnancy in each female by evaluating fecal progestagens in a single bison scat collected in winter. Our field protocol required determining the presence of a calf with marked, reproductive-aged cows at least weekly during the parturition period to estimate date of birth. Since our study population was wide-ranging, this was a labor intensive task. Our objective was to remove animals that we were confident were not pregnant from the group of those checked for calves each week, thus improving efficiency.

Since our study area did not have continuous snow cover in winter, urine collection was not feasible. Therefore, we focused on obtaining fecal samples from marked females and

quantifying fecal progestagens for estimating the probability of pregnancy. Based on past research (Kirkpatrick et al. 1992, 1993), we expected concentrations of fecal progesterone metabolites to differ between pregnant and non-pregnant bison. However, in contrast to similar studies on ungulates, our goal was not to simply compare mean level of fecal hormone between pregnant and non-pregnant groups, but instead to estimate the probability of pregnancy for individual cows based on hormone level, and use this probability to predict pregnancy status. Using a binary indicator of pregnancy status as the response variable, we focused on development of a logistic regression model that would allow prediction of pregnancy status through estimation of the probability of pregnancy. We could thus gauge our confidence in the prediction using the quantified uncertainty in the probability estimates.

STUDY AREA

Our study area in the southern Greater Yellowstone Ecosystem (GYE) focused on Grand Teton National Park (GTNP) and the National Elk Refuge (NER), which comprised the primary range of the free ranging Jackson bison herd. This area included the upper Snake River drainage in a high elevation valley commonly referred to as Jackson Hole, bound by the Teton Range to the west, the Gros Ventre and Absaroka Mountains to the east, the Yellowstone Plateau to the north, and the town of Jackson, Wyoming to the south. Elevations ranged from 1,890 m in the valley floor to 4,197 m atop surrounding peaks. The climate is characterized by long, cold, snowy winters and short, cool summers. The 49-year (1958-2006) average high and low temperatures in January were -3.4° C and -17.2° C and in July were 26.8° C and 5.2° C at Moose, Wyoming (Western Regional Climate Center 2009), located roughly in the center of the study area. Approximately 70% of precipitation fell as snow. Bisected by the Snake River riparian floodplain, the valley floor was dominated by numerous river terraces with minor elevational relief, glacial moraines,

and several timbered buttes rising 150 to 300 m above the valley floor, and contained large areas of previously or currently irrigated and grazed hay lands (Smith et al. 1998). Vegetation was described by Dunk et al. (1997).

Bison were extirpated from this area in the late 1800s but were reintroduced from Yellowstone and Theodore Roosevelt National Parks in 1948 and 1964, respectively (Grand Teton National Park and National Elk Refuge 1996). The herd grew from 37 individuals in 1980, when they discovered winter elk feeding operations on the NER, to over 1,000 in 2006 (S. L. Cain, Grand Teton National Park, unpublished data). Since 1980 the herd established a regular pattern of summering in GTNP and wintering on the lower elevation NER (Grand Teton National Park and National Elk Refuge 1996). During our 1997-2005 study, herd size ranged from 320-950 (S.L. Cain, GTNP, unpublished data) and pelleted alfalfa was provided to elk and bison on the NER for an average of 57 days (range 28-85) each winter (E. Cole, National Elk Refuge, unpublished data).

Brucellosis was discovered in the Jackson bison herd in 1989 (Williams et al. 1993). Its source has been attributed to the original Yellowstone reintroduction, since these bison were known to be infected (Meagher 1973), or sympatric elk from the Jackson elk herd, also known to be infected (Thorne 1982). Based on a criterion that required a minimum of 2 out of 6 serologic tests to be positive by bovine Uniform Methods and Rules (USDA 2003), at least one of which had to be a quantitative test, brucellosis annual seroprevalence in non-calf predominantly female bison averaged 69% during our study (T.J. Roffe, USFWS, unpublished data). We considered

bison sampled more than once during the year as positive if they met the criteria for positive on any sample.

METHODS

Our study protocol required marking, taking samples from, and monitoring the reproductive performance of a sample of female bison each year. We immobilized bison mid-winter (late Jan-mid Mar) from a tracked snow vehicle when they were on or near NER feedlines, using a carfentanil/xylazine drug combination administered from a projectile dart using the methods of Roffe et al. (2001). We deployed Telonics very high frequency radio collars (Telonics, Inc., Mesa, AZ) and numbered 5.8 x 7.3 cm bangle ear tags on each individual, initially selected randomly from groups of several hundred bison on feedlines. We did this by generating random numbers from 1-20 to determine the *nth* adult cow we would choose as we approached or worked our way through a group of bison. After marking our original sample in 1997, during 1998-2002 we attempted to re-capture each individual annually to determine pregnancy status, take blood and fecal samples, and replace aged radio-collars. During 2003-2005 we re-captured only half of our marked sample annually to address a separate objective of evaluating the effects of handling on reproduction. We added new animals to our sample each year, both randomly and targeted by approximate age, to replace animals that died or whose radio collar expired or malfunctioned.

For each anesthetized bison, transrectal palpation was performed by a veterinarian to assess pregnancy status, 20cc of whole blood was taken by jugular venipuncture to evaluate brucellosis exposure and pregnancy-specific protein B (PSPB), and approximately 20-50g of feces was obtained unless the animal's colon was void. For animals not immobilized, we obtained fecal

174 samples by locating individuals with radio-telemetry, verifying their identity with ear tags when
175 multiple radioed animals were in a group, watching them until they defecated, then collecting the
176 target sample. We minimized potential for sampling non-target fecal matter by approaching
177 bison closely, working in pairs whenever possible (allowing 1 person to remain focused on the
178 target sample while the other approached and collected), using a stationary spotting scope
179 focused on the target sample to corroborate location, sampling only fresh fecal matter, and
180 rejecting samples we could not confidently attribute to the target animal. We stored fecal
181 samples frozen in sealed plastic bags until thawed for laboratory analyses.
182
183 Endocrine assessments were conducted at the Smithsonian Conservation Biology Institute (Front
184 Royal, Virginia, USA). All fecal samples were processed and extracted using fecal hormone
185 procedures that have been previously described (Garrott et al. 1998). Fecal progestagens were
186 quantified with a monclonal antibody (CL425, C. Munro, U.C. Davis, Davis, CA) used in a
187 radioimmunoassay (1997-2000; Wasser et al. 1994) or enzyme-immunoassay (2001-2005;
188 Graham et al. 2001). For both assays, serially diluted fecal extracts demonstrated displacement
189 curves that were parallel to those of standard hormone preparations. All samples for each year
190 were evaluated in a single assay, and intra-assay variation was < 5%.
191
192 We centrifuged blood, harvested serum and froze all samples until processing. Sera were
193 submitted to the American Association of Veterinary Diagnostic Laboratories certified Montana
194 Veterinary Diagnostic Laboratory, Bozeman, Montana, for 8 *Brucella* serologic tests: card (Rose
195 Bengal), buffered antigen plate agglutination (BAPA), standard plate agglutination (STP),
196 standard tube agglutination (STT), rivanol (RIV), and complement fixation (CF). Samples from

2002 to 2005 were also processed using fluorescent polarization (FP). We interpreted tests as positive or negative based on US Department of Agriculture bovine brucellosis uniform methods and rules (USDA 2003). We classified high titered animals based on complement fixation of a 4+ reaction at 1:80 dilution or greater (Roffe et al. 1999).

We also submitted sera to Biotracking, Inc (Moscow, ID, biotracking.com) to determine concentrations of PSPB (Sasser et al. 1986). PSPB was determined by radioimmunoassay (earlier samples) and enzyme-linked immunoassay (later samples) with a reported 95% accuracy at 40 days pregnant in bison (Biotracking, Inc, Moscow, ID).

To evaluate the relationship between fecal progestagens and pregnancy, we needed an independent measure of pregnancy against which to assess fecal progestagen levels. We defined "pregnant" as all adult females observed with a calf in the spring or summer following winter fecal sampling. We determined "cows with calves" by locating each marked reproductive-aged bison at least once weekly during the calving season (Berger and Cain 1999), and assigning calves to specific cows only after a suckling was observed (Berger and Cunningham 1994).

Because "absence of a calf" alone had a greater margin of error for including pregnant cows that lost calves through fetal or neonatal mortality, we defined as "non-pregnant" those cows that were both negative by palpation and were without a calf. We eliminated cows of uncertain pregnancy status (positive palpation, absence of calf) from our analysis because of the small potential for error by palpation. We recognized that using these criteria incurred a small potential

219 to include cows palpated falsely as not-pregnant that subsequently lost their fetus or neonate
220 (therefore were absent calf) in the "not-pregnant" group.
221
222 We modeled the relationship between pregnancy status and winter-collected (late gestation) fecal
223 progestagens (WP) for the marked cows meeting the criteria for pregnant or not-pregnant using
224 binary logistic regression (Hosmer and Lemeshow 2000). Logistic regression models the log of
225 the odds (or logit) of being pregnant as a linear function of explanatory variables, such as WP.
226 The odds of being pregnant are defined as the probability of being pregnant, divided by the
227 probability of not being pregnant. R statistical software was used for all analyses (R
228 Development Core Team 2007).
229
230 We compared two transformations of the WP explanatory variable to better meet assumptions
231 and increase the predictive ability of the logistic regression model: (1) natural log transformed
232 WP (ln (WP)), and (2) year-centered ln (WP). Year centering -- subtracting the year average
233 from each observation within that year -- served to standardize and account for substantial year-
234 to-year variability in the range of WP values (Fig. 1A). We evaluated the logistic regression
235 linearity assumption by calculating empirical logits for artificially created bins of WP (or
236 transformed WP) values for data collapsed over cows and years. Plots of the empirical logits
237 versus the explanatory variable clearly showed the linearity assumption was best met for year
238 centered ln(WP).
239
240 For each cow, the logistic regression model provided an estimated logit of the probability of
241 pregnancy, $\ln(p/(1-p))$, along with standard errors. We then obtained estimated probabilities of

pregnancy simply by inverting the logit transformation. Likewise, we calculated confidence intervals for probabilities by applying the inverse transformation to the endpoints of the confidence interval calculated on the logit scale (Agresti 2002).

Because our goal was to predict pregnancy status for non-handled cows, we focused our model assessment efforts on quantifying predictive success. Classification of a cow as pregnant or not-pregnant based on the magnitude of her WP required specification of a threshold probability. We chose the classification threshold minimizing the mis-prediction rate (MPR), as calculated using leave-one-out cross validation. That is, we removed one observation from the data set, fit the model, and then predicted pregnancy status for that observation, allowing for comparison of predicted status to true status. Repeating this for all observations allowed us to calculate MPR as the number of incorrect predictions divided by the total number of observations. We also calculated an average estimation error (AEE) using cross-validation by taking the absolute value of the difference between the estimated probability of pregnancy and the true response (1 for pregnant, 0 for not pregnant), averaged over all observations. For example, an estimated probability of 0.78 gave an estimation error of 0.22 if the cow was pregnant and 0.78 if the cow was not pregnant.

The use of MPR and AEE relied only on the point estimates of the probabilities and the chosen classification threshold, and did not take into account the width of the associated confidence intervals (CIs) for the probabilities. To assess this uncertainty in the context of prediction, we evaluated whether CIs crossed the classification threshold, coupled with whether the pregnancy status was correctly predicted. For correct predictions, it is undesirable for the CI to cross the

threshold, indicating a lack of confidence in a correct prediction. On the other hand, for incorrect predictions, it is undesirable for the CI to not cross the threshold, indicating high confidence in a wrong prediction. In such cases, simply relying on MPR would give a falsely optimistic or pessimistic view, respectively. We considered these bad predictive behaviors for a model, and calculated the proportion of 95% CIs in each scenario for a particular model, allowing comparison of these proportions across models with different covariates. Thus, we evaluated predictive ability by comparing predictions to known values, and utilizing the information in the CI regarding uncertainty in the estimate used to make the prediction.

We addressed several other issues related to the assumptions made in our statistical analyses. First, since data were collected across years, we assessed the appropriateness of assuming homogeneity in the relationship between (transformed) WP and the probability of pregnancy across years. Initially, we found clear year-to-year differences in the range of WP (Fig. 1 A). Factors that can modulate year-to-year variability in fecal steroid excretion include nutrition, disease, stage of pregnancy, and methodological variation (Cook et al. 2001, 2002). When seasonal sample collection intervals span several months, cows are sampled at different stages of gestation, and this can impact both within – and between-year variation in absolute hormone concentrations detected, and potentially impact the predictive value of hormonal measures (Garrott et al. 1998). Even when feces are collected during the same timeframe each year, variation in the timing of reproductive season onset could result in between-year differences in the stage of pregnancy being assessed (Monfort et al. 1993). Nutritional plane and variation in dietary fiber can also impact fecal steroid excretion, and species-specific differences exist, but these factors have not been systematically examined in bison. However, controlled nutritional

studies in elk revealed that even emaciated animals, near death, did not abort their fetuses, and fecal progestagen concentrations remained above the level indicative of pregnancy (see Cook et al. 2002). In addition to sampling strategies and physiology, intra- and inter-assay variation in hormone tests also contribute to the year-to-year variation (see reviews, Monfort 2003, Schwartz and Monfort 2008). To address this year-to-year variation from myriad factors, we centered the (transformed) WP values within each year by subtracting from each observation its associated year average (Fig. 1B). This provided a simple, yet effective, way to deal with the problem that can easily be applied to data collected in future years.

Second, we assessed the assumption of homogeneity among cows, since most cows were accompanied by multiple observations over time. By ignoring cow in the analysis, we assumed the relationship between probability of pregnancy and (transformed) WP was the same for each cow, and that each observation was independent, even if it came from the same cow in a different year. We relied on individual cow-specific values (Fig. 2) to assess the appropriateness of this assumption and concluded that after log transformation and centering there is no clear indication of heterogeneity among cows. For example, a visual cut-off of approximately zero for year centered ln(WP) was effective at separating pregnant from non-pregnant observations, regardless of the cow or year.

Third, we assessed the potential for serial dependence among responses within a cow. That is, within a cow, the pregnancy status may be more similar for successive (or close) years than for observations farther apart in time. Serial correlation in binary data is manifested by longer successive sequences (fewer runs) of 1's or 0's than would be expected by chance. We have

short time series, making this an inherently difficult assumption to assess (Fig. 2). Most runs containing more than one observation are for pregnant status, which is to be expected for cows of reproductive age in the absence of chronic or acute health issues where the probability of pregnancy is much greater than 0.5. A small number of cows in the sample could have had chronic or acute health issues that led to successive years of non-pregnant status. From Figure 2, it does not appear that observations within a run of non-pregnant observations are in general behaving differently from non-pregnant observations occurring amidst two runs of pregnant observations. However, we were not overly concerned with the potential violation of independence, maintaining that the information in each observation is important input into the model regardless of whether the animal was not pregnant for two or more successive years.

In the discussion of model assumptions, it is important to keep the ultimate goal of the model in sight. The primary purpose of the model is not to merely describe the relationship for this set of data, but to build a model to predict pregnancy status for unhandled and unmarked cows in any year. Thus, the identification of the cows and their reproductive history will be unknown, precluding the ability to effectively use random effects for cows or an estimate of serial correlation among observations within a cow. Therefore, coupling the goal of the model with the lack of clear evidence for heterogeneity among cows or years, led us to the decision to use all observations as independent events in building a model to estimate the probability of pregnancy from WP. We are confident the standard errors accompanying our estimates are not misleadingly small. With the potential naiveties of our model recognized, it strikes an important balance between statistical appropriateness, effective use of expensive data, and future practical usefulness.

Finally, since our study population was brucellosis infected, and brucellosis is known to heavily infect the placenta of bison (Cheville et al. 1998, Rhyan et al. 2001), an important producer of progesterone, we were concerned with the possibility that disease status could affect our fecal progestagen measurements. Our question focused on the possibility that fecal progestagen values were not only indicative of pregnancy status, but also of disease status, and that our model could misclassify pregnant animals as non-pregnant due to active brucellosis infections and reduced production of progesterone. Quantitative serology is known to be related to infection in greater Yellowstone area *Brucella*-infected bison (Roffe et al. 1999). To investigate the relationship between infection status and fecal progestagen, we ranked the *Brucella* antibody response of bison as either "high" or "not high" based on the complement fixation test (Roffe et al. 1999), and compared mean transformed WP by serological status, age, and pregnancy status using 128 cases for which we had all three variables. We found no evidence of an interaction between age and serological status ($p = 0.468$), and inconclusive evidence of a difference in mean WP between animals with "high" and "not high" titers after pregnancy status was accounted for (two-sided $p = 0.092$, Fig. 3). We estimated the mean WP for animals with "high" titers to be 0.23 units smaller than for those with "not-high" titers (95% CI: (-0.49, 0.037)). While we did find evidence of a slight decrease in WP for high-titered animals, we are confident that the magnitude was not large enough to negatively impact predictions from our model.

Procedures were approved by the Animal Care and Use Committee of Montana State University (proposal clearance #97-779) and the Biological Resources Division United States Geological Survey, and conformed to the Animal Welfare Act and United States Government principles for

the use and care of vertebrate animals used in testing, research, and training. Our work was conducted under approved permits by the National Park Service and Wyoming Game and Fish Department.

RESULTS

From 1997-2005 we assigned pregnancy status using our palpation and calf status criteria and obtained WP values for 155 observations of 42 unique bison; 114 were classified as pregnant and 41 as not pregnant. Years of data for individual cows ranged from 1 to 8 ($\bar{x}$ = 3.7). Of the 42 unique cows, 25 were not pregnant in at least one year and 34 were pregnant in at least one year. Fifteen cows had at least one observation for each status. Twelve of the 41 not pregnant observations came from cows below reproductive age (<2 yrs). We collected winter fecal samples and determined fecal progestagen levels for all samples. Of the 155 fecal samples, we collected 129 from immobilized animals and 26 from free-ranging individuals.

We obtained blood and conducted PSPB pregnancy analyses for 129 observations for which cows were assigned a pregnancy status. Twenty-five of the 129 comparisons between our palpation-calf assigned pregnancy status and PSPB derived pregnancy status were in disagreement (Table 1).

Rectal palpation was performed for 123 of the 155 observations that were assigned a pregnancy status (32 non-palpated cows were assigned solely based on presence of a calf). Of 41 animals determined not pregnant from palpation, 2 (4.9%) were subsequently documented with calves.

The 2 incorrect palpations occurred for observations with the second and third longest palpation-birth intervals (186 and 197 days; median of all intervals = 93 days, n = 81), which could have made detecting fetuses more difficult. Regardless, based on these data we estimated a false negative palpation rate across all palpation-birth intervals as about 5% in our study. In addition to 88 cows palpated as pregnant, subsequently observed with a calf, and assigned as pregnant in our analyses, we recorded 24 animals palpated as pregnant that were never observed with a calf. These observations were not used in fitting the model because they did not meet our criteria for assignment as pregnant or not pregnant. While these cases represented 22.6% (24/106) of observations recorded as pregnant from palpation, the potential for fetal or neonatal loss subsequent to palpations precluded us from estimating a false positive palpation error rate.

MPRs obtained using cross validation varied from 6.45 to 14.19% among the 4 logistic regression models we investigated (Table 2). The year-centered ln(WP) model had the lowest MPR (6.45%) and the lowest overall percentage of confidence intervals accompanying correct predictions crossing the threshold value (Table 2). Correct predictions were made with high confidence in this model and resulted in an overall successful pregnancy prediction rate of 93.5%.

The model incorporating date of fecal sample collection (WDAT) slightly reduced the AEE and the percent of confidence intervals accompanying mis-predictions that did not cross the threshold (Table 2). Fewer wrong predictions were made with this model, but many of these were made

399 with false confidence. On the other hand, the inclusion of WDAT increased the MPR from 6.45
400 % to 7.10 %. However, the small magnitudes of the differences among the models are probably
401 not practically meaningful.

402 The estimated logistic regression equation using year centered ln(WP) was:

403 *Logit* (p) = 1.685 + 3.940 * (YC ln (WP))

404 Standard errors for the regression coefficients were 0.338 (intercept) and 0.696 (slope)..
405 Performing the inverse logit transformation to obtain the equation for estimating the probability
406 of pregnancy we obtained:

407 p = exp (1.69 + 3.94 * (YC ln (WP))) / 1 + exp (1.69 + 3.94 * (YC ln (WP)))

408 The threshold for determining pregnant vs. not pregnant status resulting in the best predictive
409 success was 0.60, and this was used to make all predictions from the year centered ln(WP)
410 model.

411

412 Using the year centered ln(WP) model, 110 of the 114 observations assigned pregnant status
413 were correctly predicted as pregnant (Table 3). Of the 110, only three CIs crossed the threshold,
414 indicating a lack of confidence in the prediction. The model correctly predicted 35 of the 41 not
415 pregnant observations, with 3 confidence intervals crossing the threshold. Of the 10 mis-
416 predictions, 3 of the associated CIs overlapped the threshold, indicating an appropriate lack of
417 confidence in the prediction. For those not overlapping the threshold, indicating false confidence

418 in the predictions, most came very close to the threshold and thus in practice would be flagged as
419 uncertain predictions (Fig. 4).

420

421 In addition to estimating the probability of pregnancy for a particular cow given their fecal
422 progestagen levels, the logistic regression model allowed interpretation of the relationship
423 between the progestagen level and the odds of pregnancy. We found convincing evidence that
424 year centered fecal WP level was associated with pregnancy status ($Z_{Wald's}$ = 5.66, P < 0.001). A
425 doubling of year centered WP was associated with an estimated 15.4 (= $2^{3.94}$) fold increase in the
426 odds of a cow being pregnant (95% CI from 6.0 to 39.5 times).

427 **DISCUSSION**

428 Our logistic regression model provided predictions of mid- to late-gestation bison pregnancy
429 status based on one fecal sample with high accuracy (93.5%). Other studies have reported
430 successful pregnancy prediction rates based on fecal hormones of 100% in a small sample of
431 bison (14 pregnant, 4 non-pregnant based on palpation; Kirkpatrick et al. 1992), 85% in captive
432 moose (*Alces alces;* Monfort et al. 1993), approximately 97-100% in elk (Garrott et al. 1998,
433 Stoops et al. 1999), and 93-100% in bighorn sheep (*Ovis canadensis*; Borjesson et al. 1996,
434 Schoenecker et al. 2004). However, unlike these studies, using logistic regression we derived a
435 measure of confidence for each individual prediction, allowing practitioners to further evaluate
436 estimates for individual animals, thereby improving prediction interpretations.

437

438 Our study underscores the importance of independently validating fecal hormones with other
439 objective measures of pregnancy when developing pregnancy-fecal hormone relationships.

Using calf status alone as an indicator of pregnancy, for example, is not sufficient, because fetal and neonatal losses from predation, disease or other environmental factors are possible. In our study area, where brucellosis and predators capable of taking bison -- wolves and grizzly bears -- were common, we documented a 22.6% reduction between bison cows palpated as pregnant and those observed with a calf within about a week of birth. Fetal losses due to brucellosis infections in our population could partially explain this reduction, but abortion rates are generally low in chronically infected herds like the Jackson and Yellowstone bison herds (Cheville et al 1998). We also considered the possibility that our palpations could have contributed to fetal losses. To evaluate this, we compared fetal/neonatal loss rates in our bison that were and were not palpated, using the model presented herein to retrospectively determine pregnancy status of animals that were not palpated and not observed with calves. We found that 11/41 (26.9%) pregnant animals fell into this category, which was similar but slightly higher than the fetal/neonatal loss rate in our palpated sample (22.6%), suggesting that palpation had no effect. Neonatal predation could also have contributed, but our study was not designed to measure this. Nevertheless, our data make it clear that pregnancy estimates based on calf status alone can be prone to underestimating bias.

Transrectal uterine palpation is generally regarded as an accurate indicator of pregnancy status, but its reliability is highly dependent on the experience of the practitioner and the extent of fetal development (Weber et al. 1982). Using experienced veterinarians with mid-late gestation bison, we documented a 5% false negative palpation rate, which may have been related to expectations of a larger fetal size at this time in gestation during palpation. The two false negative palpations were in earlier stages of development than May parturient bison. We were unable to estimate

463 false positive rates because of confounding factors discussed above. While the accuracy of
464 palpation is probably sufficient for most studies, we recommend that it be coupled with or
465 substituted by ultrasonography whenever possible to improve our understanding of fecal
466 hormones as indicators of pregnancy.
467
468 Interestingly, based on our pregnancy assignment protocol using both palpation and calf status,
469 we found PSPB to be an unreliable indicator of pregnancy status and thus excluded it from our
470 analyses. PSPB, which Haigh et al. (1991) showed was 93% accurate in wood bison (*Bison*
471 *bison athabascae*), and which has been used in other studies as an independent indicator of
472 pregnancy status (Joly and Messier 2005, Garrott et al. 1998, Russell et al. 1998) disagreed with
473 19% (25/129) of our pregnancy assignments based on field measurements, including 10/36
474 (28%) assigned "not pregnant" by PSPB and confirmed to have calves later in the season (Table
475 1). The laboratory that conducted our analyses specified 95% accuracy in bison.
476
477 Detailed knowledge of the normal endocrine excretion dynamics of bison was essential for
478 determining the optimal time for sampling (i.e., early- versus mid- or late-pregnancy). In our
479 study, we used information derived from monthly assessments of urinary and fecal progestagen
480 in bison by Kirkpatrick et al. (1992). However, these data were limited to monthly mean
481 progestagens assessed in 14 cows sampled before conception through the first 4.5 months of
482 gestation (September, November, and January). Sampling strategies for pregnancy detection
483 would likely benefit from more comprehensive longitudinal assessments of fecal progestagens
484 excretion in pregnant and non-pregnant bison sampled before and throughout the entire period of
485 gestation. Aligning sampling effort with the gestational stage when females excrete maximal

fecal progestagens would increase the likelihood of discriminating pregnant from non-pregnant or cycling females (Monfort 1993). Furthermore, more intensive sampling, perhaps in a captive herd, would help to elucidate the extent of among-individual variation, including age-related effects on fecal hormone production.

MANAGEMENT IMPLICATIONS

Knowledge of pregnancy rates is often lacking in demographic models used to manage ungulate populations, but their inclusion would greatly increase the informative power, problem solving attributes, and overall utility of such models. The results of our study provide wildlife managers and researchers with an effective and efficient method for confidently estimating pregnancy in free-ranging bison, at scales ranging from individual animals to populations. The approach is attractive to practitioners because fecal samples necessary for evaluating progestagens are easily collected and preserved, laboratory procedures are well documented, and logistic regression is readily available in statistical computer software. Furthermore, samples can be obtained non-invasively, without the need to handle animals and incur associated drawbacks associated with immobilization, such as cost, human and animal safety, potential data bias, and lack of social acceptability. The benefits of the latter should not be underestimated in contemporary wildlife management settings, given the increasing concern among the general public and special interests for manipulation of wild animals (see Peterson et al. 2003 and Farnsworth and Rosovsky 1993). Our approach could be applied to other species as well.

ACKNOWLEDGMENTS

Funding was provided by the National Park Service, U.S. Fish and Wildlife Service, U.S. Geological Survey, and Wildlife Conservation Society. C. Cunningham, K. McFarland, K. Dhillon, and K. Gasaway collected fecal samples. K. Mashburn, D. Kersey and N. Presley

assisted with hormone analyses. C. Cunningham, K. McFarland, M. Reid, K. Cannon, S. Sweeney and K. Coffin assisted with bison immobilizations. B. Smith, B. Reiswig, and F. Escobedo helped with field logistics. The manuscript benefitted from the comments of ?? anonymous reviewers. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United States Government.

LITERATURE CITED

Agresti, A. 2002. Categorical Data Analysis, 2nd Edition. John Wiley & Sons, Inc. Hoboken, New Jersey.

Association for the Study of Animal Behavior. 2003. Animal Behaviour 65:249-255.

Ballard, W. B., and R. W. Tobey. 1981. Decreased calf production of moose immobilized with anectine administered from helicopter. Wildlife Society Bulletin 9:207-209.

Barrett, R. H. 1981. Pregnancy diagnosis with Doppler ultrasonic fetal pulse detectors. Wildlife Society Bulletin 9:60-63.

Bekoff, M., and D. Jamieson. 1996. Ethics and the study of carnivores: doing science while respecting animals. Pages 15-45 *in* J. L. Gittleman, editor. Carnivore behavior, ecology, and evolution, II. Cornell University Press, Ithaca, New York.

Berger, J. and C. Cunningham. 1994. Bison: mating and conservation in small populations. Columbia University Press. New York.

Berger, J., and S. L. Cain. 1999. Reproductive synchrony in brucellosis-exposed bison in the southern Greater Yellowstone Ecosystem and in non-infected populations. Conservation Biology 13:357-366.

530 Berger, J., W. Testa, T. Roffe, and S. L. Monfort. 1999. Conservation endocrinology: a
531 noninvasive tool to understand relationships between carnivore colonization and ecological
532 carrying capacity. Conservation Biology 13:980-989.
533 Berger, J., K. M. Murray, Bayarbaatar Buuveibaatar, M. R. Dunbar, and Badamjav
534 Lkhagvasuren. 2010. Capture of ungulates in central Asia using drive nets: advantages
535 and pitfalls as illustrated by endangered Mongolian Saiga. Oryx. Available on CJO 20
536 Aug 2010 doi:10.1017/S003060531000058X
537 Borjesson, D. L., W. M. Boyce, I. A. Gardner, J. DeForge, and B. Lasley. 1996. Pregnancy
538 detection in bighorn sheep (*Ovis Canadensis*) using a fecal-based enzyme immunoassay.
539 Journal of Wildlife Diseases 32:67-74.
540 Cheville, N. F., D. R. McCullough, and L. R. Paulson. 1998. Brucellosis in the Greater
541 Yellowstone Ecosystem. National Research Council, National Academy Press,
542 Washington, D.C.
543 Cook, R. C., D. L. Murray, J. G. Cook, P. Zager, and S. L. Monfort. 2001. Nutritional
544 influences on breeding dynamics in elk. Canadian Journal of Zoology 79:845-853
545 Cook, R. C., J. G. Cook, R. A. Garrott, L. L. Irwin, and S. L. Monfort. 2002. Effects of diet and
546 body condition on fecal progestagen in elk. Journal of Wildlife Diseases 38: 558-565.
547 Coulson, T, J.-M. Gaillard, and M. Festa-Bianchet. 2005. Decomposing the variation in
548 population growth into contributions from multiple demographic rates. Journal of Animal
549 Ecology 74:789-801.
550 Dein, F. J., D. E. Toweill, and K. P. Kenow. 2005. Care and use of wildlife in field research.
551 Pages 185-196 *in* C. E. Braun, editor. Techniques for wildlife investigations and
552 management. Sixth edition. The Wildlife Society, Bethesda, Maryland, USA.

553 Dunk, J. R., R. N. Smith, and S. L. Cain. 1997. Nest-site selection and reproductive success in
554 common ravens. Auk 114:116-120.
555 Farnsworth, E. J. and J. Rosovsky. 1993. The ethics of ecological field experimentation.
556 Conservation Biology 7:463-472.
557 Fuller, J. A., R. A. Garrott, P. J. White, K. E. Aune, T. J. Roffe, and J. C. Rhyan. 2007.
558 Reproduction and survival of Yellowstone bison. Journal of Wildlife Management
559 71:2365-2372.
560 Garrott, R. A., S. L. Monfort, K. L. Mashburn, P. J. White, and J. G. Cook. 1998. One-sample
561 pregnancy diagnosis in elk using fecal steroid metabolits. Journal of Wildlife Diseases
562 34:126-131.
563 Graham, L., F. Schwarzenberger, E. Mostl, W. Galama, and A. Savage. 2001. A versatile
564 enzyme immunoassay for the determination of progestogens in feces and serum. Zoo
565 Biology 20:227-236.
566 Grand Teton National Park and National Elk Refuge. 1996. Jackson bison herd: long term
567 management plan and environmental assessment. Grand Teton National Park, Moose,
568 Wyoming, USA.
569 Haigh, J. C. 1990. Opiods in zoological medicine. Journal of Zoo and Wildlife Medicine
570 21:391-413.
571 Haigh, J. C., C. Gates, A. Ruder, and R. Sasser. 1991. Diagnosis of pregnancy in wood bison
572 using a bovine assay for pregnancy-specific protein B. Theriogenology 36:749-754.
573 Haigh, J. C., M. Cranfield, and R. G. Sasser. 1988. Estrus synchronization and pregnancy
574 diagnosis in red deer. Journal of Zoo Animal Medicine 19:202-207.

575 Hein, R. G., J. L. Musser, and E. F. bracken. 1991. Serologic, parasitic and pregnancy survey of
576 the Colockum elk herd in Washington. Northwest Science 65:217-222.
577 Hosmer, D. W., and S. Lemeshow. 2000. Applied Logistic Regression, 2nd Edition, John Wiley
578 & Sons Inc., New York, USA.
579 Houston, D. B., C. T. Robbins, C. A. Ruder, and R. G. Sasser. Pregnancy detection in mountain
580 goats by assay for pregnancy-specific protein B. Journal of Wildlife Management 50:740-
581 742.
582 Jessup, D. A., R. K. Clark, R. A. Weaver, and M. D. Kock. 1988. The safety and cost
583 effectiveness of net-gun capture of desert bighorn sheep (*Ovis Canadensis nelson*).
584 Journal of Zoo Animal Medicine 19:208-213.
585 Joly, D. O., and F. Messier. 2005. The effect of bovine tuberculosis and brucellosis on
586 reproduction and survival of wood bison in Wood Buffalo National Park. Journal of
587 Animal Ecology 74:543-551.
588 Kirkpatrick, J. F., S. E. Shideler, and J. W. Turner, Jr. 1990. Pregnancy determination in
589 uncaptured feral horses based on free steroids in feces and steroid metabolites in urine-
590 soaked snow. Canadian Journal of Zoology 68:2576-2579.
591 Kirkpatrick, J. F., K. Bancroft, and V. Kincy. 1992. Pregnancy and ovulation detection in bison
592 (*Bison bison*) assessed by means of urinary and fecal steroids. Journal of Wildlife
593 Diseases 28:590-597.
594 Kirkpatrick, J. F., D. F. Gudermuth, R. L. Flagan, J. C. McCarthy, and B. L. Lasley. 1993.
595 Remote monitoring of ovulation and pregnancy of Yellowstone bison. Journal of Wildlife
596 Management 57:407-12.

Larsen, D. G., and D. A. Gauthier. 1989. Effects of capturing pregnant moose and calves on calf survivorship. Journal of Wildlife Management 53:564-567.

Lasley, B. L., and J. F. Kirkpatrick. 1991. Monitoring ovarian function in captive and free-roaming wildlife by means of urinary and fecal steroids. Journal of Zoo and Wildlife Medicine 22:23-31.

Meagher, M. M. 1973. The bison of Yellowstone National Park. National Park Service Scientific Monograph Series Number One. Washington D.C.

Messier, F., D. M. Desaulniers, A. K. Goff, R. Nault, R. Patenaude, and M. Crete. 1990. Caribou pregnancy diagnosis from immunoreative progestins and estrogens excreted in feces. Journal of Wildlife Management 54:279-283.

Monfort, S. L., N. P. Arthur, and D. E. Wildt. 1990. Monitoring ovarian function and pregnancy by evaluating excretion of urinary oestrogen conjugates in semi-free-ranging Przewalski's horses (*Equus przewalskii*). Journal of Reproduction and Fertility 91:155-164.

Monfort, S. L., C. C. Schwartz, and S. K. Wasser. 1993. Monitoring reproduction in captive moose using urinary and fecal steroid metabolites. Journal of Wildlife Management 57:400-407.

Monfort, S. L. 2003. Non-Invasive Endocrine Measures of Reproduction and Stress in Wild Populations. Pages 147-165 *in* D. E. Wildt, W. Holt, and A. Pickard editors. Reproduction and integrated conservation science. Cambridge University Press, Cambridge, United Kingdom.

Peterson, M. N., R. R. Lopez, P. A. Frank, M. J. Peterson, and N. J. Silvy. 2003. Evaluating capture methods for urban white-tailed deer. Wildlife Society Bulletin 31:1176-1187

Poole, J.H., L. H. Kasman, E. C. Ramsay, and B. L. Lasley. 1984. Musth and urinary testosterone in the African elephant (*Loxodonta africana*). Journal of Reproduction and Fertility 70:255-290.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL http://www.R-project.org.

Ramsay, M. A., and R. M. F. S. Sadleir. 1979. Detection of pregnancy in living bighorn sheep by progestin determination. Journal of Wildlife Management 43:970-973.

Rhyan, J. C., T. Gidlewski, T. J. Roffe, K. Aune, L. M. Philo, and D. R. Ewalt. 2001. Pathology of brucellosis in bison from Yellowstone National Park. Journal of Wildlife Diseases 37:101-109.

Roffe, T., J, J. C. Rhyan, K. Aune, L. M. Philo, D. R. Ewalt, T. gidlewski, and S. G. Hennager. 1999. Brucellosis in Yellowstone National Park bison: quantitative serology and infection. Journal of Wildlife Management 63(4):1132-1137

Roffe, T. J., K. Coffin, and J. Berger. 2001. Survival and immobilizing moose with carfentanil and xylazine. Wildlife Society Bulletin. 29:1140-1146.

Roffe, T. J., S. J. Sweeney, and K. Aune. 2005. Chemical immobilization of North American mammals. Pages 286-302 *in* C. E. Braun, editor. Wildlife Investigational Techniques. Sixth Edition. The Wildlife Society, Bethesda, Maryland, U.S.A.

Russell, D. E., K. L. Gerhart, R. G. White, and D. Van De Wetering. 1998. Detection of early pregnancy in caribou: evidence for embryonic mortality. Journal of Wildlife Management 62:1066-1075.

641 Sasser, R. G., C. A. Ruder, K. A. Ivani, J. E. Butler, and W. C. Hamilton. 1986. Detection of
642 pregnancy by radioimmunoassay of a novel pregnancy-specific protein in serum of cows
643 and a profile of serum concentrations during gestation. Biology of Reproduction 35:936-
644 942.
645 Schoenecker, K. A., R. O. Lyda, and J. Kirkpatrick. 2004. Comparison of three fecal steroid
646 metabolites for pregnancy detection used with single sampling in bighorn sheep (*Ovis*
647 *canadensis*). Journal of Wildlife Disease 40:273-281.
648 Schwartz, M. K., and S. L. Monfort. 2008. Genetic and Endocrine Tools for Carnivore Surveys.
649 Pages 228-250 *in* R. A. Long, P. MacKay, J. C. Ray, and W. J. Zielinski, editors.
650 Noninvasive survey methods for North American carnivores. Island Press, Washington
651 D.C.
652 Smith, R. N., S. L. Cain, S. H. Anderson, J. R. Dunk, and E. S. Williams. 1998. Blackfly-
653 induced mortality of nestling red-tailed hawks. Auk 115:368-375.
654 Stephenson, R. R., J. W. Testa, G. P. Adams, R. G. Sasser, C. C. Schwartz, and K. J.
655 Hundertmark. 1995. Diagnosis of pregnancy and twinning in moose by ultrasonography
656 and serum assay. Alces 31:167-172.
657 Stoops, M. A., G. B. Anderson, B. L. Lasley, and S. E. Shideler. 1999. Use of fecal steroid
658 metabolites to estimate the pregnancy rate of a free-ranging herd of tule elk. Journal of
659 Wildlife Management 63:561-569.
660 Testa, J. W. 2004. Population dynamics and life history trade-offs of moose (*Alces alces*) in
661 south-central Alaska. Ecology 85:1439–1452.

662 Thorne, E. T. 1982. Brucellosis. Pages 54-63 *in* E. T. Thorne, N. Kingston, W. R. Jolley, and R.
663 C. Bergstrom, editors. Diseases of wildlife in Wyoming. Second edition. Wyoming Game
664 and Fish Department, Cheyenne, USA.
665 United States Department of Agriculture, Animal and Plant Health Inspection Service. 2003.
666 Brucellosis Eradication: Uniform Methods and Rules, Effective October 1, 2003. APHIS
667 publication 91-45-013.
668 Valkenburg, P., R. O. Boertje, and J. L. Davis. 1983. Effects of darting and netting on caribou
669 in Alaska. Journal of Wildlife Management 47:1233-1237.
670 Wasser, S. K., S. L. Monfort, J. Southers, and D. E. Wildt. 1994. Excretion rates and
671 metabolites of oestradiol and progesterone in baboon *(Papio cynocephalus cynocephalus)*
672 faeces. Journal of Reproduction and Fertility 101:213-220.
673 Webber, B. J., M. L. Wolfe, and G. C. White. 1982. Use of serum progesterone levels to detect
674 pregnancy in elk. Journal of Wildlife Management 46:835-838.
675 Welsh, T. H., and B. H. Johnson. 1981. Stress-induced alterations in secretion of
676 corticosteroids, progestins, luteinizing hormone and testosterone in bulls. Endocrinology
677 109:185-190.
678 Western Regional Climate Center. 2009. Western Regional Climate Center,
679 http://www.wrcc.dri.edu. Accessed 4 Mar 2009.
680 White, P. J., R. A. Garrott, J. F. Kirkpatrick, and E. V. Berkeley. 1995. Diagnosing pregnancy
681 in free-ranging elk using fecal steroid metabolites. Journal of Wildlife Diseases 31:514-
682 522.

683 Williams, E. S., E. T. Thorne, S. L. Anderson, and J. D. Herriges, Jr. 1993. Brucellosis in free-
684 ranging bison (*Bison bison*) from Teton County, Wyoming. Journal of Wildlife Diseases
685 29:118-122.
686 Wood, A. K., R. E. Short, A. Darling, G. L. Dusek, R. G. Sassar, and C. A. Ruder. 1986. Serum
687 Assays for Detecting Pregnancy in Mule and White-Tailed Deer. Journal of Wildlife
688 Management 50: 684-687

689

690 **Tables:**

691 Table 1. Comparison of pregnancy determination by pregnancy specific protein B (PSPB) from
692 blood samples and a combination of rectal palpation and annual calf status (Assigned Status) in
693 Jackson bison, northwest Wyoming, USA, 1997-2005 (shaded cells show disagreement between
694 the two methods).

	Assigned Status		
PSPB Results	Not pregnant	Pregnant	Total
Not Pregnant	26	10	36
Pregnant	15	78	93
Total	41	88	129

695

696

Table 2. Leave-one-out cross-validation measures of pregnancy status predictive success using winter collected (late gestation) fecal progestins (WP) in 4 logistic regression models for Jackson bison, northwest Wyoming, USA, 1997-2005 (YC = year-centered, WDAT = sample collection date) .

Model	% mis-predicted	Average Estimation Error (AEE)	% mis-prediction CIs not overlapping cutoff	% correct prediction CIs covering cutoff	Pregnancy Probability cut-off value
WP	14.19	0.283	4.52	27.74	0.5
ln(WP)	12.26	0.223	5.81	13.55	0.5
YC ln(WP)	6.45	0.123	4.52	3.87	0.6
YC ln(WP) + WDAT	7.10	0.117	3.23	5.81	0.5

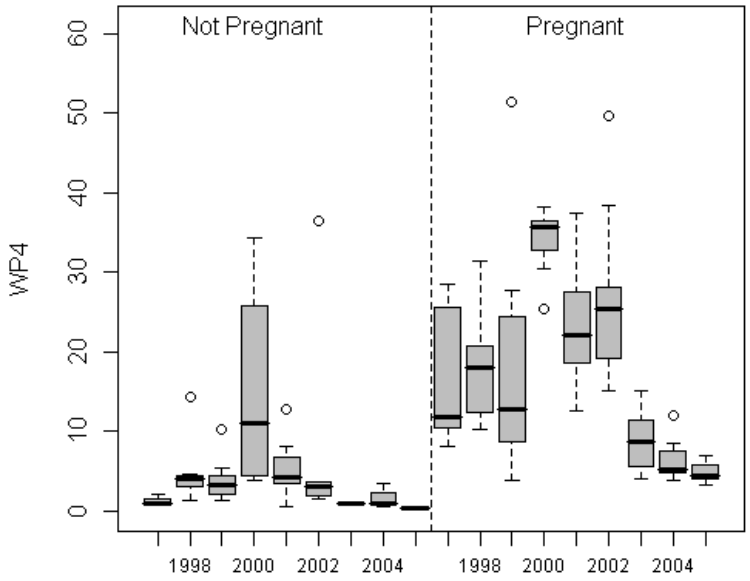
704 Table 3. Pregnancy predictions using logistic regression and year-centered, log transformed
705 values of winter collected (late gestation) fecal progestins (WP) from Jackson bison, northwest
706 Wyoming, USA, 1997-2005.

Model Predictions	Assigned Pregnancy Status		
	Pregnant	Not Pregnant	Total
CIs do not overlap cutoff:			
Pregnant	107	5	112
Not Pregnant	2	32	34
CIs overlap cutoff:			
Pregnant	3	1	4
Not Pregnant	2	3	5
Total	114	41	155

707

708

709 **Figures:**



710

711 Figure 1A.

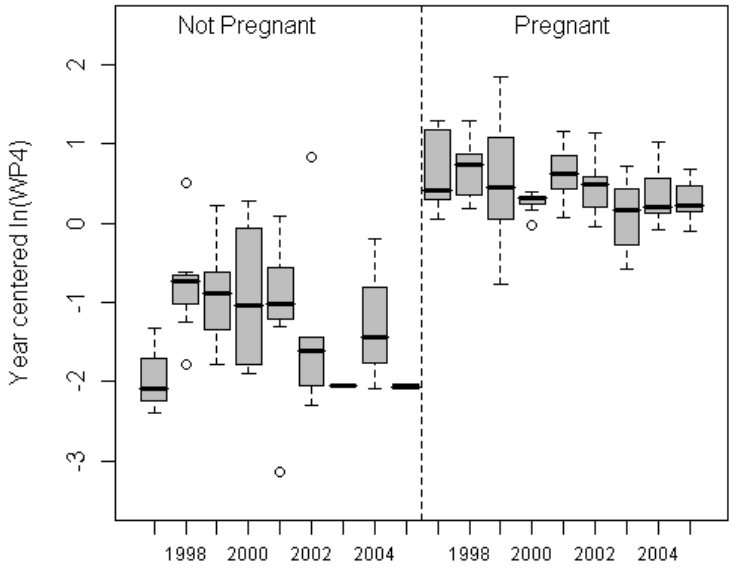


Figure 1B.

Figure 1. Winter collected (late gestation) fecal progestin levels (WP) (A), and year-centered ln(WP) (B) by year and assigned pregnancy status for Jackson bison, northwest Wyoming, USA, 1997-2005 (dark horizontal lines in each box represent medians, boxes span the 0.25 to 0.75 quantile, representing the interquartile range [IQR], whiskers extend to min and max observations within 1.5 times the IQR, and points denote observations > 1.5 times the IQR).

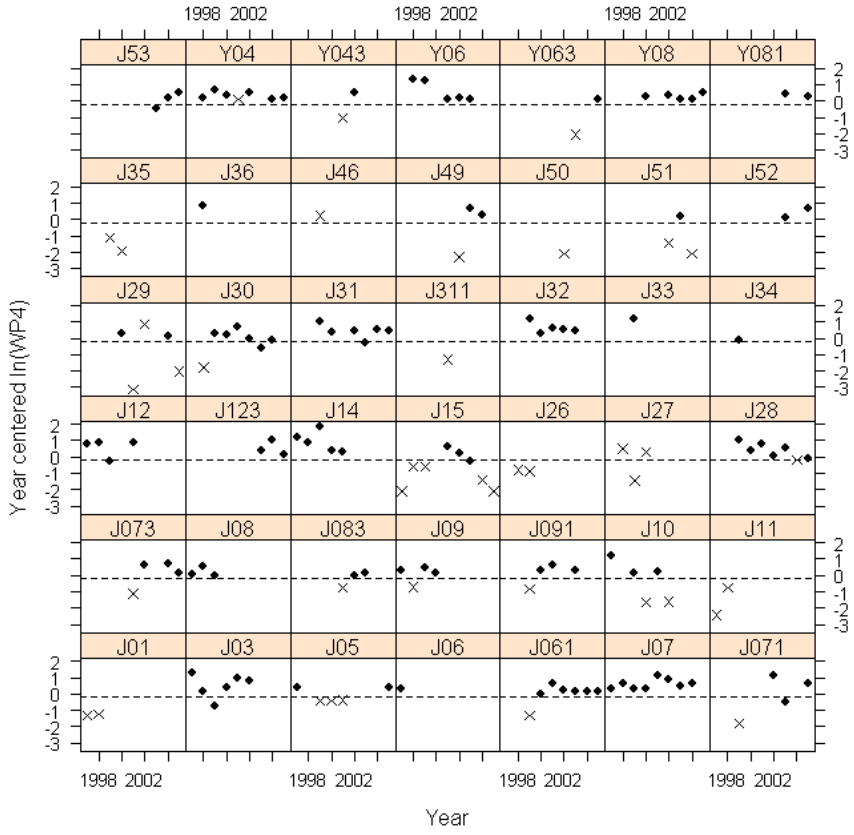
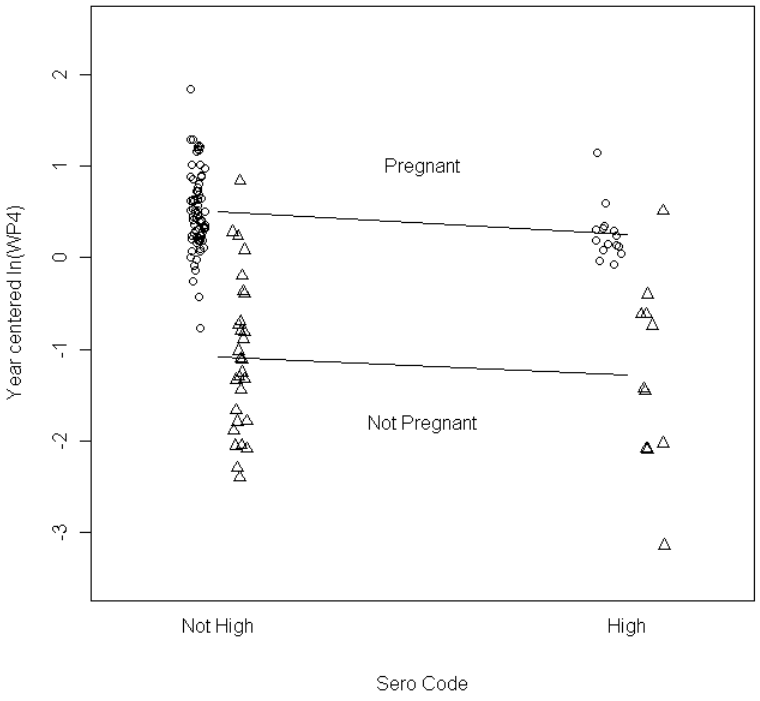


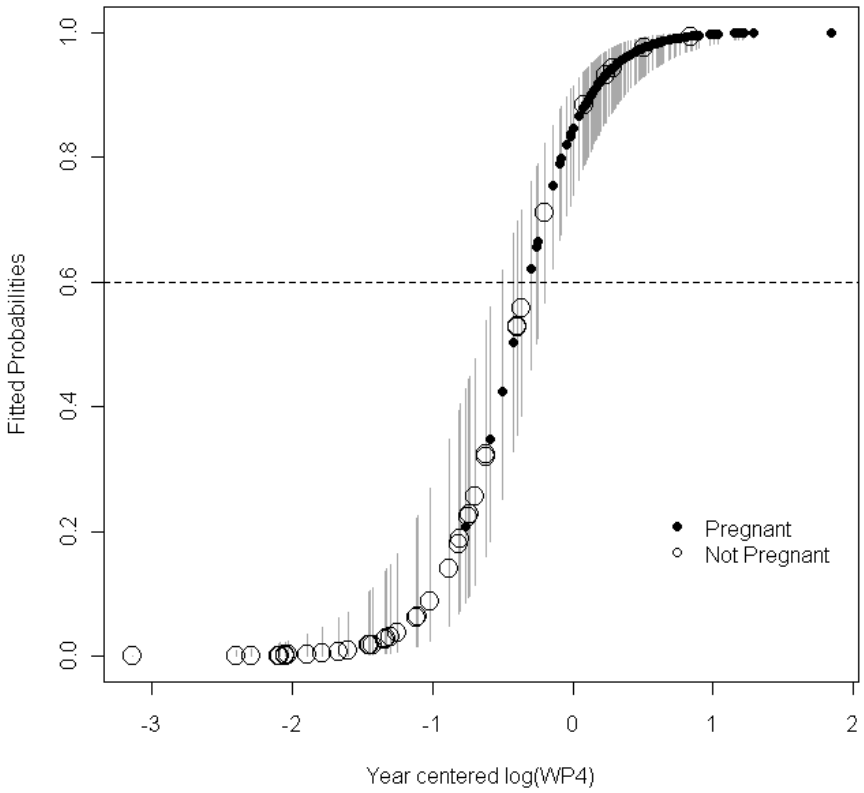
Figure 2. Year-centered, log transformed values of winter collected (late gestation) fecal progestins (WP) by year and assigned pregnancy status (solid diamond = pregnant, X = not pregnant) for 42 female bison from the Jackson bison herd, northwest Wyoming, USA, 1997-2005.



730

731 Figure 3. Year centered ln(WP) values grouped by disease status (Sero Code) and coded by
732 pregnancy status (circle = pregnant, triangle = not pregnant). The lines connect the average
733 for each disease status within pregnancy status.

734



735

736 Figure 4. Fitted assigned pregnancy probabilities and confidence intervals for the logistic
737 regression model using year centered, log transformed values of winter collected (late
738 gestation) fecal progestins (WP) for Jackson bison, northwest, Wyoming, USA, 1997-2005
739 (dashed line represents proposed cut off for model-predicted pregnancy status).

740

741