Prediction of ossification from live and carcass traits in young beef cattle: model development and evaluation¹

Boyd W. Gudex,^{*,†} Malcolm J. McPhee,^{*,2,®} Victor H. Oddy,^{*} and Brad J. Walmsley^{*,‡}

^{*}NSW Department of Primary Industries, Livestock Industries Centre, University of New England, Trevenna Road, Armidale, New South Wales 2351, Australia; [†]Present address: Agricultural Business Research Institute, University of New England, Armidale, New South Wales 2351, Australia; and [‡]Animal Genetics and Breeding Unit³, University of New England, Armidale, New South Wales 2351, Australia

ABSTRACT: Physiological maturity, measured as carcass ossification [10 unit increments (100, 110, 120, ...)], is used by the United States Department of Agriculture and the Meat Standards Australia carcass grading systems to reflect age-associated differences in beef tenderness and determine producer payments. In most commercial cattle herds, the exact age of animals is unknown; thus, prediction of ossification in association with phenotypic prediction systems has the capacity to assist producer decision making to improve carcass and eating quality. This study developed and evaluated prediction equations that use either live animal or carcass traits to predict ossification for use in phenotypic prediction systems to predict meat quality. The average ossification in the model development dataset was 138 with a SD of 21 and a range between 100 and 200. Model development involved regressing various combinations of live animal traits: age at recording, sex, live weight (BW), average daily gain, ultrasound scanned eye muscle area, 12/13th rib and subcutaneous P8 rump fat thickness; or carcass traits: age at slaughter, sex, hot standard carcass weight (HSCW), carcass eye muscle area, marble score, rib, and P8 rump fat (CP8) thickness, against ossification. The models

were challenged with data from 3 independent datasets: 1) Angus steers produced by divergent selection for visual muscle score; 2) temperate (Angus, Hereford, Shorthorn and Murray Grey) steers and heifers; and 3) tropically adapted (Brahman and Santa Gertrudis) steers and heifers. Five models with adjusted R^2_{adj} above 0.55 were evaluated. When challenged with dataset 1, the absolute mean bias (MB) and root mean square error of prediction (RMSEP) ranged from 0.1 to 4.2, and 9.8 to 10.7, which are within the bounds of the 10 point increment on the ossification scale. When subsequently challenged with dataset 2, MB and RMSEP ranged from 2.8 to 13.4, and 19.6 to 23.7, respectively; and with dataset 3, MB and RMSEP ranged from 14.4 to 17.5, and 23.3 to 31.9, respectively. Generally, when compared in relation to the ossification scale, all evaluated models had similar accuracy. For predicting meat quality, the model containing live animal traits considered most useful was $[85.35 + 0.16 \times BW + 10.94 \times \text{sex} - 0.09 \times \text{sex} \times BW]$ (adjusted $R^2 = 0.59$; SE = 13.51) and the most useful model containing carcass traits was $[107.15 + 11.53 \times \text{sex} + 1.10 \times \text{CP8} + 0.16 \times \text{HSCW} - 0.15 \times \text{sex} \times \text{HSCW}]$ (adjusted $R^2 = 0.60$; SE = 13.39).

Key words: beef cattle, carcass, live animal, maturity, meat quality, ossification

© The Author(s) 2018. Published by Oxford University Press on behalf of the American Society of Animal Science. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

J. Anim. Sci. 2019.97:144–155
doi: 10.1093/jas/sky422

¹The authors acknowledge all participants, both scientists and technical staff, who contributed to or supported the work, including those involved in experimental design, cattle management, data collection, and data handling. Funding for this study was supplied by Meat and Livestock Australia.

²Corresponding author: malcolm.mcphee@dpi.nsw.gov.au

³The Animal Genetics and Breeding Unit is a joint institute of NSW Department of Primary Industries and the University of New England.

Received June 14, 2018.

Accepted October 30, 2018.

INTRODUCTION

Market signals provided the emphasis during the 1990's and 2000's for the beef industry to increase the focus on improving carcass and meat quality (Johnston et al., 2003). This is embodied in carcass grading systems developed by the United States Department of Agriculture (USDA, 1997) and Meat Standards Australia (MSA, Watson et al., 2008). Carcass ossification is used by both the MSA and USDA carcass grading systems when predicting the eating quality of beef which impacts on the payments made to producers. Ossification is an assessment of the calcification of the spinous processes in the sacral, lumbar, and thoracic vertebrae (AUS-MEAT, 2005; Watson et al., 2008) and is used within grading systems as an assessment of physiological maturity. Physiological maturity is used to reflect age-associated differences in beef tenderness (Schonfeldt and Strydom, 2011). In live commercial cattle, exact age or other measures of physiological maturity are often unknown.

The predictions made by carcass grading systems can be used in association with phenotypic prediction systems, such as the BeefSpecs calculator (Walmsley et al., 2014), to provide predictions of eating quality up to 200 d prior to slaughter. This has the capacity to assist producers make management decisions to improve compliance rates with carcass market specifications and improve eating quality. To achieve this, the ossification input required by the carcass grading systems could be predicted from the live animal [e.g., age or live weight (BW)] or carcass traits [e.g., hot standard carcass weight (HSCW) or carcass P8 rump fat (CP8)] that are provided as outputs by the BeefSpecs calculator. The authors are unaware of any models for predicting ossification.

The objective of this study was to develop a prediction equation that uses either live animal or carcass characteristics to predict ossification which would be used as an input for phenotypic prediction of meat quality on live animals.

MATERIALS AND METHODS

Animal care and use committee approval was not obtained for this study because data were obtained from existing databases [New South Wales Department of Primary Industries (NSW DPI) and co-operative research center (CRC) for Beef Genetic Technologies databases].

This study developed models that describe the relationships between carcass ossification and either live animal or carcass traits in beef cattle.

Data used during model development were taken from a growth and slaughter experiment that used progeny born in 2012 from the NSW DPI muscling herd. Datasets used for model evaluation were taken from the NSW DPI muscling herd (born 2011) and the CRC for Cattle and Beef Quality (born 1993 to 1998) separated on the basis of *Bos taurus* or *Bos indicus* breed type (Robinson, 1995, Upton et al., 2001).

Data for Model Development

Thirty-six Angus heifers and 58 Angus steers (94 animals in total) were used to develop ossification prediction models. Cafe et al. (2018) described the design and management of the NSW DPI muscling herd. In brief, the NSW DPI muscling herd was established in 1992 from an unselected Hereford herd that was mated to Angus bulls selected on visual muscle score (McKiernan, 2007). To create divergent muscle lines, the female progeny and subsequent generations were also selected on muscle score and mated to Angus bulls purchased from within the commercial Australian beef industry again based on visual muscle score (Walmsley and McKiernan, 2011).

The 2012 born cohort used to develop the models in this study were backgrounded following weaning on temperate pastures in North Eastern New South Wales. The animals were serially slaughtered during late 2013 and 2014 in 3 groups at average ages of 448, 532, and 626 d, respectively. The final slaughter group was grain finished for 97 d prior to slaughter.

All animals had BW (kg), visual muscle scores (McKiernan 2007), sex (heifer or steer), birth, and measurement dates recorded. The birth and recording dates were used to calculate age at recording (AGEREC, days) and ADG (kg/d) between weaning and slaughter for each individual.

Ultrasound scanning was used to measure rib fat (USRIB, mm), subcutaneous rump fat (USP8, mm), and eye muscle area (USEMA, cm²) on the live animal. These traits and the corresponding BW were recorded between 3 and 5 times for each animal, including immediately prior to slaughter, with individuals killed at older ages having more opportunities for measurement. The USRIB and USEMA (*M. longissimus thoracis et lumborum*) measurements were taken at the 12th and 13th ribs (Barwick et al., 2009), the USP8 measurement was taken at the intersection between a line parallel to the spine from the *tuber ischium* and a line perpendicular to it from the spinous process of the third sacral vertebra (Johnston

et al., 2003; Wolcott et al., 2009). All ultrasound measurements were taken by an accredited technician (Upton et al., 1999) with an Aloka 500V real-time ultrasound scanner using a 17-cm transducer (Corometrics Medical Systems Inc., Wallingford, CT), with vegetable oil as the coupling agent, and the fat depths were recorded using callipers built into the scanner (Wolcott et al., 2001).

Following slaughter, carcass measurements were taken by MSA grading and AUS-MEAT chiller assessment accredited graders. Ossification score, the degree of conversion of cartilage to bone, was measured at the sacral, lumbar, and thoracic vertebrae on a scale between 100 and 590 in increments of 10 (Wolcott et al., 2009). Hot standard carcass weight was measured in kilograms from untrimmed carcasses. Carcass EMA (CEMA) was measured using the AUS-MEAT EMA standard grid as the number of square centimetres (cm²) of *longissimus thoracis et lumborum* at the quartering site. Carcass rib fat (CRIB, mm) was measured as the depth of subcutaneous fat at the quartering site in the chilled carcass approximately 75% of the way along the rib eye muscle. Carcass P8 rump fat was measured in millimeters as the depth of subcutaneous fat at the intersection of a line parallel to the spine, from the *tuber ischium*, and a line perpendicular to it, from the spinous process of the third sacral vertebra (Johnston et al., 2003; Wolcott et al., 2009). MSA marbling score, a measure of the fat deposited between myofiber bundles in the rib eye muscle, was

subjectively assessed using a scale between 100 and 1,100 in increments of 10. Marbling was assessed at the quartering site of the chilled carcass a minimum of 45 min after the carcass was quartered and is calculated by evaluating the amount, piece size, and distribution of marbling in comparison to the MSA reference standards (AUS-MEAT, 2005; Wolcott et al., 2009).

Animals with missing measurements were removed from the dataset. Summary statistics for the traits recorded on the live animal and carcass are shown in Table 1.

Data for Model Evaluation

Three datasets were used to evaluate the models developed and are described below. The traits were recorded in a similar manner to that described for the development dataset. Summary statistics for OSS, AGEREC, BW, AGESL, HSCW, and CP8 in these datasets are given in Table 2.

Dataset 1. Data were obtained from an earlier cohort of the NSW DPI muscling herd as described above for the development dataset. This dataset contained 74 steers born in 2011 that were managed/recorded in a similar manner to the animals in the development dataset and slaughtered during early 2013.

Dataset 2. Data from *Bos taurus* breed types (Angus, Hereford, Shorthorn and Murray Grey)

Table 1. Summary statistics for the live animal and carcass traits recorded on Angus heifers and steers in the development dataset

Sex	Live Animal						Carcass					
	Trait ¹	<i>n</i>	Min.	Max.	Avg.	SD	Trait ²	<i>n</i>	Min.	Max.	Avg.	SD
Heifer	AGEREC (d)	36	434	647	531.6	72.5	AGESL (d)	36	436	649	533.6	72.5
	BW (kg)	36	309	640	427.3	98.5	HSCW (kg)	36	146	332	215.8	56.8
	ADG (kg/d)	36	0.34	0.92	0.59	0.17	CRIB (mm)	36	1	12	6.1	3.5
	USRIB (mm)	36	2	14	7.1	3.2	CP8 (mm)	36	1	27	11.4	7.0
	USP8 (mm)	36	3	20	10.6	4.9	Marble Score	36	200	510	329.2	74.4
	USEMA (cm ²)	36	50	84	64.6	9.9	CEMA (cm ²)	36	44	99	66.1	13.3
	Muscle Score	36	3	11	7.2	2.1	OSS	36	120	200	153.3	21.1
Steer	AGEREC (d)	58	420	648	531.4	75.1	AGESL (d)	58	422	650	533.4	75.1
	BW (kg)	58	327	740	493.9	111.5	HSCW (kg)	58	163	380	250.6	63.5
	ADG (kg/d)	58	0.29	1.15	0.70	0.20	CRIB (mm)	58	1	13	5.0	3.4
	USRIB (mm)	58	2	14	6.3	3.2	CP8 (mm)	58	1	26	7.8	6.2
	USP8 (mm)	58	2	17	8.1	4.4	Marble Score	58	190	480	305.7	59.7
	USEMA (cm ²)	58	52	92	69.5	11.3	CEMA (cm ²)	58	43	110	69.4	13.5
	Muscle Score	58	2	13	7.8	2.6	OSS	58	100	150	129.0	14.8

¹AGEREC = age at recording, BW = live weight, ADG = average daily gain from weaning to slaughter, USRIB = ultrasound scanned subcutaneous 12/13th rib fat thickness, USP8 = ultrasound scanned subcutaneous rump fat thickness and USEMA = ultrasound scanned eye muscle area.

²AGESL = age at slaughter, HSCW = hot standard carcass weight, CRIB = carcass subcutaneous 12/13th rib fat thickness, CP8 = carcass subcutaneous rump fat thickness, CEMA = carcass eye muscle area and OSS = ossification score.

Table 2. Summary statistics for age at measurement (AGEREC), live weight (BW), age at slaughter (AGESL), hot standard carcass weight (HSCW), carcass P8 fat depth (CP8), and ossification (OSS) in each evaluation dataset

Sex	Trait	Dataset 1					Dataset 2					Dataset 3				
		<i>n</i>	Min.	Max.	Avg.	SD	<i>n</i>	Min.	Max.	Avg.	SD	<i>n</i>	Min.	Max.	Avg.	SD
Heifer	Live															
	AGEREC (d)	-	-	-	-	-	83	496	821	617.2	93.0	404	307	1131	729.6	156.8
	BW (kg)	-	-	-	-	-	83	342	614	472.7	75.8	404	172	678	434.7	85.0
	Carcass															
	AGESL (d)	-	-	-	-	-	83	501	838	625.5	95.7	404	429	1153	747.7	148.6
	HSCW (kg)	-	-	-	-	-	83	155	324	238.3	48.4	404	140	358	236.0	41.0
	CP8 (mm)	-	-	-	-	-	83	2	22	10.0	4.8	404	3	30	12.3	5.0
	OSS	-	-	-	-	-	83	140	200	169.8	16.8	404	120	200	171.0	17.4
Steer	Live															
	AGEREC (d)	74	519	690	609.9	52.5	1202	340	961	630.3	154.2	482	318	1290	756.5	173.1
	BW (kg)	74	391	736	538.1	87.3	1202	232	762	504.7	104.5	482	216	778	519.3	97.7
	Carcass															
	AGESL (d)	74	544	693	623.9	42.2	1202	343	964	659.6	140.0	482	428	1297	768.3	169.6
	HSCW (kg)	74	226	438	312	57	1202	121	400	275.0	54.1	482	146	416	286.2	50.9
	CP8 (mm)	74	4	26	12.1	5.6	1202	1	34	9.4	4.6	482	1	27	11.1	4.8
	OSS	74	100	150	133.6	10.8	1202	100	200	143.4	20.5	482	100	200	148.9	19.8

were collected within the CRC for Cattle and Beef Quality (Robinson, 1995; Upton et al., 2001). Data was collected between 1997 and 1999, and after removing animals ($n = 19$) with missing measurements or ossification scores outside of the range (OSS: between 100 and 200) contained in the development dataset, this dataset contained 83 heifers and 1,202 steers. Calves were purchased at weaning by the Beef CRC from 34 commercial breeder herds throughout eastern Australia and half were grown out on pasture and half in a feedlot (both occurred in North Eastern New South Wales). Grain finished cattle were targeted for feedlot entry weights of 300, 400 (short fed), and 400 kg (long fed) for 400, 520, and 640 kg slaughter BWs, respectively. Live animal traits were recorded at intervals of 6 mo and animals were slaughtered when the average BW of the cohort was predicted to achieve target carcass weights of 220 kg (domestic market), 280 kg (Korean market) or 340 kg (Japanese market) (Upton et al., 2001). For more detail on the recording of slaughter traits, see Perry et al. (2001).

Dataset 3. Data from *B. indicus* breed types (Brahman and Santa Gertrudis) was obtained from the Beef CRC as described for dataset 2. These calves were bred between 1995 and 1998 on 12 Queensland commercial properties and, after removing animals ($n = 73$) with missing measurements or ossification scores outside of the range (OSS: between 100 and 200) contained in the development dataset, this

dataset contained 404 heifers and 482 steers. At weaning, one-third of the calves were transferred to join the animals in dataset 2 in Northern New South Wales and the remainder was kept in Central Queensland for either pasture or feedlot finishing (Upton et al., 2001). These animals were slaughtered between 1997 and 2000. The same traits and measurement procedures were used for these animals as described for dataset 2.

Statistical Analysis

Linear regression analyses were conducted using the linear model procedure in the R statistical package (R Development Core Team, 2014) to develop prediction models with ossification as the dependent variable (Eq. [1]), as follows:

$$\text{Ossification} = \beta_0 + \beta_1 \text{age} / \text{BW} / \text{HSCW} + \beta_2 \text{sex} + \dots + \text{interactions} + e, \quad (1)$$

where combinations of traits were fitted as covariates and fixed effects, with β_i representing the regression coefficient of each fitted covariate/fixed effect; and e the residual error. All 2-way interactions were evaluated. Sex was coded as heifer = 0 and steer = 1. The linear regressions were progressively developed, beginning with models that only contained age (AGEREC or AGESL), BW, or HSCW as covariates. The effect of sex was then included in association with age, BW, and HSCW. Models

that included either all live animal (USRIB, USP8, USEMA, muscle score and either BW or ADG) or carcass (CRIB, CP8, marbling score, CEMA and either HSCW or ADG) traits plus all 2-way interactions were also evaluated. Only one measure of growth was used in each of the models; thus, only BW or ADG were included in each live trait model and only HSCW or ADG in the carcass trait models. Effects that were not significant ($P < 0.05$) were removed. The Box-Cox procedure (Box and Cox, 1964) in R was used to determine that the ossification data did not require transformation. The adjusted R^2 (R^2_{adj}) and residual standard error (SE) were used to assess how well models fitted the development data. The R^2_{adj} (Eq. [2]) was used to adjust for the different number of parameters fitted in each model during model development:

$$R^2_{\text{adj}} = 1 - \left[\frac{SSE(n-1)}{SST(n-p-1)} \right], \quad (2)$$

where SST is the total sums of squares, SSE is the error sums of squares, n is the number of observations, and p is the number of parameters in the model.

The prediction equations with R^2_{adj} above 0.55 were challenged using the three evaluation datasets. The R statistical package (R Development Core Team, 2014) was used to plot observed versus predicted ossification with a line of unity ($y = x$) and the residuals (observed – predicted) *versus* predicted. The observed *versus* predicted plots with the line of unity do not have regression lines because the accuracy of prediction is how close a dataset can lie on the line of unity rather than the strength of an R^2 value (Tedeschi 2006). Model evaluation was conducted using a customized procedure in the R statistical package. Model predictions of ossification were evaluated using mean bias (MB; Eq. [3]):

$$\frac{\sum (O_i - P_i)}{n}, \quad (3)$$

where n is the number of data points, O_i is the observed ossification, and P_i is the predicted ossification, respectively ($i = 1$ to n). The error of prediction was assessed by the mean square error of prediction (MSEP; Eq. [4]):

$$\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}, \quad (4)$$

where the terms are as defined above. The root mean square error of prediction (RMSEP) was used as a measure of the accuracy of prediction. The MSEP was decomposed into bias, slope, and

random components as a proportion of MSEP to assess the error structure, following the method of Bibby and Toutenburg (1977). The statistical significance of each MB was evaluated using a paired t -test of the mean of the differences between the observed and model-predicted values.

RESULTS

Model Development

Comparison of models where either BW or ADG were included for the live animal traits and either HSCW or ADG were included for the carcass traits found in all instances that ADG was not significant in explaining variation in ossification and these models had lower fits to the ossification data than models containing either BW or HSCW. Consequently, only the results for models that contained BW or HSCW are presented.

The coefficients for the linear regressions of ossification on either live animal (sex, AGEREC, USRIB, USP8, USEMA, muscle score, grass or grain finishing, and BW) or carcass (sex, AGESL, CRIB, CP8, marbling score, CEMA, and HSCW) traits, including all 2-way interactions, are shown in Table 3. The linear regression between ossification and AGEREC (model A) demonstrates that AGEREC accounted for only a moderate proportion of the variation in ossification ($R^2_{\text{adj}} = 0.30$). Models C and E that contained either BW or HSCW only accounted for a very low proportion of the variation in ossification (both $R^2_{\text{adj}} = 0.08$). Models B, D, and F demonstrate that including sex and its interaction with other traits accounted for more of the observed variation in ossification ($P < 0.05$). Models G and H demonstrate that CP8 fitted in conjunction with sex or sex and HSCW also accounted for more of the observed variation in ossification ($P < 0.05$). Other live and carcass traits evaluated did not significantly ($P > 0.05$) account for any of the variation observed in ossification. An unexpected outcome was that USP8 was not significant ($P > 0.10$) when tested during the development of model D, but CP8 was significant in both models G and H.

The R^2_{adj} and SE shown in Table 3 indicate models A, C, and E fit the ossification data less well than the other models. Model B, which contains sex, AGEREC, and their interaction, has the highest R^2_{adj} and lowest SE ($R^2_{\text{adj}} = 0.66$, SE = 12.38). Models D, F, G, and H, which do not include AGEREC, but include BW, CP8, and/or HSCW, have slightly

lower R^2_{adj} and slightly higher SE ($R^2_{adj} = 0.59, 0.58, 0.60$, and 0.60 , SE = 13.51, 13.70, 13.35, and 13.39, respectively; Table 3) than model B.

Predictions of ossification are illustrated in Figs. 1 to 3 for models B, D (live inputs), F, G, and H (carcass inputs). The effect AGEREC, BW, HSCW, and CP8 have on the prediction of ossification is clearly demonstrated. Comparison of Fig. 2b and Fig. 3a illustrate the different influences CP8 has on the prediction of ossification, with divergent responses to CP8 in steers and heifers predicted by model G and parallel responses predicted by model H.

Model Evaluation

The 5 models with R^2_{adj} above 0.55 in the development analyses were evaluated using the three evaluation datasets with the results reported in Table 4. It is important to bear in mind when examining these evaluation results that ossification is assessed in increments of 10 (Wolcott et al. 2009) as opposed to on a continuous scale. The MB in Table 4 for dataset 1 indicates that all models

predicted ossification accurately with model B having the largest bias (-4.2). Model D slightly under predicted ossification while models B, F, G, and H tended to slightly over predict. The MBs found in dataset 2 were higher than those found in dataset 1 except model B which had an MB of 2.8. Models D and G had the highest MB while F had the lowest, after model B. The MBs found in dataset 3 were generally higher than those found in datasets 1 and 2 with model D having the highest MB and model F the lowest. In general, all models tended to under predict ossification in datasets 2 and 3, except model B in dataset 3, which over predicted ossification by a similar magnitude to the under prediction made by the other models.

The RMSEP followed a similar trend to that observed for MB across the three evaluation datasets (Table 4). The average RMSEP was lowest in dataset 1, increased in dataset 2 and increased again in dataset 3. The RMSEP were similar in dataset 1 for all models. In dataset 2 model B had the lowest RMSEP (19.6), while it had the highest (31.9) in dataset 3. The remaining models had similar RMSEP in both datasets 2 and 3. The

Table 3. Linear regression coefficients (SE) for the prediction of ossification based on either live animal or carcass traits; adjusted R^2 (R^2_{adj}) and residual standard error (SE) of model fit are presented

Model	Equation ¹	R^2_{adj}	SE
<i>Live Animal</i>			
A	$54.28_{(13.33)} + 0.16_{(0.02)} \times \text{AGEREC}$	0.30	17.66
B	$27.83_{(15.49)} + 41.03_{(19.41)} \times \text{sex} + 0.23_{(0.03)} \times \text{AGEREC} - 0.12_{(0.04)} \times \text{sex} \times \text{AGEREC}$	0.66	12.38
C	$111.79_{(9.09)} + 0.06_{(0.02)} \times \text{BW}$	0.08	20.23
D	$85.35_{(10.16)} + 0.16_{(0.02)} \times \text{BW} + 10.94_{(13.01)} \times \text{sex} - 0.09_{(0.03)} \times \text{sex} \times \text{BW}$	0.59	13.51
<i>Carcass traits</i>			
E	$115.10_{(8.18)} + 0.10_{(0.03)} \times \text{HSCW}$	0.08	20.27
F	$92.77_{(9.08)} + 0.28_{(0.04)} \times \text{HSCW} + 10.85_{(11.70)} \times \text{sex} - 0.18_{(0.05)} \times \text{sex} \times \text{HSCW}$	0.58	13.70
G	$126.53_{(4.29)} - 6.06_{(5.15)} \times \text{sex} + 2.35_{(0.32)} \times \text{CP8} - 1.25_{(0.43)} \times \text{sex} \times \text{CP8}$	0.60	13.35
H	$107.15_{(10.92)} + 11.53_{(11.45)} \times \text{sex} + 1.10_{(0.49)} \times \text{CP8} + 0.16_{(0.07)} \times \text{HSCW} - 0.15_{(0.05)} \times \text{sex} \times \text{HSCW}$	0.60	13.39

¹Sex (heifer = 0, steer = 1), AGEREC = age in d, BW = animal body weight (kg), HSCW = hot standard carcass weight (kg), CP8 = subcutaneous carcass rump fat thickness (mm).

Figure 1. Predictions of ossification using (a) model B with age at scanning ranging from 436 to 649 d for heifers and 422 to 650 d for steers and using (b) model D with live weight ranging from 309 to 640 kg for heifers and 327 to 740 kg for steers.

Figure 2. Predictions of ossification using (a) model F with hot standard carcass weight ranging from 155 to 361 kg for heifers and 170 to 404 kg for steers and using (b) model G with carcass P8 fat (CP8) ranging from 1 to 27 mm in heifers and 1 to 26 mm in steers.

Figure 3. Predictions of ossification using model H with (a) hot standard carcass weight held constant at 253 kg while carcass P8 fat (CP8) ranged from 1 to 27 mm in heifers and from 1 to 26 mm in steers and with (b) carcass P8 fat (CP8) constant at 9 mm while hot standard carcass weight ranged from 155 to 361 kg in heifers and from 170 to 404 kg in steers.

decomposition of the MSEP into its components indicated that the majority (>50%) of error was of a random nature, except model B in dataset 3 where the random component only accounted for ~25% of the error. The proportion of random error on average decreased from dataset 1 to dataset 2, and finally to dataset 3. In dataset 1, model B had some error contained in the bias component while the remaining models had error contained in the slope component. In datasets 2 and 3, the opposite occurred, where model B had more error contained in the slope component and the remaining models had more error in the bias component. The predictive ability of model H in each dataset across the range of ossification scores present in the development dataset is shown in Fig. 4.

DISCUSSION

The key to remaining profitable and producing meat sustainably in any production system is to produce the product the consumer demands consistently and cost effectively (Egan et al., 2001; McPhee et al., 2006). Phenotypic prediction systems (e.g., BeefSpecs calculator; Walmsley et al.,

2014), have been developed to provide cattle producers with information that would otherwise not be accessible to help inform management decisions that impact carcass compliance and meat quality. The increasing focus on the consumer eating experience necessitates that these prediction systems broaden their capabilities to include not only compliance traits like HSCW and CP8 but meat quality traits.

Tenderness is well established as being critical in determining meat eating quality and acceptability (Voges et al., 2007; Destefanis et al., 2008) with consumer preference studies of sensory attributes showing tenderness is more important than flavor or juiciness (Tornberg 1996; Destefanis et al., 2008). This is further demonstrated by the positive relationship between the price of a cut of meat and its tenderness (Miller et al., 2001). The negative relationship between tenderness and animal maturity (Romans et al., 1965; Shorthose and Harris 1990; Schonfeldt and Strydom 2011) due to muscle tissue containing decreasing amounts of heat-labile collagen as animals mature (Schonfeldt and Strydom 2011) has been seized upon by meat grading systems (e.g., MSA, USDA) as a valuable indicator of eating quality and consumer acceptability.

Table 4. Evaluation of models B and D, that contain live traits, and models F, G, and H, that contain carcass traits, for their accuracy of predicting ossification

Dataset:	Dataset 1					Dataset 2					Dataset 3				
	Model					Model					Model				
	Live Traits		Carcass Traits			Live Traits		Carcass Traits			Live Traits		Carcass Traits		
Item	B	D	F	G	H	B	D	F	G	H	B	D	F	G	H
Mean observed	133.6	133.6	133.6	133.6	133.6	145.1	145.1	145.1	145.1	145.1	159.0	159.0	159.0	159.0	159.0
Mean predicted	137.8	131.9	135.2	133.7	134.1	142.3	131.7	133.3	132.0	132.5	175.3	141.5	144.6	143.0	144.1
Mean bias	-4.2	1.7	-1.6	-0.1	-0.5	2.8	13.4	11.8	13.1	12.6	-16.3	17.5	14.4	16.0	14.9
P-value	<0.01	0.130	0.175	0.984	0.700	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
MSEP ¹															
RMSEP	10.5	10.0	9.8	10.7	10.7	19.6	23.0	22.1	23.7	23.2	31.9	25.7	23.3	24.6	23.4
Bias, %	16.0	3.1	2.5	0.0	0.2	2.1	34.2	28.9	30.9	29.6	26.1	46.2	38.0	42.2	40.5
Slope, %	1.0	1.5	1.0	7.8	8.5	16.0	0.0	0.1	0.1	0.0	48.4	2.7	2.4	1.5	1.1
Random, %	83.0	95.4	96.5	92.2	91.3	81.9	65.8	71.0	69.0	70.4	25.5	51.1	59.6	56.3	58.4

¹MSEP = mean square prediction error; RMSEP = root MSEP; Bias = MSEP decomposed into error attributable to overall bias of prediction; Slope = MSEP decomposed into error attributable to deviation of the regression slope from unity; Random = MSEP decomposed into error attributable to the random variation.

Figure 4. Evaluation of model H for datasets 1, 2, and 3 showing plots of (a) observed versus predicted and (b) the residual (observed – predicted) for ossification.

Previous research has used different measures as indicators of animal maturity including chronological age (Shackelford et al., 1995; Raines et al., 2008) and ossification (Romans et al., 1965; Shackelford et al., 1995; Raines et al., 2008). Although these measures have been shown to be positively correlated (0.64, Raines et al., 2008), they are not synonymous (Romans et al., 1965). Age is a linear measure that is not affected by physiological processes such as pregnancy, lactation, or age at maturity, whereas ossification is strongly influenced by hormonal status of the animal (Field et al., 1997; Scheffler et al., 2003). This close association with physiological processes allows ossification to better reflect physiological maturity (Field et al., 1997), and thus tenderness (Weston et al., 2002) than age and is why the MSA and USDA grading systems

both use ossification as the measure for maturity when predicting eating quality.

Model Development

The linear regression between ossification and AGEREC (model A) supports that they are not synonymous in that AGEREC accounted for only a moderate proportion of the variation present in ossification ($R^2_{adj} = 0.30$, Table 3). Despite AGEREC explaining less than anticipated it explained a significantly larger proportion of the variation in ossification than either BW (model C) or HSCW alone (model E; both $R^2_{adj} = 0.08$, Table 3). Although looking at age, Raines et al., (2008) found a correlation with hot carcass weight of 0.12 which contrasts to the correlation with

ossification of 0.64 which suggests that weight in isolation is not a good indicator of age or ossification. Regressing ossification on ADG (results not presented) provided no advantage over regressing on BW or HSCW which was again unexpected given animals of a similar type with higher ADG and thus younger ages at a particular weight would be expected to be less physiologically mature. This result is supported by Field et al. (1997) who found that 2 similar groups of heifers had significantly different ($P < 0.01$) ossification, as shown by bone maturity, after similar rates of weight gain.

The rate of development of ossification and hence the ossification score of a carcass is strongly influenced by hormonal status of the animal, and in particular oestrogen (Field et al., 1997; Scheffler et al., 2003). Oestrogen levels at any point of development of the animal are influenced by sex, castration, hormonal growth promotants, and parity status (Waggoner et al., 1990; Field et al., 1996; Scheffler et al., 2003). The variation in ossification explained when sex was included with AGEREC (model B) doubled compared to AGEREC alone (model A) whereas when sex was included with either BW (model D) or HSCW (model F) the variation in ossification explained increased by more than 7 times (Table 3) compared to either BW (model C) or HSCW (model E) alone. When sex and CP8 were included, slightly more variation ($R^2_{\text{adj}} = 0.60$) was explained than when sex was included with either BW ($R^2_{\text{adj}} = 0.59$) or HSCW ($R^2_{\text{adj}} = 0.58$). The pattern of accretion of body components where fat accumulation takes higher priority than bone or lean tissue as animal's mature (Berg and Butterfield 1976) provides a possible explanation as to why CP8 when used in association with sex explained slightly more variation in ossification than either BW or HSCW. The inclusion of HSCW with sex and CP8 explained the same amount of variation in ossification ($R^2_{\text{adj}} = 0.60$) as sex and CP8 fat which could be interpreted as weight may have a limited capacity to describe physiological maturity. However, this was not observed during the development of model D where USP8 was not significant ($P > 0.10$) in explaining the variation in ossification while BW was significant ($P < 0.05$).

Model Evaluation

The purpose of this research was to develop a prediction equation that uses either live animal or derived carcass traits to predict ossification with sufficient accuracy for phenotypic prediction systems to make robust predictions of meat quality. Model

evaluation revealed that, in general, all the models tested had similar accuracies within a dataset and similar differences in accuracies between datasets. The only exception was model B which had lower absolute MB and RMSEP when tested in dataset 2 than the other models and a higher RMSEP than the other models in dataset 3. In reality, these differences are not as large as they initially appear due to the scale (increments of 10) used for assessing ossification (Wolcott et al. 2009). The RMSEP for all models in all evaluation datasets was within, or close to, 3 increments of the ossification scale while the MB was within 2 increments. In particular, the MB in dataset 1 was essentially zero and the RMSEP was 1 when the ossification scale is taken into consideration. These results indicate all models have predictive accuracy sufficient for inclusion in phenotypic prediction systems.

When examining the predictive accuracy of models between datasets all models generally had higher predictive accuracy in dataset 1 than dataset 2 and lower accuracy again in dataset 3. The loss in accuracy for all models except model B when tested in datasets 2 and 3 compared to dataset 1 is attributable to the bias component of the RMSEP. This result suggests that there is a factor(s) impacting on ossification that differs between dataset 1 and datasets 2 and 3 which are not being accounted for by any of the models. This could be explained by dataset 1 being an earlier cohort of the same herd from which the development dataset was obtained whereas datasets 2 and 3 are regarded as being unrelated to the development dataset. Model B had more error in the slope component than the bias component of RMSEP which suggests that there is a different relationship between ossification and AGEREC in datasets 2 and 3 than the development dataset or dataset 1. This again could be explained by dataset 1 being recorded in the 12 mo prior to the development dataset with the 2 datasets experiencing a similar production environment and similar genetic backgrounds. In contrast, datasets 2 and 3 were recorded approximately 15 yr earlier and both the genetics of the animals and production systems have changed.

The evolution of the method used to record ossification score in Australia could also provide some explanation of the evaluation differences seen between dataset 1 and datasets 2 and 3. The ossification data in datasets 2 and 3 was actually recorded before ossification was adopted by MSA (Perry et al., 2001) for describing the impact of animal maturity on meat quality. It is feasible that developments which have occurred during the development

of MSA and the training of the MSA technicians could have increased the quality of trait recording found in dataset 1 and the development dataset compared to datasets 2 and 3. This may also explain why unpublished results described below (using datasets 2 and 3 to develop models) produced less accurate results than those reported in this study.

The decrease in accuracy found between datasets 2 and 3 is explained by the differences in breed composition between the development dataset and the evaluation datasets. Dataset 1 was an earlier cohort of the same herd from which the development dataset was obtained; dataset 2 was unrelated but comprised, predominantly, of the same breeds as found in the development dataset, while dataset 3 was comprised of *B. indicus* content animals. The correlation between the accuracy with which the models predicted ossification and the relatedness of the validation datasets in terms of breed content suggests that the prediction model(s) may require a breed factor to be included. The absence of variation in breed in the development dataset prevented this from occurring in this study.

In an attempt to overcome this limitation, a combination of datasets 2 and 3 were used to develop prediction models that included a breed factor but this proved to be unsuccessful (unpublished). The prediction models obtained when doing this had higher MB and a greater proportion of the MSEP due to slope and bias when evaluated using dataset 1 and the development dataset. These MB and MSEP values were all larger than those values presented in Table 4 for the models in the current study. An alternative would be to include an ad hoc *B. indicus* breed correction, based on the MB found in dataset 3 evaluations, in the models to increase their transferability to other breeds. However, this would require a second *B. indicus* dataset to evaluate this approach, which was not available in this study. It would be highly desirable to include variation in breed content in any future research focused on ossification.

Model Selection for Phenotypic Prediction

The similar predictive accuracy observed for all models evaluated across the 3 evaluation datasets (Table 4) when considering ossification is recorded in 10 unit increments (Wolcott et al., 2009) indicates that any of the models could be used to make phenotypic predictions of ossification. The context in which the equations are used will ultimately determine their usefulness. Given very few cattle in commercial production systems have birth dates

recorded and thus age known, model B is anticipated to be of limited value for predicting ossification. It is desirable to minimize the number of input parameters to those necessary for accurate prediction, however models D, F, G, and H contain only 2 or 3 input parameters and one of these is sex. If ossification was to be predicted independent of any other modelling system, model D would probably be the most suitable given the inputs (BW and sex) are generally accurately recorded. The potential weakness of model D is that 2 animals of equal BW, but different mature sizes would be predicted to have the same ossification even though the animal with a lower mature size would be expected to have a higher physiological maturity (Berg and Butterfield 1976), and thus ossification (Romans et al., 1965; Shackelford et al., 1995; Raines et al., 2008) at any weight.

When fat deposition increases in priority relative to other tissues as animal's mature, animals with lower mature size would be expected to be fatter at any given BW/HSCW which is captured in model H. For both sexes at any HSCW, model H would predict ossification to be higher as CP8 fat increases (Fig. 3). This is not represented in model G, in fact model G suffers a similar weakness to model D, in that, animals which grew at 2 different rates to the same final CP8 fat depth would be predicted to have the same ossification even though the animal which grew at the slower rate would be expected to have a higher physiological maturity, and thus ossification. The weakness of model H is HSCW and even CP8 fat depth are not simple measures to record accurately on live animals. However, when used in association with a phenotypic prediction system like the BeefSpecs calculator (Walmsley et al., 2014), which predicts carcass P8 rump fat and HSCW, model H could produce predictions of ossification that align with physiological expectations.

The accuracy of the predictions of ossification would be determined by the accuracy of the predictions of P8 rump fat and HSCW. Only when HSCW is predicted with higher accuracy than the accuracy of measuring animal live weight would the accuracy of models E, F, and H be expected to be higher than model D when assessing animals ready for slaughter. When predicting into the future, as is the intent of the BeefSpecs calculator, predictions of ossification made with model D would be limited by the accuracy of predicting future live weight. Using any of the models developed and evaluated in this study in the appropriate context to inform predictors of meat quality will assist beef producers focus more of their decision making on satisfying consumer

demands, as described by carcass grading systems, consistently and cost effectively to remain profitable. The collection of data that includes variation in breed is desirable to further develop the models in this study to increase their accuracy of prediction.

LITERATURE CITED

- AUS-MEAT. 2005. Handbook of Australian meat. 7th ed. AUS-MEAT Ltd, Brisbane, QLD.
- Barwick, S. A., M. L. Wolcott, D. J. Johnston, H. M. Burrow, and M. T. Sullivan. 2009. Genetics of steer daily and residual feed intake in two tropical beef genotypes, and relationships among intake, body composition, growth and other post-weaning measures. *Anim. Prod. Sci.* 49:351–366. doi:10.1071/EA08249
- Berg, R. T., and R. M. Butterfield. 1976. New concepts of cattle growth. Sydney University press, Sydney, NSW.
- Bibby J., and H. Toutenburg. 1977. Prediction and improved estimation in linear models. John Wiley & Sons and Akademie-Verlag, Berlin, German Democratic Republic.
- Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations (with discussion). *J. R. Stat. Soc. B. Stat. Methodol.* 26:211–252. <http://www.jstor.org/stable/2984418>.
- Cafe, L. M., W. A. McKiernan, and D. L. Robinson. 2018. Selection for increased muscling is not detrimental to maternal productivity traits in Angus cows. *Anim. Prod. Sci.* 58:185–192. doi:10.1071/AN13023
- Destefanis, G., A. Brugiapaglia, M. T. Barge, and E. Dal Molin. 2008. Relationship between beef consumer tenderness perception and Warner-Bratzler shear force. *Meat Sci.* 78:153–156. doi:10.1016/j.meatsci.2007.05.031
- Egan, A. F., D. M., Ferguson, and J. M. Thompson. 2001. Consumer sensory requirements for beef and their implications for the Australian beef industry. *Aust. J. Exp. Agric.* 41:855–859. doi:10.1071/EA00065
- Field, R., R. McCormick, V. Balasubramanian, D. Sanson, J. Wise, D. Hixon, M. Riley, and W. Russell. 1996. Growth, carcass, and tenderness characteristics of virgin, spayed, and single-calf heifers. *J. Anim. Sci.* 74:2178–2186. doi:10.2527/1996.7492178x
- Field, R., R. McCormick, V. Balasubramanian, D. Sanson, J. Wise, D. Hixon, M. Riley, and W. Russell. 1997. Tenderness variation among loin steaks from A and C maturity carcasses of heifers similar in chronological age. *J. Anim. Sci.* 75:693–699. doi:10.2527/1997.753693x
- Johnston, D. J., A. Reverter, H. M. Burrow, V. H. Oddy, and D. L. Robinson. 2003. Genetic and phenotypic characterisation of animal, carcass, and meat quality traits from temperate and tropically adapted beef breeds. 1. Animal measures. *Aust. J. Agric. Res.* 54:107–118. doi:10.1071/AR02085
- McKiernan, W. A. 2007. Muscle scoring beef cattle. Primefact No. 328. NSW Department of Primary Industries, Orange, NSW.
- McPhee, M. J., J. W. Oltjen, T. R. Famula, and R. D. Sainz. 2006. Meta-analysis of factors affecting carcass characteristics of feedlot steers. *J. Anim. Sci.* 84:3143–3154. doi:10.2527/jas.2006-175
- Miller, M. F., M. A. Carr, C. B. Ramsey, K. L. Crockett, and L. C. Hoover. 2001. Consumer thresholds for establishing the value of beef tenderness. *J. Anim. Sci.* 79:3062–3068. doi:10.2527/2001.79123062x
- Perry, D., W. R. Shorthose, D. M. Ferguson, and J. M. Thompson. 2001. Methods used in the CRC program for the determination of carcass yield and beef quality. *Aust. J. Exp. Agric.* 41:953–957. doi:10.1071/EA00092
- R Development Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Raines, C. R., M. E. Dikeman, J. A. Unruh, M. C. Hunt, and R. C. Knock. 2008. Predicting cattle age from eye lens weight and nitrogen content, dentition, and United States department of agriculture maturity score. *J. Anim. Sci.* 86:3557–3567. doi:10.2527/jas.2007-0445
- Robinson, D. L. 1995. Design of the CRC straightbred genetics experiments. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 11:541–545.
- Romans, J. R., H. J. Tuma, and W. L. Tucker. 1965. Influence of carcass maturity and marbling on the physical and chemical characteristics of beef. I. Palatability, fiber diameter and proximate analysis. *J. Anim. Sci.* 24:681–685. doi:10.2527/jas1965.243681x
- Scheffler, J. M., D. D. Buskirk, S. R. Rust, J. D. Cowley, and M. E. Doumit. 2003. Effect of repeated administration of combination trenbolone acetate and estradiol implants on growth, carcass traits, and beef quality of long-fed holstein steers. *J. Anim. Sci.* 81:2395–2400. doi:10.2527/2003.81102395x
- Schönfeldt, H. C., and P. E. Strydom. 2011. Effect of age and cut on tenderness of South African beef. *Meat Sci.* 87:206–218. doi:10.1016/j.meatsci.2010.10.011
- Shackelford, S. D., M. Koohmaraie, and T. L. Wheeler. 1995. Effects of slaughter age on meat tenderness and USDA carcass maturity scores of beef females. *J. Anim. Sci.* 73:3304–3309. doi:10.2527/1995.73113304x
- Shorthose, W. R., and P. V. Harris. 1990. Effect of animal age on the tenderness of selected beef muscles. *J. Food Sci.* 55:1–8. doi:10.1111/j.1365-2621.1990.tb06004.x
- Tedeschi, L. O. 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89:225–247. doi:10.1016/j.agry.2005.11.004
- Tornberg, E. 1996. Biophysical aspects of meat tenderness. *Meat Sci.* 43:S175–S191. doi:10.1016/0309-1740(96)00064-2
- Upton, W. H., H. M. Burrow, A. Dundon, D. L. Robinson, and E. B. Farrell. 2001. CRC breeding program design, measurements and database: methods that underpin CRC research results. *Aust. J. Exp. Agric.* 41:943–952. doi:10.1071/EA00064
- Upton, W. H., K. A. Donoghue, H.-U. Graser, and D. J. Johnston. 1999. Ultrasound proficiency testing. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 13:341–344.
- USDA. 1997. United States standards for grades of slaughter cattle. Agricultural Marketing Service, USDA, Washington, DC.
- Voges, K. L., C. L. Mason, J. C. Brooks, R. J. Delmore, D. B. Griffin, D. S. Hale, W. R. Henning, D. D. Johnson, C. L. Lorenzen, R. J. Maddock, et al. 2007. National beef tenderness survey - 2006: assessment of Warner-Bratzler shear and sensory panel ratings for beef from US retail and foodservice establishments. *Meat Sci.* 77:357–364. doi:10.1016/j.meatsci.2007.03.024
- Waggoner, A. W., M. E. Dikeman, J. R. Brethour, and K. E. Kemp. 1990. Performance, carcass, cartilage calcium, sensory and collagen traits of longissimus muscles of open versus 30-month-old heifers that produced one calf. *J. Anim. Sci.* 68:2380–2386. doi:10.2527/1990.6882380x

- Walmsley, B. J., and W. A. McKiernan. 2011. Observed trends in estimated breeding values in response to selection using visual muscle score in beef cattle. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 19:123–126.
- Walmsley, B. J., M. J. McPhee, and V. H. Oddy. 2014. Development of the BeefSpecs fat calculator to assist decision making to increase compliance rates with beef carcass specifications. *Anim. Prod. Sci.* 54:2003–2010. doi:10.1071/AN14611
- Watson, R., R. Polkinghorne, and J. M. Thompson. 2008. Development of the Meat Standards Australia (MSA) prediction model for beef palatability. *Aust. J. Exp. Ag.* 48:1368–1379. doi:10.1071/EA07184
- Weston, A. R., R. W. Rogers, and T. G. Althen. 2002. Review: the role of collagen in meat tenderness. *Prof. Anim. Sci.* 18:107–111. doi:10.15232/S1080-7446(15)31497-2
- Wolcott, M. L., D. J. Johnston, S. A. Barwick, C. L. Iker, J. M. Thompson, and H. M. Burrow. 2009. Genetics of meat quality and carcass traits and the impact of tenderstretching in two tropical beef genotypes. *Anim. Prod. Sci.* 49:383–398. doi:10.1071/EA08275
- Wolcott, M. L., J. M. Thompson, and D. Perry. 2001. The prediction of retail beef yield from real time ultrasound measurements on live animals at three stages through growout and finishing. *Aust. J. Exp. Agric.* 41:1005–1011. doi:10.1071/EA00017