final report

Project code: L.EQT.1628

Prepared by: IJ Lean#, N Lees#, HM Golder#, P McGilchrist* and JEP Santos‡
#Strategic Bovine Services Pty Ltd (Scibus), *School of Environmental and Rural Science, University of New England ‡ Department of Animal Sciences, University of Florida

Date published: 12 September 2017

PUBLISHED BY
Meat and Livestock Australia Limited
PO Box 1961
NORTH SYDNEY NSW 2059

Meta-analysis and Literature Review on the Ageing and Hormonal Growth Promotant Interventions in the MSA Model

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

This publication is published by Meat & Livestock Australia Limited ABN 39 081 678 364 (MLA). Care is taken to ensure the accuracy of the information contained in this publication. However MLA cannot accept responsibility for the accuracy or completeness of the information or opinions contained in the publication. You should make your own enquiries before making decisions concerning your interests. Reproduction in whole or in part of this publication is prohibited without prior written consent of MLA.

Executive summary

There are benefits of hormonal growth promotants (HGP) for production efficiency, profit, and environmental effects of beef cattle. Questions remain; however, about the effects of HGP on meat quality, particularly on measures of toughness such as Warner-Bratzler shear force (WBSF), tenderness, and other consumer tested attributes of meat such as juiciness, tenderness, flavour, and connective tissue. This meta-analysis used 31 experiments containing 181 treatment comparisons to evaluate the effects of HGP on WBSF and sensory measures on meat quality, almost all using Longissimus dorsi. The experiments varied greatly in design in using many different hormonal treatments and combinations, which were single or repeated, in different breeds and sex groups of cattle, with or without electrical stimulation, and with different lengths of meat ageing and feeding. The effects of multiple treatment comparisons in experiments were evaluated using robust regression models and compared to Knapp-Hartung and permutation meta-analytical methods. In general, the true variance of experiments, tau^2 (τ^2) was low <0.1, but heterogeneity, I^2 was high >50% indicating that much of the variance was due to measurement error. Increased WBSF was associated with HGP treatment; in particular, use of multiple HGP implants was associated with an increase in WBSF of 0.248 kg (95% CI 0.203 to 0.292), but a single implant had more limited effects with an increase in WBSF of 0.176 kg (95% CI 0.109 to 0.242). Ageing did not significantly alter the HGP association with increased WBSF (P = 0.105); however, the point direction was towards a reduced effect with ageing (ES = −0.005 per d aged). Studies using trenbolone acetate treatments did not differ in WBSF from those using other implants (P > 0.15). The experiment also provides information on other sensory aspects of meat quality. The findings on tenderness, as assessed by sensory methods differ, from those using WBSF as HGP treatment was not associated with reduced tenderness (P > 0.3) and multiple treatments increased tenderness (ES = 0.468) compared to a single implant. Further, juiciness, flavour, and connective tissue were not associated with HGP use; whereas, there was a marked 5.5-point decrease in meat quality score (Meat Standards Australia quality scoring system, CMQ4), albeit with limited experiments and treatments. There is a need for more targeted studies on the role of HGP in influencing meat quality to examine the effects of different HGP treatments and ageing on WBSF, tenderness, juiciness, and other sensory measures.

Table of contents

1 Background

Hormonal growth promotant implants (HGP) are widely used in the beef industries of major beef producing countries including USA, Brazil, Argentina, Australia, and South Africa. The impacts of these on the efficiency of meat production are substantial with many individual reports and reviews highlighting productive responses including increased weight gain and feed efficiency from the interventions. There are also substantial environmental benefits (Capper and Hayes 2012) from the use of these interventions and the production responses are profitable for producers (Hunter 2010). However, questions remain about the effect of HGP on meat quality, particularly on measures of toughness such as Warner-Bratzler shear force (WBSF), tenderness, and other consumer tested attributes of meat for instance juiciness, tenderness, flavour, and connective tissue (Watson 2008).

There have been a number of quantitative and semi-quantitative reviews of the effects of HGP on the quality of meat as assessed by WBSF. In a traditional review, there was evidence of increased toughness of the meat with HGP use that the authors chose to consider to be negligible Nichols *et al.* (2002). In a semi-quantitative review, Duckett and Pratt (2014) considered that the impacts of the increase in WBSF may be more associated with repeated treatments with HGP and with anabolic rather than oestrogenic steroids. Duckett and Pratt (2014) noted the quantitative review and meta-analysis by Watson (2008) of the effects of HGP in increasing WBSF and toughness and considered that there may be mitigating factors such as repeated number of implants and potential for post-mortem ageing to influence the responses. The aim of this meta-analysis was to evaluate the effects of HGP, primarily on WBSF, but also to consider effects on other meat quality outcomes and to evaluate whether sources of variation in the responses to HGP may be mediated by factors such as ageing, type of implant, number of implants used, and freezing.

2 Project objectives

Phase 1 – conduct a literature review to determine the availability of information
Phase 2 – conduct a meta-analysis, provided there is sufficient existing information
Phase 3 – extensive further analysis of identified literature

3 Methodology

3.1 Literature search and diet description

A comprehensive literature search of English language literature published from 1975 to 2017 was conducted to identify research experiments involving treatment comparisons designed to evaluate the effects of HGP on meat quality, primarily on the change in WBSF and taste-panel data for the juiciness, flavour, tenderness, connective tissue content, and Meat Standards Australia CMQ4 score. Three search engines, ISI Web of Science (http://wokinfo.com/), Google Scholar (http://scholar.google.com/), and Pub Med (http://www.ncbi.nlm.nih.gov/pubmed), were utilized between May 1 and 14, 2017 with a defined and repeatable search strategy using the terms "(HGP OR hormonal OR implants) AND (palatability OR shear-force OR tenderness) AND (beef or steer)" to identify relevant experiments. Additional experiments were examined from the references of experiments identified from the primary databases searches.

3.2 Inclusion and exclusion criteria

All published experiments were screened using standardized criteria. For inclusion into the meta-analysis, experiments needed to have the following: be English language, use HGP in a randomized, replicated experiment in which a reference group was present, they measured meat quality outcomes,

they included sufficient data to determine the effect size (ES), they included a measure of effect amenable to ES analysis for continuous data (e.g., standardized mean difference [SMD]), and they included a measure of variance (SE or SD) for each effect estimate or treatment and control comparisons. In order to reduce variability in the evaluation and ensure that multiple comparisons on the one carcass were not included, longissimus muscle was assessed and data from other muscle groups was excluded with a single exception.

3.3 Data extraction

Response means and measures of variance (SD or SE) were organized into an Excel spreadsheet with the following experimental details: authors, year, source of information, details of the HGP used, ageing details on the beef, country in which experiments were conducted, breed, sex, feeding system (pens or pasture), number of days that cattle were fed, whether carcasses were electrically stimulated or not, days that carcasses were chilled before processing, the cut or muscle group tested, whether meat was frozen or not, number of cattle (or pens) per treatment, and details of the outcomes and their measures of dispersion. Outcomes for this experiment included WBSF, and detailed meat quality responses, and taste-panel data for the juiciness, flavour, tenderness, connective tissue content, and Meat Standards Australia CMQ4 score. Some experiments reported different units of shear strength and Newtons were corrected to kg by dividing by 9.807. Some experiments reported different scales on which sensory outcomes were evaluated and these, with their respective measures of dispersion were retained on the basis that these were amenable to ES analysis, but would not allow a weighted mean difference (WMD) to be calculated.

3.4 Statistical analysis

Data was structured to allow a classical meta-analytical evaluation of differences in responses of the experimental groups to be assessed. There is a hierarchical structure in these data as many experiments used multiple treatment comparisons. Consequently, there is dependence within experiment and the effects of experiment and treatment need to be evaluated by meta-regression using multi-level models (Hedges *et al.* 2010; St-Pierre 2001; Van den Noortgate *et al.* 2013).

Variables that were examined by meta-regression included the length of time that beef was aged (‘ageing’), use of multiple implants or not (yes or no), use of trenbolone acetate (yes or no), breed (British, European, Holstein and crosses; Brahman and Brahman crosses; crossbred undescribed; not stated), sex [steer, bull, heifer, mixed (steers and heifers)], days on feed, and electrical stimulation of the carcass (yes, no, not stated). Freezing of the meat before evaluation was almost universal and length of time that meat was frozen before evaluation was not often reported. Consequently, this was not evaluated, nor was days chilled or vacuum packing of the meat as these were not consistently reported.

Model development. Initial data exploration included production of basic statistics using Stata (Version 14.2, StataCorp. LP, College Station, TX) to examine the data for errors and to estimate the means and measures of dispersion. Normality of the data was examined for continuous variables, by visual and statistical appraisal.

Further exploration of the factors that influenced outcomes was conducted using variables that had a $P < 0.2$ on univariable correlation analysis with the outcome variables. This method was used to reduce the potential for over-fitting models to the data (Dohoo *et al.* 2009). The effect of treatment within experiment was examined as a random effect using GLAMM (Stata 14.2) to partition the variance components of the nested model (Rabe-Hesketh and Skrondal 2005), and this effect explained a substantial amount (43.6%) of variation in responses above that explained by experiment alone.

Stata was also used to analyze differences in meat quality responses by SMD analysis which is also called ES analysis. These methods have been published in detail in Lean *et al.* (2009) and Golder and Lean (2016). The difference between treatment and reference groups means, which is termed 'treatment' in the following description, was standardized using the SD of reference and treatment groups. The SMD estimates were pooled using the DerSimonian and Laird (1986) random effects models. Only random effects models were used, as previous work concluded that when there was uncertainty in the evaluative units caused by clustering of observations, the random effects model was appropriate (White and Thomas 2005).

If a paper reported separate estimates of measures of variance (SE or SD) for each group, these were recorded as such. Many experiments reported a common SE or SD and these estimates were applied to both reference and treatment groups. A random-effects WMD between treatment and reference is provided for WBSF and CMQ4, with the weighting reflecting the inverse of the variance of the treatments included according to the nostandard method in the metan model of Stata to allow an interpretation of treatment effects in familiar units (kg of force), rather than ES. The other variables studied used scales that differed within the variable and were not amenable to WMD analysis.

Assessment of heterogeneity. Variations among the treatment level SMD were assessed using a chi-squared (Q) test of heterogeneity. Heterogeneity in treatments reflects underlying differences in clinical diversity of the herds and interventions, differences in experimental design and analytical methods, and statistical variation around responses. The clinical diversity of the herd includes all the non-study design aspects of variation, such as facility design, environment, animal management that may be measured and controlled for in meta-analysis, but are often not reported or measured. Identifying the presence and sources of the heterogeneity improves understanding of the responses to the interventions used. An α level of 0.10 was used because of the relatively poor power of the χ2 test to detect heterogeneity among small numbers of treatment comparisons (Egger and Smith 2001). Heterogeneity of results among the treatments was quantified using the I^2 statistic (Higgins and Thompson 2002), which was developed to measure the impact of heterogeneity on a meta-analysis from mathematical criteria that are independent of the number of treatment comparisons and the treatment effect measure. The measure, I^2 is a transformation of the square root of the χ2 heterogeneity statistic divided by its degrees of freedom and describes the proportion of total variation in treatment estimates that is due to heterogeneity. Further, I^2 provides an estimate of the proportion of the true variance of effects of the treatment, that is the true variance, tau2 (τ2) divided by the total variance observed in the treatment (Borenstein *et al.* 2017) that reflect measurement error. Negative values of I^2 are assigned a value of 0, consequently the value I^2 lies between 0 and 100%. An I^2 value between 0 and 40% might not be important, 30 to 60% may represent moderate heterogeneity, 50 to 90% might represent substantial heterogeneity, and 75 to 100% might represent considerable heterogeneity (Higgins and Green 2011). A 95% CI for I^2 was calculated using the heterogi command in Stata according to methods recommended by (Ioannidis *et al.* 2007). Both I^2 and τ^2 are provided to allow readers the opportunity to evaluate both metrics.

Meta-regression. A key focus of meta-analysis is to identify and understand the sources of heterogeneity or variation of response among treatments. Meta-regression analyses were used to explore the source of heterogeneity of response, using the individual SMD for each treatment as the outcome and the associated SE as the measure of variance. Meta-regression is also a technique that can formally test whether there is evidence of different effects in different subgroups of treatments (Knapp and Hartung 2003). We have previously published the equations used in meta-regression (Rabiee *et al.* 2012) and refer readers to these for a description of meta-regression using the methods of Thompson and Sharp (1999) and Knapp and Hartung (2003).

Backward stepping models were used for meta-regression that included variables with a univariable value of *P*-value < 0.2 obtained using the Knapp-Hartung method (Knapp and Hartung 2003). Models

were derived using the Knapp-Hartung method until the variables retained had a *P*-value < 0.1 when a permutation model was used to develop final models. The permutation test approach for assessing the statistical significance of meta-regression methods suggested by Higgins and Thompson (2004), and programmed by Harbord and Higgins (2008) and Harbord and Steichen (2004), was used to reduce the risk of type I error as described by Rabiee *et al.* (2012). The data are simulated under the null hypothesis of no association between effect estimates and any covariate, yet with an unexplained component of heterogeneity according to the standard random effects meta-analysis model (Higgins and Thompson 2004). Without loss of generality the average effect was assigned to zero:

$$\theta_i \sim N(0, \tau^2)$$
$$y_i \sim N(\theta_i, v_i) \text{ for } i=1,....,k \qquad \text{(Higgins and Thompson, 2004)}$$

Where an ES θ_i is estimated by y_i in treatment *i* for experiment 1,...., *k* with a mean of zero and variance τ^2 and v_i represents the within experiment variances.

Covariates are simulated from a multivariable (standard) normal distribution so that correlation is imposed between pairs of covariates. This process provides an assessment less likely to produce Type I statistical error (Higgins and Thompson 2004).

The results of the permutation test, which do not account for the hierarchical structure of the effects of treatment, are provided for comparison to robust regression models derived using the same starting variables that account for the nested effect of treatments within experiment (Hedges *et al.* 2010) and programmed as "robumeta" (Stata) and applied by Tanner-Smith and Tipton (2014). Hedges *et al.* (2010) developed the robust regression models to account for the two-stage cluster sampling inherent when the ES estimates are derived from a total of $n = k1 + k2 + \cdots + km$ estimates from treatments that were collected by sampling m clusters of experiments, that is several treatment estimates are derived from the same experiment. Hence, sampling $kj \geq 1$ estimates within the j^{th} cluster for $j = 1, \ldots, m$. Briefly, in this test the mean ES from a series of experiments is described as follows: In this case, the regression model has only an intercept *b*1 and the weighted mean has the form:

$$b1 = \frac{\sum_{j=1}^{m}\sum_{j=1}^{k1} wijTij}{\sum_{j=1}^{m}\sum_{j=1}^{k1} wij}$$

where m is the total number of studies, k the total number of treatments and wij is the weighting for treatments within experiments and Tij is the vector of the ES estimates of treatments within experiments. If all the estimates in the same experiment are given identical weights, the robust variance estimate (vR) reduces to:

$$v^{\mathrm{R}} = \frac{\sum_{j=1}^{m} w_j^2 (\check{\mathrm{T}}\mathrm{j} - \mathrm{b1})^2}{\left(\sum_{j=1}^{m} w\mathrm{j}\right)^2}$$

where Ťj is the unweighted mean of the estimates in the j^{th} cluster, *b*1 is the estimate of the weighted mean, and *wj* is the total weight given to estimates in the j^{th} cluster. This is a kind of weighted variance which reduces to (*m*-1)*/m*2 times the variance, when the weights within experiment are identical, and (since the correlation coefficient = 1 in this case) the robust regression standard error equals 1*/ m* times the variance of Ťj estimated when the weights are equal. Hedges et al. (2010) highlight several important aspects of the robust model and the underlying assumptions that; the correlation structure of the Tj does not need be known to compute the pooled ES or VR, only that the vectors of estimates from different experiments are independent and that regularity conditions are satisfied; the experiment or treatment level regressors

do not need to be fixed; the theorem is asymptotic based on the number of experiments, rather than the number of treatments; and the theorem is relatively robust to regularity assumptions.

Publication bias. Presence of publication bias was investigated using funnel plots, which are a simple scatter plot of the intervention effect estimates from individual treatments plotted against treatment precision. The name 'funnel plot' arises because precision of the intervention effect increases as the size and precision of a study increases. Effect estimates from treatments with a small number of animal units will scatter more widely at the bottom of the graph and the spread narrows for those with higher numbers of units. In the absence of bias, the plot should approximately resemble a symmetrical (inverted) funnel. If there is bias, for example because smaller treatments without statistically significant effects remain unpublished, this will lead to an asymmetrical appearance of the funnel plot and a gap will be evident in a bottom corner of the graph. In this situation, the effect calculated in a meta-analysis will tend to overestimate the intervention effect. The more pronounced the asymmetry, the more likely it is that the bias will be substantial. Data were screened for plausible quadratic relationships for these variables by visual appraisal of univariable scatter plots between the covariate and SMD of each treatment.

4 Results

Over 3000 experiments resulted from the literature searches with 182 experiments identified for review based on the pertinence of the title to this experiment and only 129 were pertinent and not repeated. Of these, 59 were excluded that did not meet the topic of interest or were rejected as review papers. Of the seventy remaining experiments, 38 were rejected for reasons that are outlined in Table 1. This left 32 experiments, one of which was rejected on the basis that the units of variation (rsd) produced an improbable SD, leaving 31 experiments containing 181 treatment comparisons. A PRISMA flow chart of the exclusions is provided as Figure 1. The tabulation of information on treatment comparisons is provided in Table 2 that lists authors, year, number of reference and treatment animals, sex of animals, first hormonal implant used, use of multiple implants (yes or no), use of trenbolone implants (yes or no), days meat was aged, days on feed, and mean WBSF for reference and treatment comparisons. Countries where treatment comparisons were conducted are USA (157), Australia (25), UK (1), and France (1). Information on descriptive statistics for the treatment comparisons is provided in Tables 3 and 4. There were relatively few observations in some categories for breed, for example undescribed cross-breeds, and sex for example bulls, or mixed heifers and steers. The lack of observations for breeds, other than the British category, Brahman and Brahman crosses and sex groups other than steers, limited the opportunities evaluate these effects in detail.

Table 1. List of experiments not included in meta-analysis and reason

Author	Year	Source	Reason
Berthiaume et al	2006	J. Anim. Sci. 84:2168-2177	No sensory evaluation
Café et al	2010	J. Anim. Sci. 88:3047-3058	No sensory evaluation
Café et al	2011	J. Anim. Sci. 89:1452-1465	No hormonal growth promotant
Choi et al	2013	Livestock Sci 157: 435-441	No hormonal growth promotant
Cleale et al	2013	J. Anim. Sci. 91:970-977	No sensory evaluation
Cranwell et al	1996	J. Anim. Sci. 1996. 74:1777-1783	Cows
Crouse et al			RSDs produced improbable SDs
Faucitano et al	2008	J. Ani. Sci. 86:1678-1689	Control confounded by monensin
Forest et al	1975	Can. J. Anim. Sci. 55:287-290	Pre 1980 and no Warner-Bratzler shear force data
Foutz et al	1989	Oklahoma Experimental Station	No reference group
Girard et al	2012	Can. J. Anim. Sci. 92:175-188	Calves mixed with yearlings in implant groups
Greathouse et al	1983	J. Anim. Sci. 57:355-363	No SD
Gruber et al	2011	J. Anim. Sci. 2011. 89:1401–1411	No reference group
Jeremiah et al	1988	Meat Sci. 22:83-101	Inadequate provision of SE values
Johnson et al	1986	J. Anim. Sci. 62:399-406	Pseudo replicated
Johnson et al	1984	J. Anim. Sci. 58: 920	Pseudo replicated
Jones et al	1991	J. Anim. Sci. 69:1363-1369	Calves
Kellermeier et al	2009	J. Anim. Sci. 87:3702-3711	No untreated reference group
Kellermeier et al	2010	J. Anim. Sci. 87:3702–3711	No data
Lowman et al	1991	Livestock Prod Sci 28;37-52	Calves
McEvers et al	2012	J. Anim. Sci. 90:4140-4147	No hormonal growth promotant
Monson et al	2007	J. Muscle Foods 18:173-185	Confounded
Morgan	1997	Oklahoma State University Animal Science	Review
Ouali et al	1988	Meat Sci. 24:151-161	Measurements on the triceps
Pritchard et al	2003	South Dakota Beef Report Paper 12	Wrong unit of analysis
Pruneda et al	1999	Oklahoma State University Animal Science Research Report	Wrong units
Reinhardt et al	2014	J. Anim. Sci. 92:4711-4718	Meta-analysis
Roy et al	2015	Meat Sci. 110: 109-117	No relevant data
Schoonmaker et al	2001	J. Anim. Sci. 79:1074-1084	No reference group
Schutt et al	2009	Anim. Prod. Sci. 49: 439-451	Calves
Simone et al	1958	J. Anim. Sci 17; 834-840	No Warner-Bratzler shear force measure
Strydom et al	2010	Animal 4: 653-660	Cows
Tait et al	2014	J. Anim. Sci. 92:456-466	No hormonal growth promotant
Thonney et al	1991	J. Anim. Sci. 69:4866-4870	No Warner-Bratzler shear force measure
Vanderwert et al	1986	J. Anim. Sci. 63:114-120	Age not reported
Watson et al	2008	Aust. J. Exp. Agric. 48; 1415-1424	No Warner-Bratzler shear force measure

Wilson et al	1999	J. Anim. Sci. 1999. 77:3133–3140	No Warner-Bratzler shear force measure
Woerner et al	2011	J. Anim. Sci. 2011. 89:201-209	No reference group
Xiong et al.	1996	Food Res. Int. 29:27-34	Sensory data only

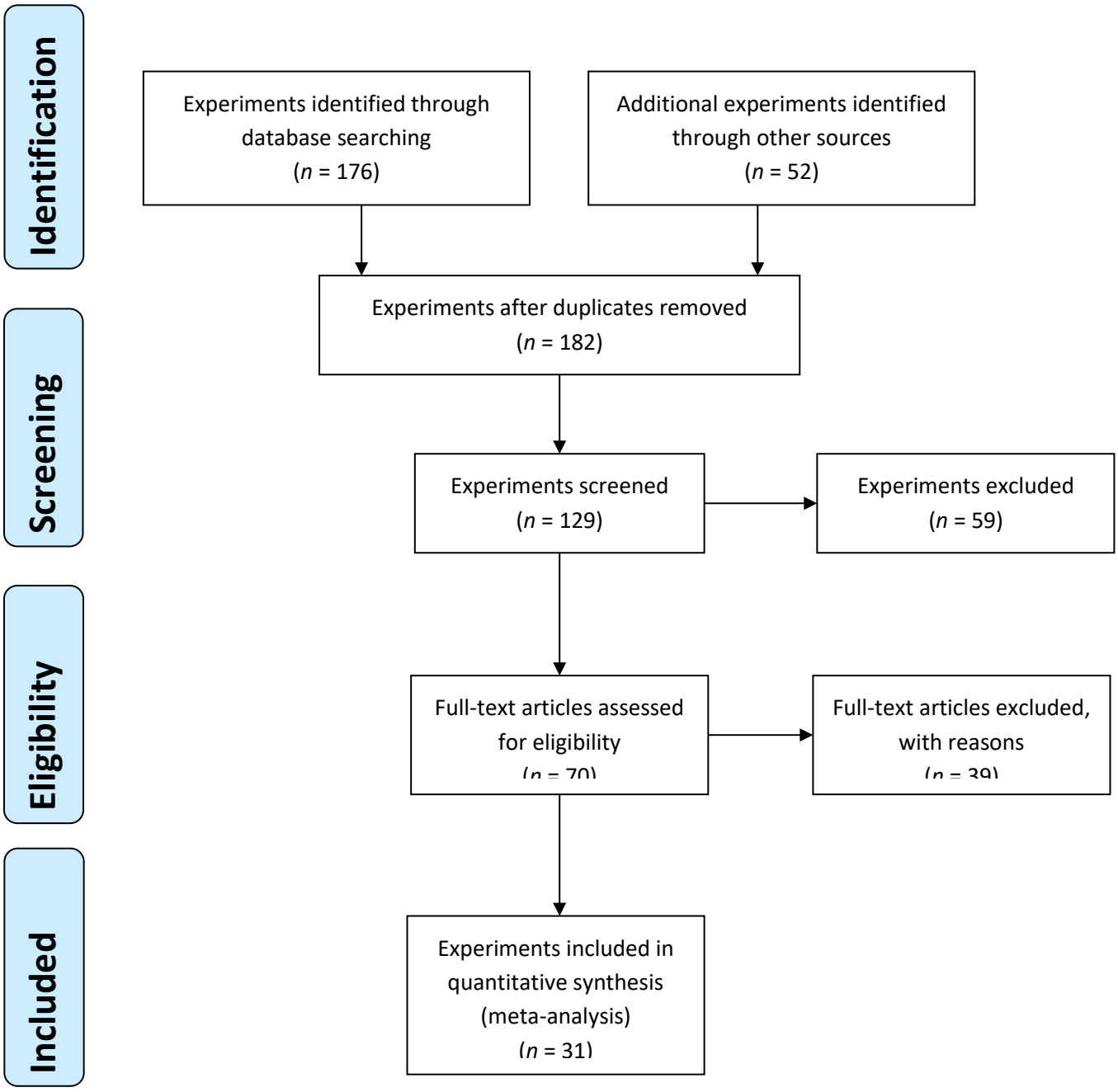

Fig 1. PRISMA flow diagram showing the number of experiments identified through database searching, other sources, experiments remaining after duplicates removed, experiments screened and excluded, full text articles accessed for eligibility and reasons for exclusion, and experiments included in quantitative synthesis. The concept of a PRIMA statement was adapted from Moher *et al.* (2009)

Almost all studies were conducted on longissimus dorsi or longissimus muscle (LD) (which was variously described by terms including strip loin), except for one (Hunter *et al.* 2000) that used semitendinosis. Two studies (Cheatham *et al.* 2008; Foutz *et al.* 1997) used rib cross-sections that would have contained longissimus muscle. The sensory measures were inconsistently reported and the most frequently reported term relating to those measures was the one selected for inclusion. However, where this term was not reported, alternate, but similar, measures were used. Specifically, the term juiciness included “juiciness”, “initial juiciness”, and “sustained juiciness”. If more than one of these three measures were used in a study, “juiciness” was used by preference. “Tenderness” included “myofibrillar tenderness”, “overall tenderness”, “initial tenderness” and “sustained tenderness”. By preference, when more than one measure was present, “overall tenderness” was used. “Flavour” included “flavour intensity”, “flavour desirability”, and “beef flavour”. The terms “off flavour” or “flavour of lean” were not used. Table 1 shows that there were many different HGP treatments applied and that these were used in a large variety of different combinations. In order to evaluate some aspects of the treatment regimens the use of trenbolone acetate (TBA) in a treatment comparison was examined as was the use of multiple implants. It was unfortunate that chilling effects and days frozen were inconsistently reported and could not be analyzed.

Table 2. Summary of descriptors for each treatment comparison used in the meta-analysis including a list of authors, year of publication, number of animals in the reference and treatment comparisons, sex of cattle, name of first hormonal implant used, use of multiple implants (yes or no), the number of days meat was aged, the number of days cattle were fed, and the mean Warner-Bratzler shear force (WBSF) for the reference and treatment groups

		Number of animals								Mean WBSF[3], kg	
Author	Year	Reference	Treatment	Sex[1]	Hormonal implant 1	Multiple implants	TBA[2] use	Days aged	Days fed	Reference	Treatment
Apple et al	1991	3	3	S	Ralgro	No	No	6	249	4.01	4.01
Apple et al	1991	3	3	S	Synovex-S	No	No	6	249	4.01	3.93
Apple et al	1991	3	3	S	Finaplix-S	No	Yes	6	249	4.01	4.06
Apple et al	1991	3	3	S	Finaplix-S	Yes	Yes	6	249	4.01	4.35
Apple et al	1991	3	3	S	Finaplix-S	Yes	Yes	6	249	4.01	4.30
Barham et al	2003	1368	660	S	Synovex-S	Yes	No	3	210	3.44	3.57
Barham et al	2003	1368	720	S	Synovex-S	Yes	Yes	3	210	3.44	3.51
Boles et al	2009	32	32	S/H	Ralgro	Yes	Yes		120	5.90	6.50
Boles et al	2009	37	37	S/H	Vet Life	No	Yes		120	6.80	7.90
Café et al	2010	83	81	S/H	Revalor-H	No	Yes	1	117	7.59	8.42
Café et al	2010	83	81	S/H	Revalor-H	No	Yes	7	117	7.29	7.66
Café et al	2010	83	81	S/H	Revalor-H	No	Yes	1	117	4.55	4.90
Café et al	2010	83	81	S/H	Revalor-H	No	Yes	7	117	4.50	4.84
Café et al	2010	71	72	S	Revalor-H	No	Yes	1	80	4.98	5.59
Café et al	2010	71	72	S	Revalor-H	No	Yes	7	80	4.77	5.41
Café et al	2010	71	72	S	Revalor-H	No	Yes	1	80	5.19	5.65
Café et al	2010	71	72	S	Revalor-H	No	Yes	7	80	4.54	4.87
Calkins et al	1986	4	4	B	Ralgro	Yes	No	10	232	2.31	2.32
Calkins et al	1986	4	4	S	Ralgro	Yes	No	10	232	2.16	2.31
Calkins et al	1986	4	4	B	Compudose 200	Yes	No	10	232	2.31	2.20
Calkins et al	1986	4	4	S	Compudose 200	Yes	No	10	232	2.16	2.33
Cheatham et al	2008	5	5	S	Ralgro	Yes	No	2	259	1.98	2.14
Cheatham et al	2008	5	5	S	Ralgro	Yes	Yes	2	259	1.98	2.25
Cheatham et al	2008	5	4	S	Ralgro	Yes	Yes	2	259	1.98	2.52

Ebarb et al	2016	11	11	H	Component TE-200	No	No	35	75	4.37	4.52
Ebarb et al	2017	11	11	H	Component TE-200	No	No	2	90	5.09	5.54
Ebarb et al	2017	11	11	H	Component TE-200	No	No	7	90	4.27	4.78
Foutz et al	1997	4	4	S	Synovex S	No	Yes	7	119-126	4.00	4.43
Foutz et al	1997	4	4	S	Revalor	No	Yes	7	119-127	4.00	4.32
Foutz et al	1997	4	4	S	Finaplix S	No	Yes	7	119-128	4.00	4.12
Foutz et al	1997	4	4	S	Finaplix S	Yes	Yes	7	119-129	4.00	4.41
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	7	152-174	2.43	2.79
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	14	152-174	2.55	2.78
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	21	152-174	2.50	2.63
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	28	152-174	1.87	2.12
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	35	152-174	2.60	2.87
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	7	152-174	2.43	2.74
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	14	152-174	2.55	2.95
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	21	152-174	2.50	2.90
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	28	152-174	1.87	2.30
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	35	152-174	2.60	2.62
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	7	152-174	3.58	4.19
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	14	152-174	3.59	4.14
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	21	152-174	3.29	3.86
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	28	152-174	2.58	3.42
Garmyn et al	2011	16	16	S	Revalor S	No	Yes	35	152-174	2.89	3.21
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	7	152-174	3.58	3.80
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	14	152-174	3.59	4.06
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	21	152-174	3.29	3.68
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	28	152-174	2.58	2.85
Garmyn et al	2011	16	16	S	Revalor XS	No	Yes	35	152-174	2.89	2.88
Gerken et al	1995	6	6	S	Synovex-S	No	No	14	112	3.98	4.56
Gerken et al	1995	6	6	S	Finaplix-S	No	Yes	14	112	3.98	3.93

Gerken et al	1995	6	6	S	Revalor-S	No	Yes	14	112	3.98	4.65
Hopkins and Dikeman	1987	3	3	B	Compudose	Yes	No	10	205	5.20	4.40
Hunt et al	1991	5	5	S	Finnaplix-120	Yes	Yes	7	160	3.40	3.30
Hunt et al	1991	5	5	B	Finnaplix-120	Yes	Yes	7	160	4.40	3.50
Hunt et al	1991	5	5	S	Finnaplix-120	Yes	Yes	7	160	3.40	3.20
Hunt et al	1991	5	5	B	Finnaplix-120	Yes	Yes	7	160	4.40	3.60
Hunter et al	2000	17	16	S	Compudose 400	No	No	Unknown	420	5.10	5.50
Hunter et al	2000	17	16	S	Compudose 100	Yes	No	Unknown	420	5.10	5.60
Hunter et al	2001	20	17	S	Compudose 100	No	No	1	100	4.30	4.80
Hunter et al	2001	16	16	S	Compudose 100	No	No	1	150	4.70	5.40
Hunter et al	2001	18	17	S	Compudose 100	No	No	1	70	4.40	4.50
Hunter et al	2001	17	12	S	Compudose 100	No	No	1	Unknown	6.00	6.30
Igo et al	2011	4	7	S	Revalor XS	No	Yes	14	145-174	3.20	3.00
Igo et al	2011	4	7	S	Revalor IS	Yes	Yes	14	145-174	3.20	3.20
Igo et al	2011	4	7	S	Revalor XS	No	Yes	21	145-174	2.90	2.90
Igo et al	2011	4	7	S	Revalor IS	Yes	Yes	21	145-174	2.90	2.90
Igo et al	2011	4	7	S	Revalor XS	No	Yes	14	145-174	3.00	2.90
Igo et al	2011	4	7	S	Revalor IS	Yes	Yes	14	145-174	3.00	3.30
Igo et al	2011	4	7	S	Revalor XS	No	Yes	21	145-174	2.70	2.60
Igo et al	2011	4	7	S	Revalor IS	Yes	Yes	21	145-174	2.70	2.80
Kerth et al	2003	8	8	H	Revalor-H	No	Yes	16	Unknown	3.49	3.54
Kerth et al	2003	8	8	H	Revalor-H	No	Yes	16	Unknown	3.49	2.93
Kerth et al	2003	8	8	H	Revalor-H	Yes	Yes	16	Unknown	3.49	3.18
Kerth et al	2003	8	8	H	Revalor-IH	Yes	Yes	16	Unknown	3.49	3.34
Kerth et al	2003	8	8	H	Synovex-H	Yes	Yes	16	Unknown	3.49	3.39
Nute and Dransfield	1984	12	12	S	Ralgro	No	No	6	Unknown		
Ouali et al	1988	10	10	S	Revalor S	No	Yes	7	130		
Packer et al	In press	100	100	S	Compudose 100	No	No	7	73	4.40	4.60

Packer et al	In press	100	100	S	Compudose 100	No	No	35	73	3.40	3.50
Packer et al	In press	100	100	S	Component TE-200	No	Yes	7	73	4.40	4.70
Packer et al	In press	100	100	S	Component TE-200	No	Yes	35	73	3.40	3.50
Phelps et al	2014	16	16	S	Component E-S	Yes	No	21	175	3.20	3.42
Phelps et al	2014	16	16	S	Component E-S	Yes	No	21	175	3.00	3.55
Platter et al	2003	50	50	S	Synovex-S	Yes	Yes	17.5	Various	3.54	3.95
Platter et al	2003	50	50	S	Ralgro	Yes	Yes	17.5	Various	3.54	4.46
Platter et al	2003	50	50	S	Synovex-S	Yes	Yes	17.5	Various	3.54	4.19
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.19
Platter et al	2003	50	50	S	Ralgro	Yes	Yes	17.5	Various	3.54	4.15
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.12
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.05
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.05
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.14
Platter et al	2003	50	50	S	Synovex-C	Yes	Yes	17.5	Various	3.54	4.38
Reiling and Johnson	2003	40	41	S	Ralgro	Yes	Yes	14	105	3.06	3.28
Reiling and Johnson	2003	40	42	S	Revalor-S	Yes	Yes	14	105	3.06	3.58
Reiling and Johnson	2003	41	41	S	Component TE-S	Yes	No	5	105	3.76	4.09
Reiling and Johnson	2003	41	41	S	Component TE-S	Yes	No	14	105	3.54	3.72
Robinson et al	2012	187	176	S/H	Revalor–H	No	Yes	7	390-660		
Robinson et al	2012	187	176	S/H	Revalor–H	No	Yes	7	390-661		
Roeber et al	2000	36	39	S	Encore	Yes	Yes	14	140 or 141	2.97	3.18
Roeber et al	2000	36	38	S	Ralgro	Yes	Yes	14	140 or 141	2.97	3.41

Roeber et al	2000	36	38	S	Ralgro	Yes	Yes	14	140 or 141	2.97	3.31
Roeber et al	2000	36	36	S	Revalor-S	Yes	Yes	14	140 or 141	2.97	3.28
Roeber et al	2000	36	36	S	Revalor-S	No	Yes	14	140 or 141	2.97	3.51
Roeber et al	2000	36	37	S		No	Yes	14	140 or 141	2.97	3.42
Roeber et al	2000	36	37	S	Synovex Plus	No	Yes	14	140 or 141	2.97	3.29
Rumsey et al	1990	10	10	S	Synovex-S	Yes	No	2	160	3.69	3.87
Rumsey et al	1990	19	19	S/H	Synovex-S	Yes	No	2	160	4.70	6.05
Samber et al	1996	8	8	S	Ralgro	Yes	Yes	14	212	2.58	2.74
Samber et al	1996	8	8	S	Ralgro	Yes	Yes	14	212	2.58	2.75
Samber et al	1996	8	8	S	Synovex-S	Yes	Yes	14	212	2.58	2.64
Samber et al	1996	8	8	S	Revalor-S	Yes	Yes	14	212	2.58	3.01
Samber et al	1996	8	8	S	Revalor-S	Yes	Yes	14	212	2.58	2.92
Scheffler et al	2003	4	4	S	Component TE-S	Yes	Yes	14	269	2.50	2.60
Scheffler et al	2003	4	4	S	Component TE-S	Yes	Yes	14	269	2.50	2.80
Scheffler et al	2003	4	4	S	Component TE-S	Yes	Yes	14	269	2.50	3.00
Schneider et al	2007	42	41	H	TBA	No	Yes	3	140	4.67	4.51
Schneider et al	2007	42	41	H	TBA	No	Yes	7	140	4.22	4.22
Schneider et al	2007	42	41	H	TBA	No	Yes	14	140	3.80	3.59
Schneider et al	2007	42	41	H	TBA	No	Yes	21	140	3.33	3.36
Schneider et al	2007	42	41	H	TBA	No	Yes	28	140	3.27	3.24
Schneider et al	2007	42	42	H	TBA + E2	No	Yes	3	140	4.67	4.57
Schneider et al	2007	42	42	H	TBA + E2	No	Yes	7	140	4.22	4.06
Schneider et al	2007	42	42	H	TBA + E2	No	Yes	14	140	3.80	3.56
Schneider et al	2007	42	42	H	TBA + E2	No	Yes	21	140	3.33	3.26
Schneider et al	2007	42	42	H	TBA + E2	No	Yes	28	140	3.27	3.13
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	3	140	4.67	4.67
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	7	140	4.22	4.33

Schneider et al	2007	42	41	H	TBA + E2	No	Yes	14	140	3.80	3.84
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	21	140	3.33	3.45
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	28	140	3.27	3.23
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	3	140	4.67	4.74
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	7	140	4.22	4.37
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	14	140	3.80	3.71
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	21	140	3.33	3.44
Schneider et al	2007	42	41	H	TBA + E2	No	Yes	28	140	3.27	3.19
Schneider et al	2007	42	41	H	TBA	Yes	Yes	3	140	4.67	4.65
Schneider et al	2007	42	41	H	TBA	Yes	Yes	7	140	4.22	4.30
Schneider et al	2007	42	41	H	TBA	Yes	Yes	14	140	3.80	3.73
Schneider et al	2007	42	41	H	TBA	Yes	Yes	21	140	3.33	3.43
Schneider et al	2007	42	41	H	TBA	Yes	Yes	28	140	3.27	3.39
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	3	140	4.67	5.03
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	7	140	4.22	4.47
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	14	140	3.80	3.87
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	21	140	3.33	3.51
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	28	140	3.27	3.26
Schneider et al	2007	42	41	H	TBA + E2	Yes	Yes	3	140	4.67	5.06
Schneider et al	2007	42	41	H	TBA + E2	Yes	Yes	7	140	4.22	4.66
Schneider et al	2007	42	41	H	TBA + E2	Yes	Yes	14	140	3.80	4.05
Schneider et al	2007	42	41	H	TBA + E2	Yes	Yes	21	140	3.33	3.67
Schneider et al	2007	42	41	H	TBA + E2	Yes	Yes	28	140	3.27	3.39
Schneider et al	2007	42	40	H	TBA + E2	Yes	Yes	3	140	4.67	5.41
Schneider et al	2007	42	40	H	TBA + E2	Yes	Yes	7	140	4.22	4.87
Schneider et al	2007	42	40	H	TBA + E2	Yes	Yes	14	140	3.80	4.20
Schneider et al	2007	42	40	H	TBA + E2	Yes	Yes	21	140	3.33	3.74
Schneider et al	2007	42	40	H	TBA + E2	Yes	Yes	28	140	3.27	3.50
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	3	140	4.67	5.31
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	7	140	4.22	4.73
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	14	140	3.80	4.11

Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	21	140	3.33	3.62
Schneider et al	2007	42	42	H	TBA + E2	Yes	Yes	28	140	3.27	3.42
Schneider et al	2007	42	44	H	TBA + E2	Yes	Yes	3	140	4.67	5.46
Schneider et al	2007	42	44	H	TBA + E2	Yes	Yes	7	140	4.22	5.00
Schneider et al	2007	42	44	H	TBA + E2	Yes	Yes	14	140	3.80	4.21
Schneider et al	2007	42	44	H	TBA + E2	Yes	Yes	21	140	3.33	3.77
Schneider et al	2007	42	44	H	TBA + E2	Yes	Yes	28	140	3.27	3.36
Schneider et al	2007	42	43	H	TBA + E2	Yes	Yes	3	140	4.67	5.56
Schneider et al	2007	42	43	H	TBA + E2	Yes	Yes	7	140	4.22	5.09
Schneider et al	2007	42	43	H	TBA + E2	Yes	Yes	14	140	3.80	4.36
Schneider et al	2007	42	43	H	TBA + E2	Yes	Yes	21	140	3.33	3.76
Schneider et al	2007	42	43	H	TBA + E2	Yes	Yes	28	140	3.27	3.66
Shackelford et al	1992	48	48	B	Ralgro	No	No	10	190, 246, 315	4.30	5.10
Shackelford et al	1992	48	48	B	Synovex-S	No	No	10	190, 246, 315	4.30	5.10
Thompson et al	2008	20	20	S	Revalor-S	No	Yes	5	55 or 65	3.60	4.00
Thompson et al	2008	20	20	H	Revalor-H	No	Yes	5	55 or 65	4.30	5.20
Thompson et al	2008	20	20	S	Revalor-S	No	Yes	21	55 or 65	3.00	3.30
Thompson et al	2008	20	20	H	Revalor-H	No	Yes	21	55 or 65	3.20	3.60
Thompson et al	2008	240	235	S	Compudose 100	No	No	1	55 or 65	5.80	5.80

[1]Sex categories; S; steers; H; heifers; B; bulls;

[2]TBA; trenbolone acetate implants;

[3]WBSF; Warner-Bratzler shear force

Table 3. Descriptive statistics for number of experiments, treatment comparisons used for multiple hormonal growth promotant implants, treatments using trenbolone acetate, length of time that meat was aged before evaluation, length of time that cattle were fed, and number of animals or pens per treatment

Variable	Number of treatment comparisons	Percentage or mean	SD	Minimum	Maximum
Multiple implants, % of treatments	181	50	0.5	NA	NA
Trenbolone acetate, % of treatments	181	83	0.4	NA	NA
Ageing of meat, d	177	13	8.8	1	35
Length of feeding, d	160	151	54.1	60	420
Number of animals or pens per treatment	181	39.9	75.5	3	720

Table 4. Frequency distribution of breed, sex, and electrical stimulation at slaughter categories for 181 treatments comparisons

Variable	Frequency	Percentage, %
Breed		
British and European breeds, British and European cross, and Holstein	129	71.3
Brahman and Brahman crosses	32	17.7
Crossbred (undescribed)	16	8.8
Not stated	4	2.2
Sex		
Steers	100	55.3
Bull	7	3.9
Heifers	65	35.9
Mixed (steers and heifers)	9	5.0
Electrical stimulation at slaughter		
Not stimulated	23	12.7
Stimulated	77	42.5
Not stated	81	44.8

There was no evidence of publication bias in the funnel plots. The funnel plot for WBSF, tenderness, juiciness, and flavour are provided in Figures 2-5, respectively.

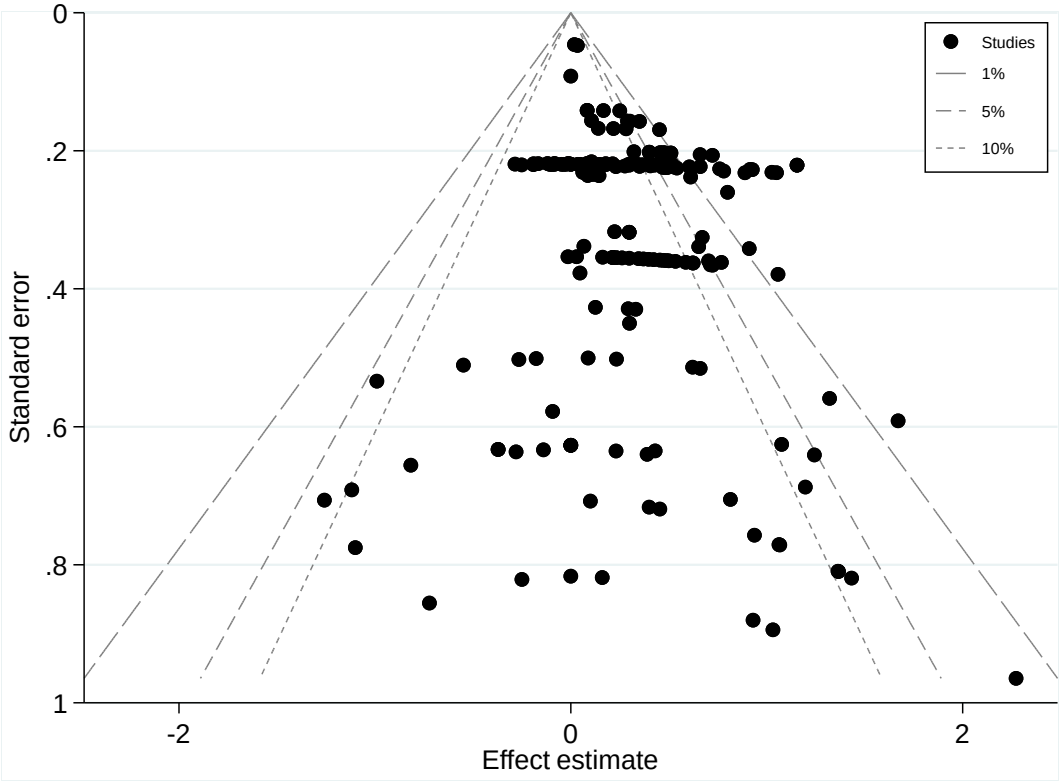

Fig 2. Contour-enhanced funnel plot for effects of hormonal growth promotants on the difference in Warner-Bratzler shear force (kg) of primarily the Longissimus dorsi muscle in beef cattle. The grey broken lines represent the 90, 95, and 99% CI for treatment comparisons. Effect estimates from small

studies will scatter more widely at the bottom of the graph and the spread narrows for larger treatments (Sterne and Harbord 2004). In the absence of heterogeneity or bias the plot should approximately resemble a symmetrical (inverted) funnel with studies lying within these lines. If there is bias, for example because smaller treatments without statistically significant effects remain unpublished, this will lead to an asymmetrical appearance of the funnel plot and a gap will be evident in a bottom corner of the graph.

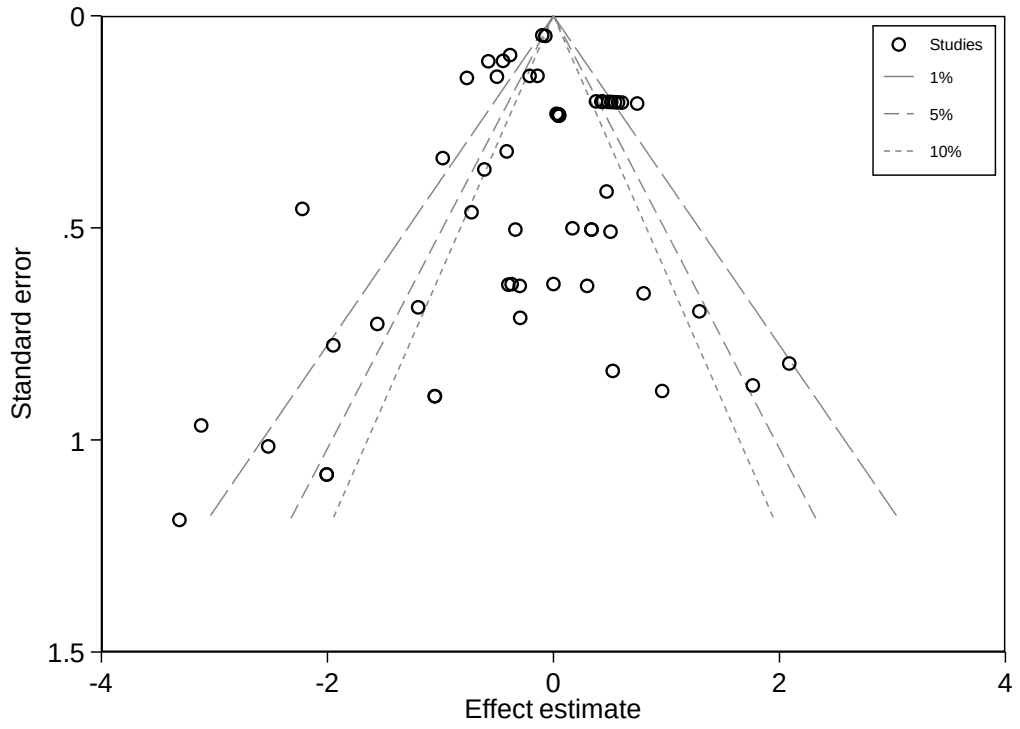

Fig 3. Contour-enhanced funnel plot for effects of hormonal growth promotants on the difference in tenderness of primarily the Longissimus dorsi muscle in beef cattle. The grey broken lines represent the 90, 95, and 99% CI for treatment comparisons. Effect estimates from small studies will scatter more widely at the bottom of the graph and the spread narrows for larger treatments (Sterne and Harbord 2004). In the absence of heterogeneity or bias the plot should approximately resemble a symmetrical (inverted) funnel with studies lying within these lines. If there is bias, for example because smaller treatments without statistically significant effects remain unpublished, this will lead to an asymmetrical appearance of the funnel plot and a gap will be evident in a bottom corner of the graph.

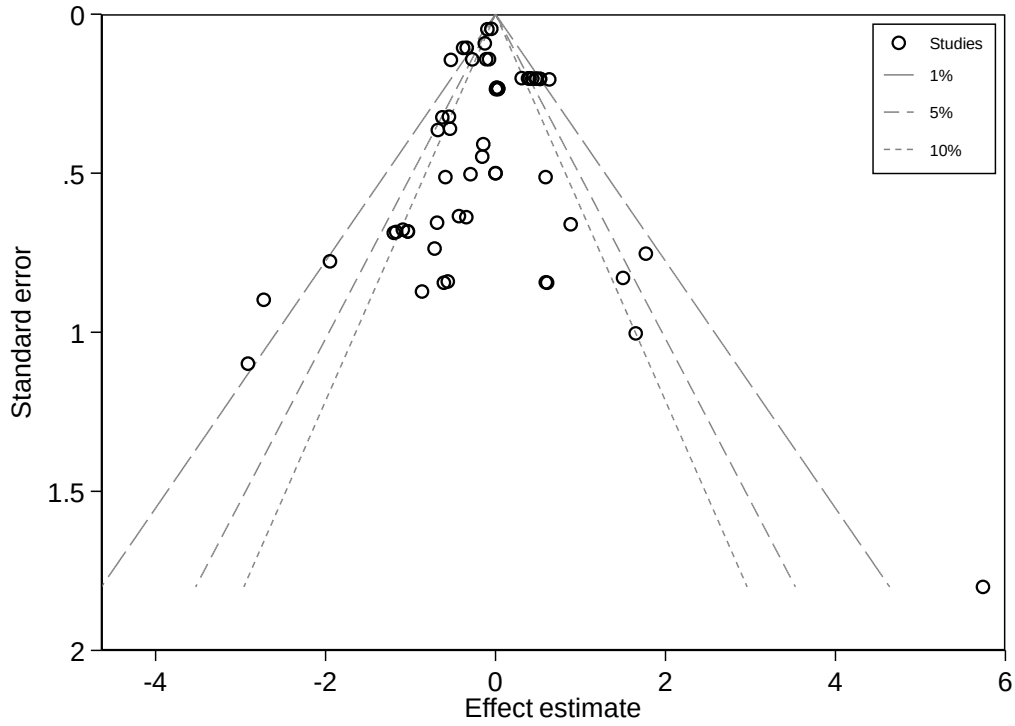

Fig 4. Contour-enhanced funnel plot for effects of hormonal growth promotants on the difference in juiciness of primarily the Longissimus dorsi muscle in beef cattle. The grey broken lines represent the 90, 95, and 99% CI for treatment comparisons. Effect estimates from small studies will scatter more widely at the bottom of the graph and the spread narrows for larger treatments (Sterne and Harbord 2004). In the absence of heterogeneity or bias the plot should approximately resemble a symmetrical (inverted) funnel with studies lying within these lines. If there is bias, for example because smaller treatments without statistically significant effects remain unpublished, this will lead to an asymmetrical appearance of the funnel plot and a gap will be evident in a bottom corner of the graph.

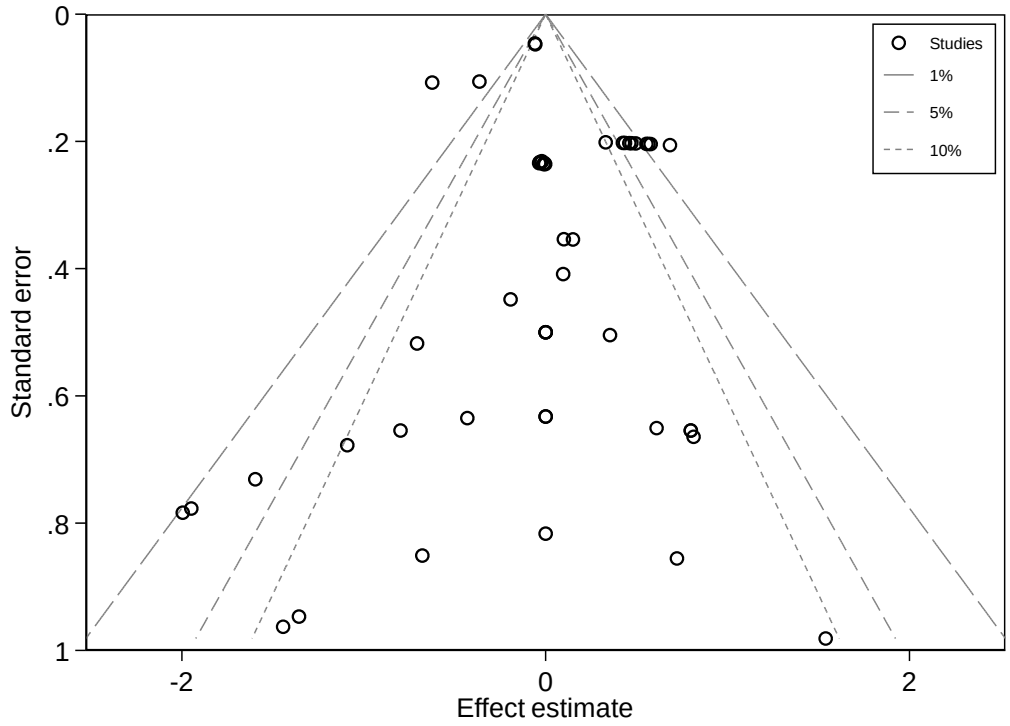

Fig 5. Contour-enhanced funnel plot for effects of hormonal growth promotants on the difference in flavour of primarily the Longissimus dorsi muscle in beef cattle. The grey broken lines represent the 90, 95, and 99% CI for treatment comparisons. Effect estimates from small studies will scatter more widely at the bottom of the graph and the spread narrows for larger treatments (Sterne and Harbord 2004). In the absence of heterogeneity or bias the plot should approximately resemble a symmetrical (inverted) funnel with studies lying within these lines. If there is bias, for example because smaller treatments without statistically significant effects remain unpublished, this will lead to an asymmetrical appearance of the funnel plot and a gap will be evident in a bottom corner of the graph.

Forest plots of the responses were created and associations between HGP treatments and tenderness, juiciness, and flavour are displayed in Figures 6 to 8, using the estimated SMD of the outcomes with both the DerSimonian and Laird (1986) and the Knapp-Hartung summary estimates. Due to the large number of treatment comparisons for WBSF the forest plot for this outcome is not shown.

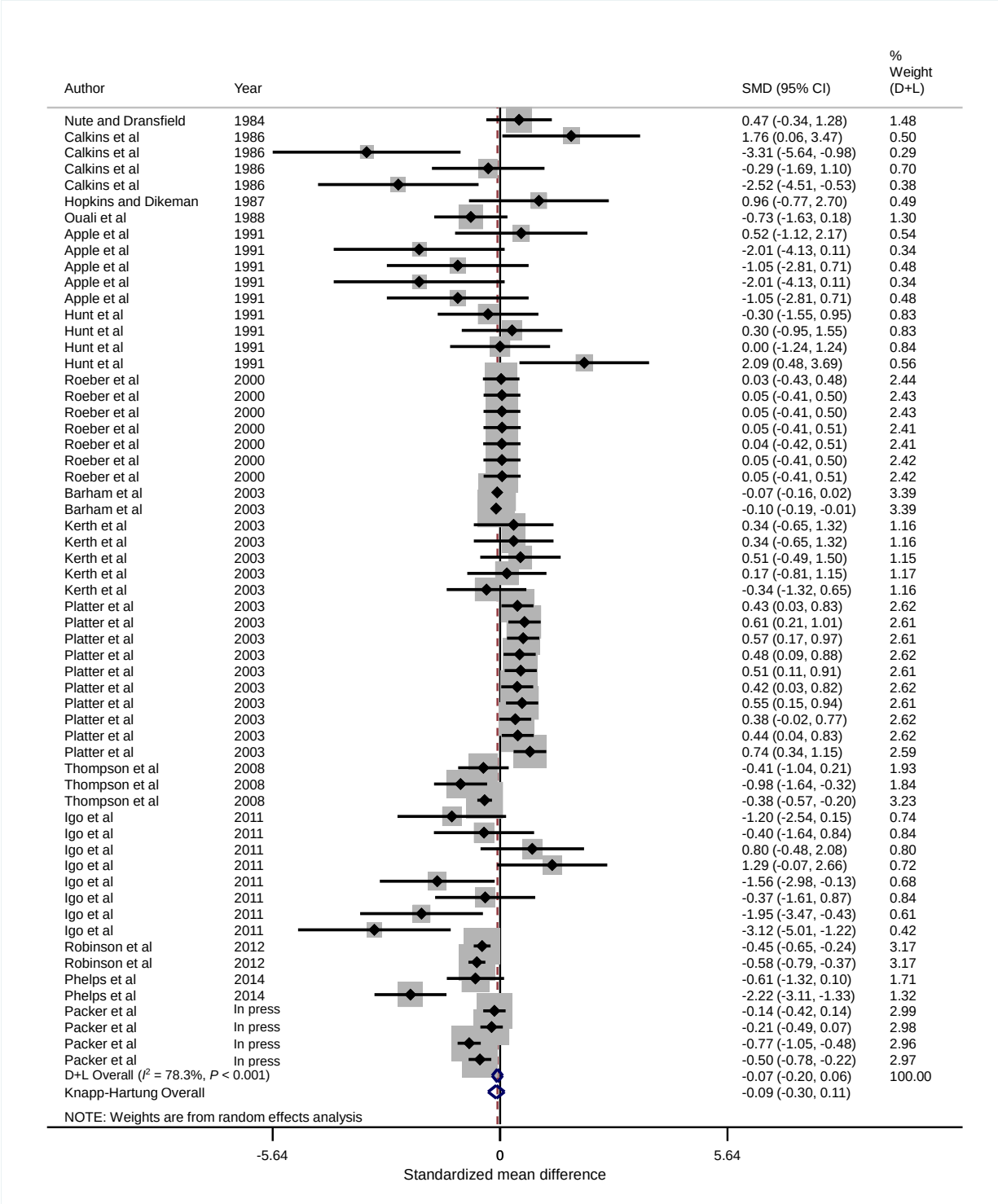

Fig 6. Forest plot of the effect size or standardized mean difference (SMD; standardized using the z-statistic) and 95% CI of the effect of hormonal growth prototants on tenderness of primarily the Longissimus dorsi muscle in beef cattle. The solid vertical line represents a mean difference of zero or no effect. Points to the left of the line represent a decrease in retained body nitrogen, while points to the right of the line indicate an increase. Each square around the point effect represents the mean effect size for that study and reflects the relative weighting of the study to the overall effect size estimate. The larger the box, the greater the study contribution to the overall estimate. The weight that each study contributed are in the right-hand column. The upper and lower limit of the line connected to the square represents the upper and lower 95% CI for the effect size. The overall pooled

effects size or SMD and 95% CI pooled using the DerSimonian and Laird (D + L; DerSimonian and Laird 1986) and Knapp-Sidak-Jonkman (Knapp-Hartung; IntHout *et al.* 2014) methods for random effects models are indicated by the respective diamonds at the bottom. The heterogeneity measure, I^2 is a measure of variation beyond chance among treatments included in the meta-analysis. Tenderness was substantially heterogeneous as indicated by the I^2 of 78.3%.

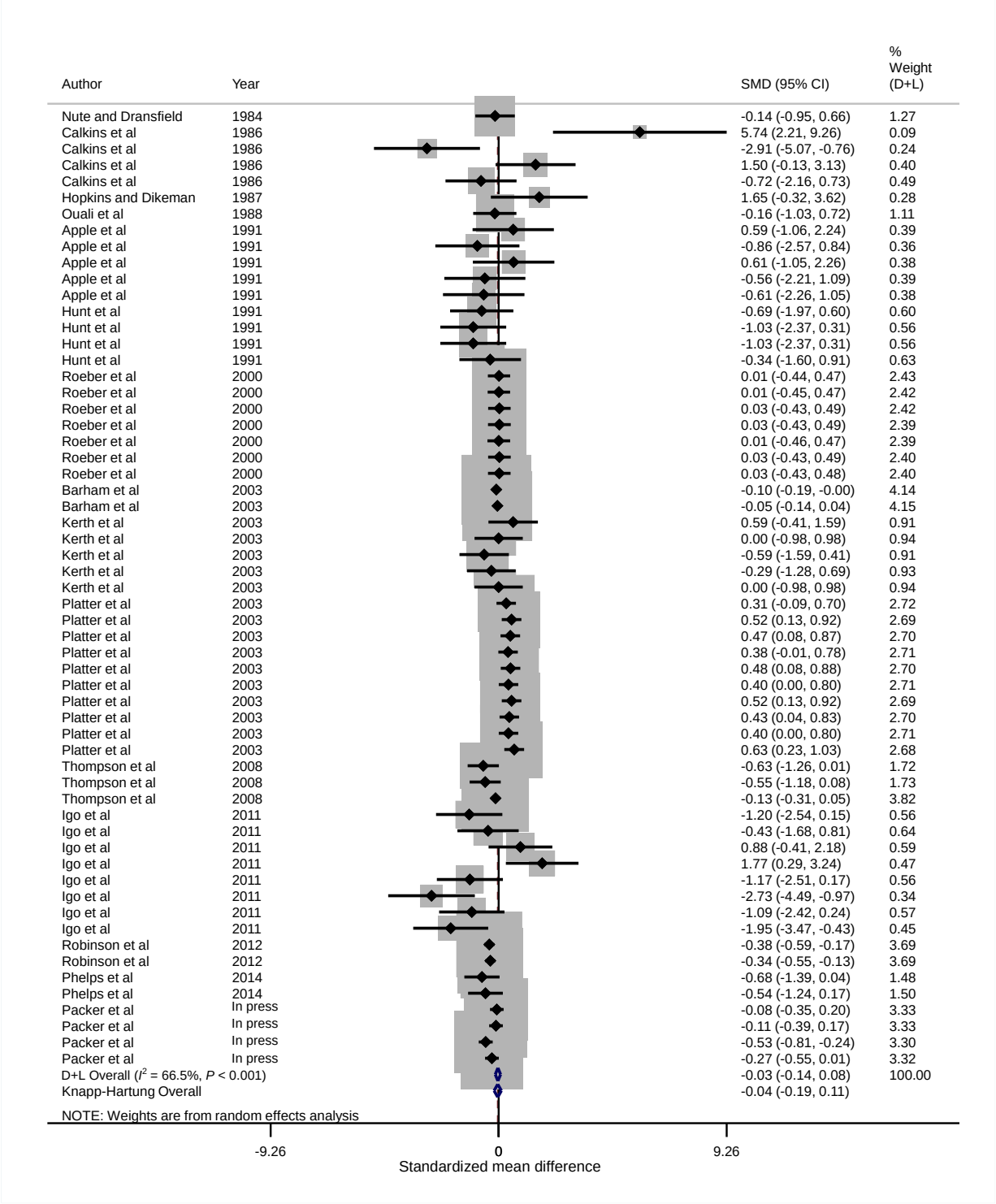

Fig 7. Forest plot of the effect size or standardized mean difference (SMD; standardized using the z-statistic) and 95% CI of the effect of hormonal growth prototants on juiciness of primarily the Longissimus dorsi muscle in beef cattle. The solid vertical line represents a mean difference of zero or

no effect. Points to the left of the line represent a decrease in retained body nitrogen, while points to the right of the line indicate an increase. Each square around the point effect represents the mean effect size for that study and reflects the relative weighting of the study to the overall effect size estimate. The larger the box, the greater the study contribution to the overall estimate. The weight that each study contributed are in the right-hand column. The upper and lower limit of the line connected to the square represents the upper and lower 95% CI for the effect size. The overall pooled effects size or SMD and 95% CI pooled using the DerSimonian and Laird (D + L; DerSimonian and Laird 1986) and Knapp-Sidak-Jonkman (Knapp-Hartung; IntHout *et al.* 2014) methods for random effects models are indicated by the respective diamonds at the bottom. The heterogeneity measure, I^2 is a measure of variation beyond chance among treatments included in the meta-analysis. Juiciness was moderately heterogeneous as indicated by the I^2 of 66.5%.

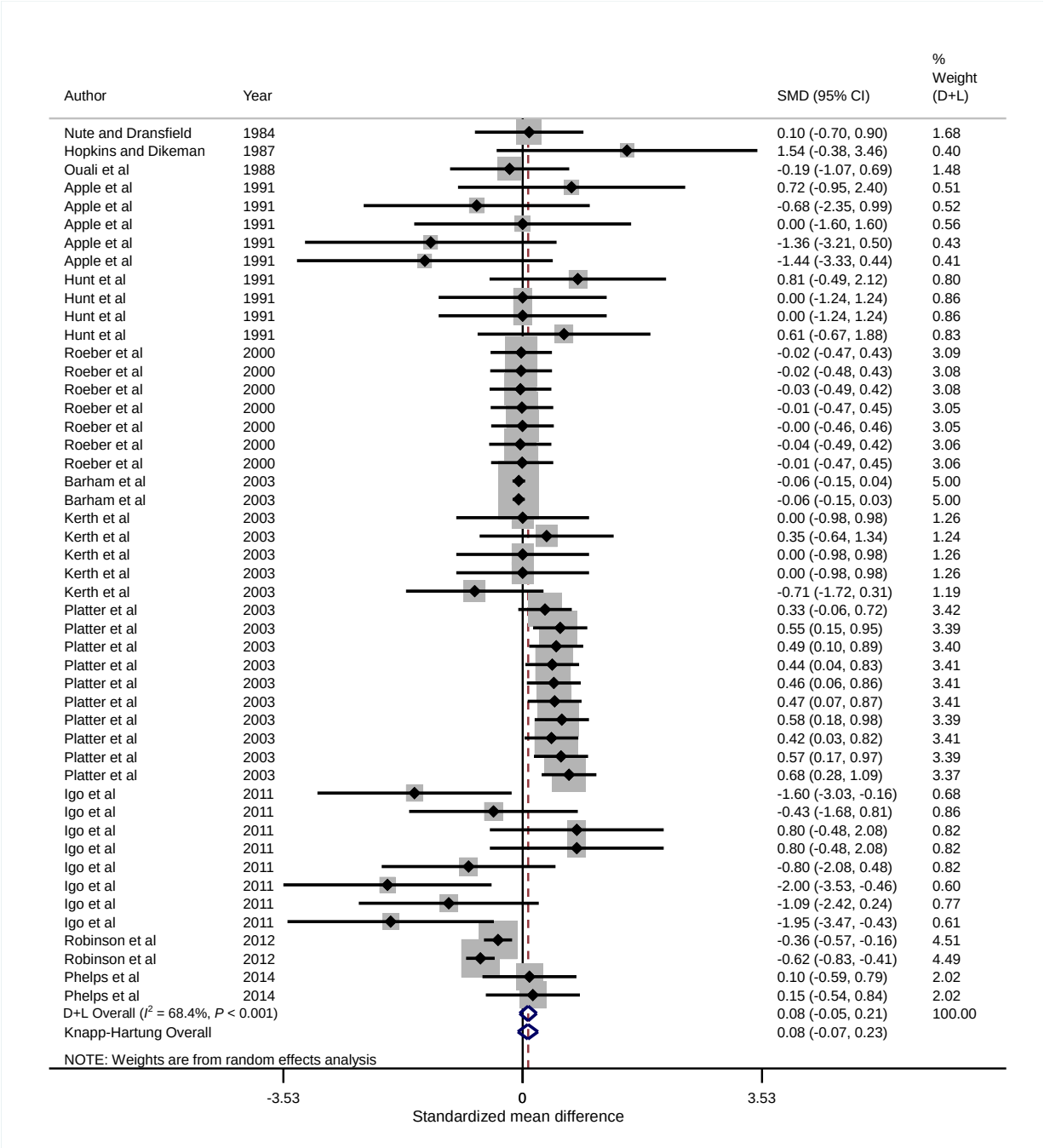

Fig 8. Forest plot of the effect size or standardized mean difference (SMD; standardized using the z-statistic) and 95% CI of the effect of hormonal growth prototants on flavour of primarily the Longissimus dorsi muscle in beef cattle. The solid vertical line represents a mean difference of zero or no effect. Points to the left of the line represent a decrease in retained body nitrogen, while points to the right of the line indicate an increase. Each square around the point effect represents the mean effect size for that study and reflects the relative weighting of the study to the overall effect size estimate. The larger the box, the greater the study contribution to the overall estimate. The weight that each study contributed are in the right-hand column. The upper and lower limit of the line connected to the square represents the upper and lower 95% CI for the effect size. The overall pooled effects size or SMD and 95% CI pooled using the DerSimonian and Laird (D + L; DerSimonian and Laird 1986) and Knapp-Sidak-Jonkman (Knapp-Hartung; IntHout *et al.* 2014) methods for random effects models are indicated by the respective diamonds at the bottom. The heterogeneity measure, I^2 is a measure of variation beyond chance among treatments included in the meta-analysis. Flavour was moderately heterogeneous as indicated by the I^2 of 68.4%.

Table 5 provides detail on the ES estimates of the effect of HGP on meat quality outcomes. The estimates are based on Knapp-Hartung methods and provide the ES, SE and 95% CI of the ES, *P*-value, I^2 and 95% CI of I^2, and τ^2. The estimates of effect based on robust regression methods provide the ES, SE and 95% CI of the ES, *P*-value, and I^2; however, the low number of treatment comparisons and experiments available precluded evaluation based on robust regression of the effects on connective tissue and CMQ4. Of the outcomes investigated only WBSF and CMQ4 were significantly affected by HGP treatment. The WMD of WBSF was 0.248 kg with a 95% CI of 0.203 to 0.292. The estimates of effect were similar for the Knapp-Hartung and robust models for WBSF (Table 5). The estimates of I^2 for all meat quality outcomes were all moderate to substantial and the 95% CI indicated that all estimates had significant heterogeneity associated with treatment, but estimates of τ^2 were low, almost all being close to or below 0.1, indicating that there was considerable variance in response that is not explained by the true effects.

Table 5. Effect size estimates of the effect of hormonal growth promotants on meat quality outcomes. The estimates are based on Knapp-Hartung methods (KH) and provide effect size (ES), SE and 95% CI of the ES, *P*-value, and measures of heterogeneity I^2 and tau^2 (τ^2). Estimates based on robust regression methods (robust) provide the ES, SE and 95% CI of the ES, and *P*-value. Treatment and experiment numbers were too small to evaluate robust regression results for the amount of connective tissue or CMQ4

Variable	ES	SE	95% CI	*P*-value	I^2, % (95% CI)	τ^2
WBSF[1], kg (KH)	0.299	0.027	0.246 - 0.352	0.001	47.3 (37-56)	0.046
WBSF[1], kg (robust)	0.306	0.053	0.181 - 0.431	0.001		0.001
Juiciness (KH)	-0.038	0.075	-0.189 - 0.112	0.610	66.5 (56-75)	0.102
Juiciness (robust)	-0.115	0.137	-0.424 - 0.193	0.421		0.001
Tenderness (KH)	-0.094	0.101	-0.296 - 0.109	0.360	78.3 (72-83)	0.129
Tenderness (robust)	-0.223	0.219	-0.717 - 0.270	0.333		0.001
Flavour (KH)	0.077	0.074	-0.071 - 0.226	0.301	68.4 (57-77)	0.101
Flavour (robust)	-0.003	0.177	-0.426 - 0.418	0.983		0.001
Connective tissue (KH)	-0.060	0.207	-0.502 - 0.382	0.776	34.1 (0-64)	0.215
CMQ4 (KH)	-0.490	0.107	-0.737 - -0.243	0.002	81.5 (66-90)	0.075

[1]Warner-Bratzler shear force

Univariable analyses were conducted using Knapp-Hartung methods to evaluate the association of potential effect modifiers with meat quality outcomes (Table 6 – 10). For WBSF, only use of multiple implants was significantly associated and had an R^2 of 18.1%. The heterogeneity for this remained

high, as was the case for the other variables examined in meta-regression. The estimates of ES of the increase in WBSF with HGP treatment were very similar for the Knapp-Hartung and Robust regression models, indicating that controlling for experiment and treatment had little effect on point estimates of effect, but *P*-values were lower and CI wider for the 3-level model, that is the robust model. A single HGP implant had more limited effects on WBSF (0.176 kg, 95% CI 0.109 to 0.242 kg). Ageing of meat (Knapp-Hartung *P* = 0.105 and robust regression *P* = 0.292) was not associated with WBSF (Table 6; Figure 9).

Table 6. Meta-regression estimates (univariable analyses) for the effects of length of time that meat was aged before evaluation, length of time that cattle were fed, use of multiple hormonal growth promotant implants (yes or no), treatments using trenbolone acetate (yes or no), breed of cattle, sex of cattle, and electrical stimulation of the carcass on Warner-Bratzler shear force responses. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2. Estimates based on robust regression methods (robust) provide the ES, SE and 95% CI of the ES, and *P*-value

Variable	ES	SE	95% CI			*P*-value	R^2	I^2, %	τ^2
Ageing of the meat, d (KH)	-0.005	0.003	-0.010	-	0.001	0.105	-0.55	46.7	0.043
Ageing of the meat, d (robust)	-0.005	0.004	-0.017	-	0.006	0.292			0.001
Length of feeding, d (KH)	0.001	0.0006	-0.0002	-	0.002	0.125	-5.06	39.9	0.035
Length of feeding, d (robust)	0.001	0.0007	-0.0004	-	0.003	0.120			0.001
Multiple implants, % of studies (KH)	0.196	0.051	0.095	-	0.296	0.001	18.1	46.9	0.036
Multiple implants, % of studies (robust)	0.172	0.125	-0.124	-	0.468	0.212			0.001
Trenbolone acetate, % of studies (KH)	-0.100	0.077	-0.252	-	0.052	0.196	-3.17	47.1	0.045
Trenbolone acetate, % of studies (robust)	-0.061	0.101	-0.276	-	0.154	0.554			0.001
Breed[1] (reference British, British cross, European, and Holstein)									
Brahman and Brahman crosses	-0.017	0.064	-0.144	-	0.110	0.789	4.39	42.7	0.042
Crossbred (undescribed)	0.189	0.087	0.018	-	0.360	0.031			
Not stated	0.423	0.217	-0.006	-	0.853	0.053			
Sex[1] (reference steers)									
Bull	0.289	0.186	-0.077	-	0.656	0.121	9.21	44.3	0.040
Heifer	-0.084	0.055	-0.193		0.024	0.127			
Mixed	0.082	0.115	-0.145	-	0.308	0.477			
Stimulation (reference not stimulated)									
Stimulated (KH)	0.059	0.090	-0.119	-	0.238	0.512	4.08	47.9	0.042
Not stated (KH)	0.197	0.094	0.012	-	0.383	0.037			

[1]The distribution of data leads to small degrees of freedom for sex and breed, resulting in unreliable *P*-values for the robust regression.

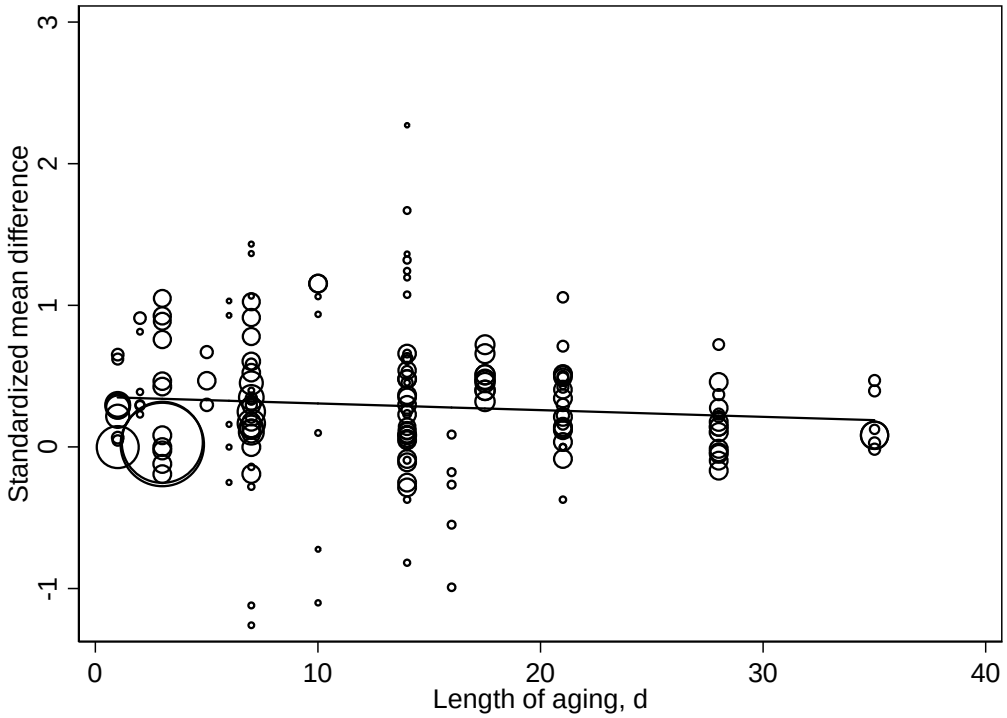

Fig 9. Standardized mean difference (SMD) between reference and hormonal growth promotant treatment for Warner-Bratzler shear force of primarily the Longissimus dorsi muscle in beef cattle with increasing length of ageing of meat in beef cattle.

The tenderness of the meat (Table 7), as assessed by consumer panels, was evaluated using different scoring systems. The only variable that was significantly associated with tenderness was the use of multiple implants that increased tenderness compared to a single implant (ES = 0.468). Treatments using crossbreds of undescribed breed and unstated breed treatments had more tender outcomes than those using British, British breed cross, European, and Holstein, cattle. The limited number of bull treatments tended to produce meat assessed as more tender. All the results had substantial heterogeneity with estimates of I^2 being all >60%. The τ^2 were moderately low (<0.3), indicating that the remaining heterogeneity was substantial and influenced by factors other than the true effects.

Table 7. Meta-regression estimates for the effects of length of time that meat was aged before evaluation, length of time that cattle were fed, use of multiple hormonal growth promotant implants (yes or no), treatments using trenbolone acetate (yes or no), breed of cattle, sex of cattle, and electrical stimulation of the carcass on tenderness responses. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2

Variable	ES	SE	95% CI			*P*-value	R^2	I^2, %	τ^2
Ageing of the meat, d	0.011	0.014	-0.167	-	0.038	0.435	0.10	78.11	0.273
Length of feeding, d	-0.001	0.002	-0.005	-	0.005	0.872	-18.6	65.3	0.277
Multiple implants, % of studies	0.468	0.182	0.104	-	0.832	0.013	41.46	71.34	0.16
Trenbolone acetate, % of studies	0.364	0.246	-0.129	-	0.858	0.145	7.06	78.43	0.254
Breed (reference British, British cross, European, and Holstein)									
Brahman and Brahman crosses	-0.211	0.182	-0.576	-	0.154	0.252	68.21	73.03	0.087
Crossbred (undescribed)	0.537	0.177	0.181	-	0.892	0.004			
Not stated	-1.167	0.547	-2.083	-	-0.251	0.014			
Sex (reference steers)									
Bull	0.974	0.493	-0.013	-	1.962	0.053	0.55	76.0	0.272
Heifer	0.068	0.349	-0.630	-	0.767	0.845			
Mixed	-0.390	0.447	-1.29	-	0.505	0.386			
Stimulation (reference not stimulated)									
Stimulated	-0.341	0.235	-0.812	-	0.129	0.151	55.02	72.25	0.123
Not stated	0.371	0.192	-0.141	-	0.756	0.059			

Juiciness of the meat (Table 8) was associated with multiple implant use (*P* = 0.008; R^2 = 56%) compared to use of a single implant; however, the overall effect was to restore juiciness towards the level of the reference group (Figure 10). Treatments using cross-bred cattle with no description of the cross were juicier than the British breed category. There was marked heterogeneity in all the results with estimates of I^2 being moderate to substantial; all were >50%. Again, the τ^2 were low (<0.05), indicating that the remaining heterogeneity was substantial and influenced by factors other than the true effects.

Table 8. Meta-regression estimates for the association of length of time that meat was aged before evaluation, length of time that cattle were fed, use of multiple hormonal growth promotant implants (yes or no), treatments using trenbolone acetate (yes or no), breed of cattle, sex of cattle, and electrical stimulation of the carcass on juiciness responses. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2

Variable	ES	SE	95% CI	*P*-value	R^2	I^2, %	τ^2
Ageing of the meat, d	0.013	0.009	-0.006 - 0.031	0.179	6.2	65.7	0.096
Length of feeding, d	0.001	0.0006	-0.0003 - 0.002	0.135	100.0	50.7	0.001
Multiple implants, % of studies	0.348	0.126	0.096 - 0.600	0.008	54.5	61.2	0.044
Trenbolone acetate, % of studies	0.134	0.185	-0.237 - 0.504	0.473	2.58	66.7	0.099
Breed (reference British, British cross, European, and Holstein)							
Brahman and Brahman crosses	-0.065	0.127	-0.321 - 0.190	0.611	73.5	54.8	0.027
Crossbred (undescribed)	0.513	0.132	0.248 - 0.778	0.001			
Not stated	-0.455	0.355	-1.167 - 0.257	0.206			
Sex (reference steers)							
Bull	0.425	0.502	-0.580 - 1.430	0.400	8.89	64.3	0.093
Heifer	-0.178	0.293	-0.765 - 0.409	0.546			
Mixed	-0.351	0.294	-0.941 - 0.308	0.239			
Stimulation (reference not stimulated)							
Stimulated	-0.117	0.168	-0.454 - 0.238	0.487	58.6	62.3	0.042
Not stated	0.325	0.143	0.012 - 0.039	0.027			

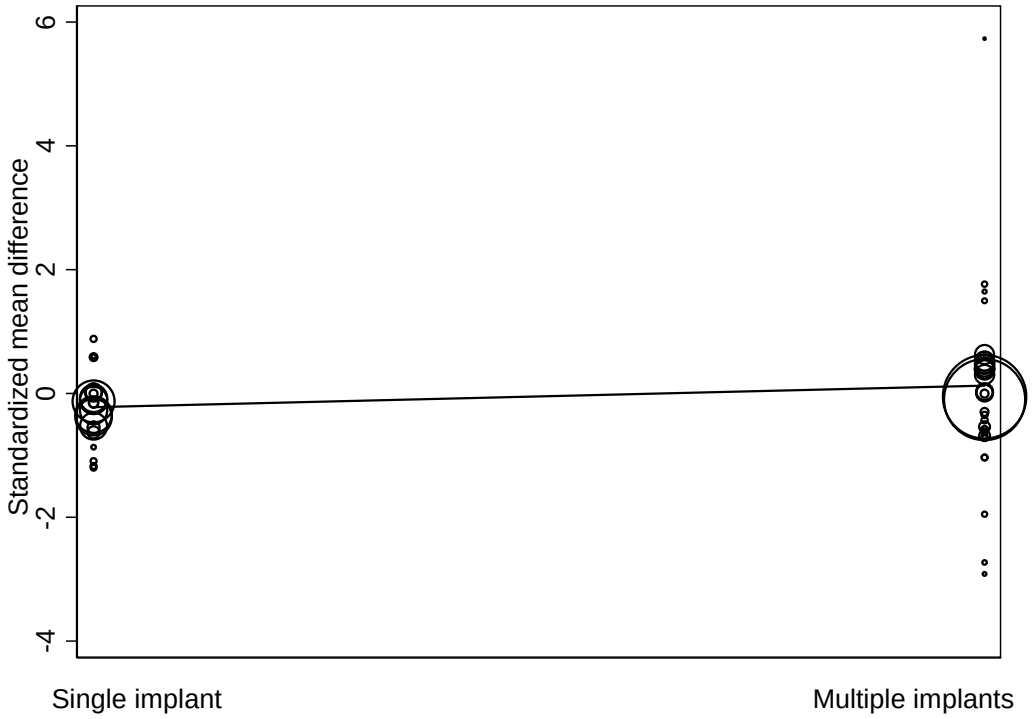

Fig 10. Standardized mean difference (SMD) between reference and hormonal growth promotant treatment for juiciness of primarily the Longissimus dorsi muscle in beef cattle implanted with single or multiple hormonal growth promotants.

While there was no significant association for the effect of treatments on flavour measures, there was a number of significant meta-regression effects (Table 9). Ageing of the meat was associated with higher flavour (R^2 = 51%) as was use of multiple implants (R^2 = 46%); however, the I^2 for these interventions were high (>50%). The mixed sex treatments were associated with less flavour than the steers. Differences in flavour were present between breeds with crossbred cattle being associated with more flavour than the British breed treatments. There was increased flavour in the treatments applied to cattle with no stated stimulation compared with those not stimulated. Again, estimates of I^2 were moderate to substantial, with the exception of breed that was moderate. Estimates of τ^2 were low.

Table 9. Meta-regression estimates for the association of length of time that meat was aged before evaluation, length of time that cattle were fed, use of multiple hormonal growth promotant implants (yes or no), treatments using trenbolone acetate (yes or no), breed of cattle, sex of cattle, and electrical stimulation of the carcass on flavour responses. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2.

Variable	ES	SE	95% CI			*P*-value	R^2	I^2, %	τ^2
Ageing of the meat, d	0.036	0.011	0.013	-	0.059	0.003	51.08	59.11	0.049
Length of feeding, d	-0.0004	0.002	-0.005	-	0.005	0.872	-18.60	55.3	0.277
Multiple implants, % of studies	0.436	0.141	0.151	-	0.722	0.004	45.89	59.79	0.055
Trenbolone acetate, % of studies	-0.023	0.229	-0.485	-	0.439	0.920	-5.28	68.98	0.107
Breed (reference British, British cross, European, and Holstein)									
Brahman and Brahman crosses	-0.158	0.114	-0.388	-	0.073	0.175	81.65	37.24	0.019
Crossbred (undescribed)	0.577	0.114	0.348	-	0.807	0.001			
Not stated	0.203	0.286	-0.373	-	0.780	0.481			
Sex (reference steers)									
Bull	0.369	0.495	-0.629	-	1.36	0.460	52.28	57.26	0.048
Heifer	-0.223	0.287	-0.802	-	0.357	0.443			
Mixed	-0.651	0.208	-1.070	-	-2.233	0.003			
Stimulation (reference not stimulated)									
Stimulated	-0.344	0.462	-1.274	-	0.585	0.460	45.20	63.09	0.055
Not stated	0.385	0.131	0.121	-	0.649	0.005			

There were limited observations (*n* = 16 treatments) on the effects of HGP on connective tissue content of meat and none of the meta-regression effects studied were significant (Table 10). For CMQ4, there were limited experiments (*n* = 9 treatments) and meta-regressions were not explored. The WMD for CMQ4 was -5.52 (95% CI = -7.94 to – 3.10).

Table 10. Meta-regression estimates for the association of length of time that meat was aged before evaluation, length of time that cattle were fed, use of multiple hormonal growth promotant implants (yes or no), and treatments using trenbolone acetate (yes or no) on connective tissue responses. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2

Variable	ES	SE	95% CI			*P*-value	R^2	I^2, %	τ^2
Ageing of meat, d	0.005	0.377	-0.076	-	0.086	0.900	-28.80	38.53	0.277
Length of feeding, d	-0.009	0.006	-0.021	-	0.003	0.115	6.8	28.3	0.200
Multiple implants, % of studies	0.729	0.611	-0.582	-	2.040	0.253	1.89	33.43	0.211
Trenbolone acetate, % of studies	0.063	0.436	-0.872	-	0.998	0.887	-16.77	38.31	0.251

These results were further investigated in multivariable models using Knapp-Hartung, permutation and robust analysis methods. In Table 11, the results of these analyses are provided for WBSF. The *P*–values for the Knapp-Hartung meta-regressions are provided as results of the permutation analyses (Harbord and Higgins 2008). These models show that the use of multiple implants was associated with a higher WBSF and that the treatments that did not include a description of electrical stimulation were associated with a higher WBSF than those that reported no stimulation. The relatively low number of experiments reporting other meat quality metrics precluded multivariable analysis.

Table 11. Multivariable meta-regression models examining effect size (ES) estimates of the effect of hormonal growth promotants on Warner-Bratzler shear force. The estimates are based on Knapp-Hartung methods and provide the effect size (ES), SE and 95% CI of the ES, significance (*P*-value), model fit (R^2), and measures of heterogeneity I^2 and τ^2. Estimates of significance (*P*-value) are provided from permutation analysis and the estimates are compared with those based on robust regression methods that provide the ES, SE and 95% CI of the ES, and I^2

Variable	ES	SE	95% CI			*P*-value	R^2	I^2, %	τ^2
Multiple implants, % of studies (KH)	0.215	0.051	0.114	-	0.315	0.001	20.4	47.4	0.035
Multiple implants, % of studies (robust)	0.188	0.124	-0.105	-	0.481	0.206			0.001
Stimulation (reference not stimulated)									
Stimulated (KH)	0.084	0.088	-0.089	-	0.257	0.654			
Not stated (KH)	0.237	0.092	0.057	-	0.419	0.035			
Stimulated (robust)	0.146	0.161	-0.268	-	0.560	0.407			
Not stated (robust)	0.271	0.190	-0.199	-	0.741	0.206			

5 Discussion

There were sufficient experiments and treatments to provide a rigorous evaluation of the effects of HGP treatment on WBSF. Our experiment refers almost exclusively to the effects of HGP on LD, with only two treatments using semitendinosis and the evidence base for muscles other than LD would have been considerably smaller. However, LD differs from other muscles in terms of ageing as Gruber *et al.* (2006) found a large ageing response for LD of 2.5 kg in muscles obtained from USDA Select grade carcasses and 2.0 kg lower shear force for muscles from USDA Choice carcasses from ageing for 26 and 15 d, respectively. These were the greatest improvements in tenderness of any of the 17 muscles tested for change in tenderness with ageing for the respective carcass quality categories.

There were essentially two approaches taken to the analysis of these data. The results of classical meta-analysis, with a random effect of experiment, are provided and meta-regression methods are used to explore the heterogeneity in the ES using Knapp-Hartung and permutation methods. The second robust method contains the random effect of experiment and treatment, and while it is possible to explore other variables using meta-regression, there were no factors that were significant in this model used to examine variability in WBSF. The two methods are included to provide a less conservative, but more informative evaluation of effects that may modify the response in WBSF with HGP treatments using the Knapp-Hartung and permutation model.

The ES for the effect of HGP on WBSF obtained from the Knapp-Hartung and robust regression are very similar and both significant, showing an increase of approximately 0.30 ES (Table 6) with a WMD of 0.25 kg of force between HGP treated and reference cattle. This increase is consistent with the

estimates of effect for HGP treatment on WBSF (WMD = 0.27 kg) derived by Watson (2008) with fewer experiments and treatments.

It has been proposed that ageing can reduce the effects of HGP on WBSF (Thompson et al., 2008). Some experiments support this finding (Igo *et al.* 2011; Packer *et al.* In Press; Schneider *et al.* 2007; Thompson *et al.* 2008), while others did not (Platter *et al.* 2003), and many experiments did not explicitly examine the effect of ageing on the WBSF response to HGP. There was limited evidence to support a diminished effect of the HGP on WBSF from this experiment (Table 6; Figure 9). However, the non-significant point effect of ageing on ES was -0.005 ES per d or -0.15 ES over 30 d; representing half the overall effect of HGP on WBSF, but ageing explained little of the overall variance in ES. The largest experiments had relatively short ageing periods. The non-significant difference in WBSF of -0.15 ES from 30 d of ageing between treatments and references estimated in this experiment is much smaller than the effect on WBSF of ageing alone over 15 and 26 d of 2.0 to 2.5 kg less force in LD reported by Gruber *et al.* (2006).

The effect of multiple implants in increasing WBSF has been consistently reported and strongly supported in this experiment. The effect of multiple implants was to increase the 0.2 ES and explained 18% of the variance in treatment (Table 6). The robust regression had a similar ES (0.17), but not significant. Further investigation, indicated that the Knapp-Hartung ES for a single implant only on WBSF was 0.195 (95% CI 0.126 to 0.264; $P < 0.001$) and had a lower heterogeneity ($I^2 = 28.9$) and very low $\tau^2 = 0.03$, suggesting that these responses were relatively consistent across treatments. Further, evaluation of the effect of a single HGP implant on WBSF using the robust regression model provided an estimate of ES of 0.219 (95% CI -0.010 to 0.447; $P = 0.06$). These results indicate that a single implant, whether this be a single agent or combination has a more limited effect on WBSF than multiple treatments.

It has been suggested that TBA may have a greater effect on increasing WBSF than other HGP treatments (Gerken *et al.* 1995; Packer *et al.* In Press). There are few experiments that test this hypothesis with single treatments, as many TBA treatments are conducted with combined TBA and estrogen treatments. Gerken *et al.* (1995) using 6 cloned steers per group found no significant difference in WBSF between treatment with a single estrogenic implant, containing 20 mg of estradiol benzoate and 200 mg of progesterone (Synovex-S) to a single androgenic implant, containing 140 mg of TBA (Finaplex). However, in our experiment, the point effect was towards TBA, associated with a reduced WBSF and the effect was not significant. The TBA implants were used in 81% of treatments either as a single, or more typically, as a combined HGP. Descriptions of the large number of different HGP products used in experiments were not always definitively provided and it was not assumed that product equivalency existed for different formulations with similar active agents. Consequently, a specific analysis for the different TBA products used was not indicated.

The evidence base for this experiment is a little unusual, because there was considerable variation in the experimental designs used. Most experiments had multiple treatment comparisons, with Schneider *et al.* (2007) containing 55 treatment comparisons. Fifty percent of treatments used more than one implant; some treatments used up to 5 implants. Experiments represented a wide range in productivity and diet composition, some reflecting feedlot practice, and some extensive pasture-based production. Further, the treatments were conducted, primarily using British and European Breeds (71%) and 18% were on Brahman and Brahman cross cattle and mostly on steers (55%) or heifers (36%). Some experiments were conducted at the pen level (Foutz *et al.* 1997; Igo *et al.* 2011; Kerth *et al.* 2003), whereas others were conducted with individual cattle as the unit of interest (Barham *et al.* 2003; Cafe *et al.* 2010; Packer *et al.* In Press). This variation in experimental design was reflected in the variance attributable to treatment within experiment being 44% of the total variance. Other meta-analytical experiments found the variance attributable to treatment level was much lower, in the order of 3 to 6% (Lean *et al.* In Press). The τ^2 representing the variance in the ES were

low, rarely exceeding 0.2 and often <0.1, but the heterogeneity attributable to random sampling errors are high, almost all with I^2 > 50 (Tables 5 to 11). The considerable variation in experimental design suggested a need for caution in interpretation of meta-regression results, such as those for TBA, because confounding of HGP treatment effects with breed, sex, or stimulation of carcass was present for single implant TBA data. However, evaluation of these TBA results controlling for the effect of breed, provided no evidence that the estimates were affected by breed 'British' or 'Brahman' and that TBA use was not associated with a higher WBSF than other HGP interventions (data not shown). There was little evidence to support breed or sex differences in modifying the effect of HGP on WBSF, with the possible exception of treatments using undescribed cross-bred cattle (Table 6). However, this effect was not present in the robust regression (results not shown). Similarly, the treatments that did not report whether electrical stimulation of the carcass was used differed for WBSF to the unstimulated studies (ES = 0.2), but only for the Knapp-Hartung and permutation model. There were few experiments represented by the undescribed cross-breds (*n* = 3) and while 19 experiments with unstated stimulation categories were present, the more conservative results of the robust regression models indicating no effect of cross-breds or electrical stimulation are appropriate.

The overall effect of HGP on tenderness was not significant, based on the sensory evaluation (*P* > 0.3; Table 5). These results are consistent with those presented in Table 2 of Nichols *et al.* (2002), but not with Watson (2008) who found that HGP reduced the tenderness of LD by approximately 5 units on a 100-point scale. None of these 3 quantitative evaluations use identical evidence bases, but many of the experiments used are the same. Watson (2008) converted the scales of assessment used in the original papers to provide a WMD, whereas Nichols *et al.* (2002) provided the data, but no pooled estimates of effect and this experiment evaluated ES, thus using the original data from experiments to provide the pooled estimate, albeit in z-score units.

The tenderness responses (Table 7) did not support the WBSF findings in that use of multiple implants was associated with increased tenderness by 0.47 ES. It should be noted that there are 13 less experiments in the tenderness and juiciness evaluations than for the WBSF database. Further, both ES, that is for a single implant or a multiple implant, had a negative association for HGP treatment on tenderness. It is also possible that time on feed, which differed between single (mean days on feed were 132 ± 15 d) against multiple implants (mean days on feed were 183 ± 8 d) may have influenced this result. While there are strong correlations between WBSF and tenderness scores for LD, Shackelford *et al.* (1995) discuss the variability and inconsistency in relationships between WBSF and tenderness scores. Duckett and Pratt (2014) also comment on the variability in responses between WBSF and sensory measures. Despite the strong correlations between WBSF and tenderness scores for LD, it appears that consumer assessed tenderness assessment of LD treated with HGP or not differed from WBSF assessed response.

Ageing did not influence the difference in tenderness; however, the point direction was to increased tenderness. Undescribed breed crosses were associated with more tenderness than 'British' cattle and 'not stated breed' were associated with being less tender than British cattle. Bulls were present in a very low number of experiments (*n* = 4), but tended (*P* = 0.055) to be associated with more tenderness than steers, possibly reflecting an earlier time to slaughter or other confounding factors.

There were limited observations for juiciness which was not significantly reduced with HGP use, nor associated with increased ageing or length of feeding. The juiciness was associated with multiple implant use, and undescribed cross-bred cattle compared to 'British' cattle, a result consistent with the findings for tenderness, but not WBSF. Similarly, the use of multiple implants, undescribed cross-bred cattle compared to 'British' cattle, and treatments that did not state whether carcass stimulation occurred were associated with increased flavour of the beef. There is a pattern of improved sensory panel performance for the treatments that had these characteristics, that is multiple implant use, undescribed cross-bred cattle compared to 'British' cattle, and treatments that did not state whether

carcass stimulation occurred for tenderness, juiciness, and flavour. It is unclear if these effects have a biological basis, or whether these findings reflect confounding for these relatively sparse observations. Both tenderness and juiciness were conducted using the same evidence base of 15 experiments and 59 treatment comparisons, but flavour had less observations. It is notable; however, that ageing was associated with increased flavour, an observation with a biological basis.

There were very limited observations on connective tissue (*n* = 16 treatments) and CMQ4 (*n* = 9 treatments) responses to HGP treatment. While connective tissue content was not altered by HGP treatment (Table 5), CMQ4 was reduced by HGP treatment by 5.54.

6 Conclusions/recommendations

The responses in this experiment were similar to those of Watson (2008) in showing an association between increased WBSF with HGP treatment, but provide new insights into the effects of HGP on WBSF. It shows that use of multiple HGP implants was associated with a large increase in WBSF, but a single implant had limited effects. Ageing did not significantly alter the HGP association with increased WBSF; however, the point direction was towards a reduced effect with ageing. The experiment also provides information on other sensory aspects of meat quality. The findings on tenderness, as assessed by sensory methods differ from those of Watson (2008) as HGP treatment was not associated with reduced tenderness. Further, juiciness, flavour, and connective tissue were not associated with HGP use, whereas there was a marked 5.5-point decrease in CMQ4, albeit in limited studies. There is a need for more targeted studies on the role of HGP in influencing meat quality.

7 Key messages

- There is an increase in WBSF with HGP treatment
- Multiple hormonal implants had a greater effect on WSBF than single implants
- Ageing did not significantly alter the HGP association with increased WBSF
- Sensory measures assessed by consumers showed different associations on meat quality with HGP use compared to WBSF
- The sensory measures tenderness, juiciness, flavour, and connective tissue content were not associated with HGP use
- There was a 5.5-point decrease in CMQ4 score in a limited number of studies with HGP use
- There is a need for more studies on the role of HGP in influencing meat quality

8 Bibliography

Apple JK, Dikeman ME, Simms DD, Kuhl G (1991) Effects of synthetic hormone implants, singularly or in combinations, on performance, carcass traits, and longissimus muscle palatability of Holstein steers. *Journal of Animal Science* **69**, 4437-4448.

Barham BL, Brooks JC, Blanton JR, Herring AD, Carr MA, Kerth CR, Miller MF (2003) Effects of growth implants on consumer perceptions of meat tenderness in beef steers. *Journal of Animal Science* **81**, 3052-3056.

Boles JA, Boss DL, Neary KI, Davis KC, Tess MW (2009) Growth implants reduced tenderness of steaks from steers and heifers with different genetic potentials for growth and marbling. *Journal of Animal Science* **87**, 269-274.

Borenstein M, Higgins JPT, Hedges LV, Rothstein HR (2017) Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods* **8**, 5-18.

Cafe LM, McIntyre BL, Robinson DL, Geesink GH, Barendse W, Pethick DW, Thompson JM, Greenwood PL (2010) Production and processing studies on calpain-system gene markers for tenderness in Brahman cattle: 2. Objective meat quality. *Journal of Animal Science* **88**, 3059-3069.

Calkins CR, Clanton DC, Berg TJ, Kinder JE (1986) Growth, Carcass and Palatability Traits of Intact Males and Steers Implanted with Zeranol or Estradiol Early and throughout Life. *Journal of Animal Science* **62**, 625-631.

Capper JL, Hayes DJ (2012) The environmental and economic impact of removing growth-enhancing technologies from US beef production. *Journal of Animal Science* **90**, 3527-3537.

Cheatham R, Duff G, Bailey C, Sanders S, Collier R, Marchello J, Baumgard L (2008) Effects of implant programs on performance, carcass characteristics, and lipogenic gene expression in Holstein steers. *South African Journal of Animal Science* **38**, 238-246.

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**, 177-188.

Dohoo IR, Martin SW, Stryhn H (2009) 'Veterinary Epidemiologic Research.' (VER, Incorporated: Charlottetown, Prince Edward Island, Canada)

Duckett S, Pratt S (2014) Meat science and muscle biology symposium—anabolic implants and meat quality. *Journal of Animal Science* **92**, 3-9.

Ebarb SM, Drouillard JS*, et al.* (2016) Effect of growth-promoting technologies on Longissimus lumborum muscle fiber morphometrics, collagen solubility, and cooked meat tenderness. *Journal of Animal Science* **94**, 869-881.

Ebarb SM, Phelps KJ*, et al.* (2017) Effects of anabolic implants and ractopamine-HCl on muscle fiber morphometrics, collagen solubility, and tenderness of beef Longissimus lumborum steaks. *Journal of Animal Science* **95**, 1219-1231.

Egger M, Smith GD (2001) Principles of and procedures for systematic reviews. In 'Systematic reviews in health care meta-analysis in context. Vol. 23-42.' (Eds M Egger, GD Smith and AD G.). (British medical journal books: London)

Foutz CP, Dolezal HG, Gardner TL, Gill DR, Hensley JL, Morgan JB (1997) Anabolic implant effects on steer performance, carcass traits, subprimal yields, and longissimus muscle properties. *Journal of Animal Science* **75**, 1256-1265.

Gerken C, Tatum J, Morgan J, Smith G (1995) Use of genetically identical (clone) steers to determine the effects of estrogenic and androgenic implants on beef quality and palatability characteristics. *Journal of Animal Science* **73**, 3317-3324.

Golder HM, Lean IJ (2016) A meta-analysis of lasalocid effects on rumen measures, beef and dairy performance, and carcass traits in cattle. *Journal of Animal Science* **94**, 306-326.

Gruber S, Tatum J, Scanga J, Chapman P, Smith G, Belk K (2006) Effects of postmortem aging and USDA quality grade on Warner-Bratzler shear force values of seventeen individual beef muscles. *Journal of Animal Science* **84**, 3387-3396.

Harbord RM, Higgins J (2008) Meta-regression in Stata. *Meta* **8**, 493-519.

Harbord RM, Steichen TJ (2004) '"METAREG: Stata module to perform meta-analysis regression," Statistical Software Components S4446201.' (Boston College, Department of Economics, Revised 05 Jan 2009.)

Hedges LV, Tipton E, Johnson MC (2010) Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* **1**, 39-65.

Higgins JPT, Green S (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. In. ' (The Cochrane Collaboration)

Higgins JPT, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* **21**, 1539-1558.

Higgins JPT, Thompson SG (2004) Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* **23**, 1663-1682.

Hopkins TD, Dikeman ME (1987) Effects of oestradiol-17β implantation on performance, carcass traits, meat sensory traits and endocrine aspects of bulls and steers. *Meat Science* **21**, 51-65.

Hunt DW, Henricks DM, Skelley GC, Grimes LW (1991) Use of trenbolone acetate and estradiol in intact and castrate male cattle: effects on growth, serum hormones, and carcass characteristics. *Journal of Animal Science* **69**, 2452-2462.

Hunter R (2010) Hormonal growth promotant use in the Australian beef industry. *Animal Production Science* **50**, 637-659.

Hunter RA, Burrow HM, McCrabb GJ (2001) Sustained growth promotion, carcass and meat quality of steers slaughtered at three liveweights. *Australian Journal of Experimental Agriculture* **41**, 1033-1040.

Hunter RA, Magner T, Allingham PG (2000) Sustained growth promotion, carcass characteristics, and meat quality of steers treated with oestradiol-17β. *Australian Journal of Agricultural Research* **51**, 133-138.

Igo J, Brooks J, Johnson B, Starkey J, Rathmann R, Garmyn A, Nichols W, Hutcheson J, Miller M (2011) Characterization of estrogen-trenbolone acetate implants on tenderness and consumer acceptability of beef under the effect of 2 aging times. *Journal of Animal Science* **89**, 792-797.

IntHout J, Ioannidis JP, Borm GF (2014) The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology* **14**, 25.

Ioannidis JPA, Patsopoulos NA, Evangelou E (2007) Uncertainty in heterogeneity estimates in meta-analyses. *BMJ* **335**, 914-916.

Kerth CR, Montgomery JL, Morrow KJ, Galyean ML, Miller MF (2003) Protein turnover and sensory traits of longissimus muscle from implanted and nonimplanted heifers. *Journal of Animal Science* **81**, 1728-1735.

Knapp G, Hartung J (2003) Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine* **22**, 2693-2710.

Lean IJ, De Ondarza MB, Sniffen CJ, Santos JEP, Griswold K (In Press) Meta-analysis to predict the effects of metabolizable amino acids on dairy cattle performance. *Journal of Dairy Science*.

Lean IJ, Rabiee AR, Duffield TF, Dohoo IR (2009) Invited review: Use of meta-analysis in animal health and reproduction: Methods and applications. *Journal of Dairy Science* **92**, 3545-3565.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PG (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine* **6**, e1000097.

Nichols WT, Galyean ML, Thomson DU, Hutcheson JP (2002) Effects of Steroid Implants on the Tenderness of Beef. *The Professional Animal Scientist* **18**, 202-210.

Nute G, Dransfield E (1984) The quality of sirloin from zeranol implanted steers. *International Journal of Food Science & Technology* **19**, 21-28.

Ouali A, Zabari M, Renou JP, Touraille C, Kopp J, Bonnet M, Valin C (1988) Anabolic agents in beef production: Effects on muscle traits and meat quality. *Meat Science* **24**, 151-161.

Packer DT, Geesink GH, Polkinghorne R, Thompson JM, Ball AJ (In Press) The impact of two different hormonal growth promotants (HGPs) on the eating quality of feedlot finished steer carcasses. *Animal Production Science*.

Phelps KJ, Drouillard JS*, et al.* (2014) Effects of the Programmed Nutrition Beef Program on meat quality characteristics. *Journal of Animal Science* **92**, 1780-1791.

Platter W, Tatum J, Belk K, Scanga J, Smith G (2003) Effects of repetitive use of hormonal implants on beef carcass quality, tenderness, and consumer ratings of beef palatability. *Journal of Animal Science* **81**, 984-996.

Rabe-Hesketh S, Skrondal A (2005) 'Multilevel and longitudinal modeling using Stata.' (STATA press: College Station, TX)

Rabiee AR, Breinhild K, Scott W, Golder HM, Block E, Lean IJ (2012) Effect of fat additions to diets of dairy cattle on milk production and components: A meta-analysis and meta-regression. *Journal of Dairy Science* **95**, 3225-3247.

Reiling BA, Johnson DD (2003) Effects of implant regimens (trenbolone acetate-estradiol administered alone or in combination with zeranol) and vitamin D3 on fresh beef color and quality. *Journal of Animal Science* **81**, 135-142.

Robinson DL, Cafe LM, McIntyre BL, Geesink GH, Barendse W, Pethick DW, Thompson JM, Polkinghorne R, Greenwood PL (2012) Production and processing studies on calpain-system gene markers for beef tenderness: Consumer assessments of eating quality. *Journal of Animal Science* **90**, 2850-2860.

Roeber DL, Cannell RC, Belk KE, Miller RK, Tatum JD, Smith GC (2000) Implant strategies during feeding: impact on carcass grades and consumer acceptability. *Journal of Animal Science* **78**, 1867-1874.

Rumsey TS, Elsasser TH, Kahl S, Solomon MB (1999) The effect of roasted soybeans in the diet of feedlot steers and Synovex-S ear implants on carcass characteristics and estimated composition. *Journal of Animal Science* **77**, 1726-1734.

Samber JA, Tatum JD, Wray MI, Nichols WT, Morgan JB, Smith GC (1996) Implant program effects on performance and carcass quality of steer calves finished for 212 days. *Journal of Animal Science* **74**, 1470-1476.

Scheffler JM, Buskirk DD, Rust SR, Cowley JD, Doumit ME (2003) Effect of repeated administration of combination trenbolone acetate and estradiol implants on growth, carcass traits, and beef quality of long-fed Holstein steers. *Journal of Animal Science* **81**, 2395-2400.

Schneider BA, Tatum J, Engle T, Bryant T (2007) Effects of heifer finishing implants on beef carcass traits and longissimus tenderness. *Journal of Animal Science* **85**, 2019-2030.

Shackelford SD, Savell JW, Crouse JD, Cross HR, Schanbacher BD, Johnson DD, Anderson ML (1992) Palatability of beef from bulls administered exogenous hormones. *Meat Science* **32**, 397-405.

Shackelford SD, Wheeler TL, Koohmaraie M (1995) Relationship between shear force and trained sensory panel tenderness ratings of 10 major muscles from *Bos indicus* and *Bos taurus* cattle. *Journal of Animal Science* **73**, 3333-3340.

St-Pierre N (2001) Invited review: Integrating quantitative findings from multiple studies using mixed model methodology. *Journal of Dairy Science* **84**, 741-755.

Sterne JA, Harbord RM (2004) Funnel plots in meta-analysis. *Stata Journal* **4**, 127-141.

Tanner-Smith EE, Tipton E (2014) Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods* **5**, 13-30.

Thompson J, McIntyre B, Tudor G, Pethick D, Polkinghorne R, Watson R (2008) Effects of hormonal growth promotants (HGP) on growth, carcass characteristics, the palatability of different muscles in the beef carcass and their interaction with aging. *Australian Journal of Experimental Agriculture* **48**, 1405-1414.

Thompson SG, Sharp SJ (1999) Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine* **18**, 2693-2708.

Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J (2013) Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods* **45**, 576-594.

Watson R (2008) Meta-analysis of the published effects of HGP use on beef palatability in steers as measured by objective and sensory testing. *Australian Journal of Experimental Agriculture* **48**, 1425-1433.

White IR, Thomas J (2005) Standardized mean differences in individually-randomized and cluster-randomized trials, with applications to meta-analysis. *Clinical Trials* **2**, 141-151.