Scotland's Rural College

A meta-analysis of cortisol concentration, vocalization, and average daily gain associated with castration in beef cattle

Canozzi, MEA; Mederos, A; Manteca, X; Turner, SP; McManus, C; Zago, D; Barcellos, JOJ

Published in:
Research in Veterinary Science

DOI:
10.1016/j.rvsc.2017.07.014

First published: 20/07/2017

Document Version
Peer reviewed version

Citation for pulished version (APA):
Canozzi, MEA., Mederos, A., Manteca, X., Turner, SP., McManus, C., Zago, D., & Barcellos, JOJ. (2017). A meta-analysis of cortisol concentration, vocalization, and average daily gain associated with castration in beef cattle. *Research in Veterinary Science*, *114*, 430 - 443. https://doi.org/10.1016/j.rvsc.2017.07.014

A meta-analysis of cortisol concentration, vocalization, and average daily gain associated with castration in beef cattle

Maria Eugênia Andrighetto Canozzi[a], America Mederos[b], Xavier Manteca[c], Simon Turner[d], Concepta McManus[e], Daniele Zago[a], and Júlio Otávio Jardim Barcellos[a, *]

[a] Department of Animal Science, Federal University of Rio Grande do Sul, Bento Gonçalves Ave., nº 7712, 91540-000, Porto Alegre, RS, Brazil.

[b] National Research Institute for Agriculture, Ruta 5 km 386, Tacuarembo 45000, Uruguay.

[c] Department of Animal and Food Science, School of Veterinary Science, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain.

[d] Animal and Veterinary Sciences Group, Scotland's Rural College, West Mains Road, Edinburgh, EH9 3JG, UK

[e] Institute of Biological Sciences, University of Brasilia, 70910-900, Brasília, DF, Brazil.

[*] Corresponding author at: Department of Animal Science, Federal University of Rio Grande do Sul, Bento Gonçalves Ave., nº 7712, 91540-000, Porto Alegre, RS, Brazil. Tel.: +55 51 3308 6042. *E-mail address:* julio.barcellos@ufrgs.br (J.O.J. Barcellos)

ABSTRACT

A systematic review and meta-analysis (MA) were performed to summarize all scientific evidence for the effects of castration in male beef cattle on welfare indicators based on cortisol concentration, average daily gain (ADG), and vocalization. We searched five electronic databases, conference proceedings, and experts were contacted electronically. The main inclusion criteria involved completed studies using beef cattle up to one year of age undergoing surgical and non-surgical castration that presented cortisol concentration, ADG, or vocalization as an outcome. A random effect MA was conducted for each indicator separately with the mean of the control and treated groups. A total of 20 publications reporting 26 studies and 162 trials were included in the MA involving 1,814 cattle. Between study heterogeneity was observed when analysing cortisol (I^2 = 56.7%) and ADG (I^2 = 79.6%). Surgical and non-surgical castration without drug administration compared to uncastrated animals showed no change ($P \geq 0.05$) in cortisol level. Multimodal therapy for pain did not decrease ($P \geq 0.05$) cortisol concentration after 30 min when non-surgical castration was performed. Comparison between surgical castration, with and without anaesthesia, showed a tendency (P = 0.077) to decrease cortisol levels after 120 min of intervention. Non-surgical and surgical castration, performed with no pain mitigation, increased and tended to increase the ADG by 0.814 g/d (P = 0.001) and by 0.140 g/d (P =0.091), respectively, when compared to a non-castrated group. Our MA study demonstrates an inconclusive result to draw recommendations on preferred castration practices to minimize pain in beef cattle.

Keywords: analgesia; animal welfare; pain

Introduction

Castration is a common livestock management procedure throughout the world. The physical procedure is the most common approach used by farmers, although it increases cortisol concentration and changes animal behaviour (Fell et al., 1986; González et al., 2010; Roberts et al., 2015). To counteract this, a hormonal method, i.e. immunocastration, has been proposed as an alternative method for castration (Martí, 2012).

The awareness of animal pain caused by routine husbandry practices has become more common, and additional studies have been motivated to determine the role of pain relief (Stafford and Mellor, 2005; Roberts et al., 2015). Currently, understanding the effects of castration methods and their relationship with pain management have been discussed in narrative reviews, highlighting that the castration cause pain, but determining the need for analgesia, as well as the dose, route, duration and frequency of drug administration in cattle are still unclear (Stafford and Mellor, 2005; Coetzee, 2011). Additionally, due the difficulties and the variability in field research, the integration of findings from many studies, using systematic review (SR) and meta-analysis (MA) can provide an equally unbiased estimate of the treatment effect, with an increase in the precision of this estimate (Egger et al., 2001; Borenstein et al., 2009).

Lean et al. (2009) reported that a rigorously conducted MA could provide new insights into animal well-being. We conducted a SR-MA to test the hypothesis that specific methods of castration and pain relief strategies can be used to prevent or minimize the adverse effects on beef cattle welfare. The purpose of this study was to identify, evaluate, critically appraise, and synthesize the available literature reports on how the castration procedures affect beef cattle welfare using a SR-MA approach.

Material and methods

Research question and protocols

This study identified the effects of castration procedure on beef cattle welfare by measuring cortisol levels, average daily gain (ADG), or vocalization.

The literature search strategy was defined based on the main concepts in terms of PICO: population (P), intervention (I), comparator (C), and outcome (O). The population studied was exclusively beef cattle up to 12 months of age (calf or yearling), since the experience of intense pain soon after birth may “programme” the animal’s subsequent sensitivity to pain challenges (Viñuela-Fernández et al., 2007). Moreover, it may not be possible to castrate young calves in extensive beef production systems where calves may not be mustered until weaning (Stafford, 2007). The invention of interest was castration, dehorning, or disbudding. The present study only shows findings on castration intervention; however, the literature search was conducted to also include other two procedures, dehorning and disbudding, as presented in the flow diagram (Fig. 1). Similar groups of animals undergoing the same procedure, with or without intervention, were considered as comparison groups. We did not exclude publications based on the type of comparison used. Outcomes of interest were vocalization, cortisol, and ADG (see Table A.1).

An *a priori* protocol was developed, and each screening tool for this study was adapted from previously available forms (Mederos et al., 2012), and pre-tested before implementation.

Search methods for identification of studies

A list of final search terms and algorithms was summarized by population, outcome, and intervention components as follows: (bovine OR “beef cattle” OR cal* OR herd) AND (disbud* OR dehorn* OR castration) AND (“animal wel*” OR “animal pain” OR “animal stress” OR cortisol OR behavio* OR vocali*). This search strategy also retrieved relevant studies of animal performance evaluation as the outcome. Therefore, ADG was not included to avoid an overload of non-relevant citations.

A systematic literature search was conducted on May 2013 and updated on May 2015 using five electronic databases – CAB Abstracts (Thomson Reuters, 1910–2015), ISI Web of Science (Thomson Reuters, 1900–2015), PubMed (MEDLINE, 1940–2015), Agricola (EBSCO, 1970–2015) and Scopus (Elsevier, 1960–2015). In addition, the following proceedings were searched for references: ADSA-ASAS Joint Annual Meeting (from 2001 to 2014) and International Society for Applied Ethology, ISAE (from 2001 to 2014). Researchers in animal welfare were contacted electronically to request unpublished data. Reference search validation was performed by searching the reference lists from four recent literature reviews (Bretschneider, 2005; Weary et al., 2006; Coetzee, 2011; Schwartzkopf-Genswein et al., 2012). All citations were imported into the reference manager RefWorks (RefWorks–COS, USA) and duplicate citations were removed manually.

Study selection criteria and relevance screening

Five reviewers contributed to the study and were trained for the relevance screening step using 30 abstracts. With that, we sought to identify potentially relevant studies, for determining relevance of the studies identified by search strategy, for critically appraising the studies, and for analysing variation among studies (Higgins and Green, 2011). The citation was considered relevant when investigating primary research; animal welfare in beef cattle; castration, dehorning, or disbudding as interventions; and measuring cortisol, vocalization, or ADG as welfare indicators. The included study designs were randomized and non-randomized clinical trials, cohort studies, and case-controls. At this stage, no limits were applied for language or publication year.

Finally, all identified citations were independently assessed for relevance by two independent reviewers using the titles and abstracts (when available). Data conflicts were determined through discussion, and an expert opinion was requested when agreement was not attained. An electronic SRSnexus review format (Möbius Analytics, Ottawa, Ontario, Canada) was used for all SR steps.

Methodological assessment and data collection process

Data extraction (DE) forms were adapted from previous work. The first author was responsible for data extraction from the eligible studies. Publications reporting more than one study design were duplicated and extracted as separate studies.

Before risk of bias assessment and DE were performed, the relevance of papers selected through abstract screening was confirmed using the full papers based on language (English, Spanish, Portuguese, or Italian); appropriate control group; sufficiently detailed results to conduct the DE and to extract quantitative data for MA. At this stage, primary research was restricted to publications in the languages in which the research team members were fluent, since translation of published articles in other languages was precluded due to financial constraints.

Information extracted from each study was divided into study population, intervention, outcome measurements, and result data. Manuscript-level information included the journal name, the author(s) name(s), the year of publication, and the original language.

Considerations for data collection and manipulation

For each outcome, we attempted to extract the mean, standard deviation (SD), or any available measure of dispersion, measurement unit, *P*-value, and the number of animals in the control and treatment groups. Cortisol and ADG data were converted to nmol/L and g/d, respectively. The included studies in our database evaluated the vocalization in the same scale, 0 to 3 (0= no vocalization; 1= snorting or grunting; 2= momentary vocalization; and 3= continuous vocalization during and immediately after testicular manipulation). These summary measures were entered into an electronic spreadsheet, and a dataset was built containing the results of controlled studies and outcomes of interest: cortisol (baseline, 20 or 30 or 40min, and 120min), ADG (during observation period), or vocalization scores (during the procedure). The early stages of pain responses, as well as the long-term pain, both induced by castration, have been assessed extensively using cortisol level (Mellor et al., 2000; Thüer et al., 2007; González et al., 2010).

For further analysis, the castration methods were stratified into three groups: 1) surgical castration including Newberry knife, knife, and scalpel blade (plus emasculator or Henderson castrating tool[1]); 2) non-surgical castration including elastrator rings, band, Callicrate bander[2], Burdizzo emasculator (plus elastrator ring), and immunocastration; and 3) surgical vs. non-surgical castration methods (comparison between non-surgical and surgical castration). The control group may be uncastrated (Group 1 and 2) or surgical (Group 1) and non-surgical (Groups 2) castration, and the treated group were always submitted to surgical (Group 1) or non-surgical (Group 2) castration. When the comparison was between two castrated groups, the intention was to compare different techniques of surgical (Group 1) and non-surgical (Group 2) castration. In addition, relevant pain mitigation was identified as analgesic-sedative (xylazine), anaesthesia (lidocaine, and combination of xylazine and ketamine), anti-inflammatory (dexamethasone, dipyrone, ketoprofen, and meloxicam), and multimodal therapy (combination of xylazine and flunixin or procaine, and lidocaine and dipyrone).

---

[1] The Henderson castration tool is clamped on each spermatic cord individually and rotated by a cordless drill approximately 20 rotations until the cord is severed.

[2] The band is applied around the scrotum proximal to the testes. The elastic band is tightened until adequate tension is achieved; a metal grommet is then crimped around the band to hold tension.

When the study reported the results in the log-transformed scales, these were transformed back to the original scale using the formula described by Mederos et al. (2012). A pooled standard deviation (Sp) was derived from the formula when an overall standard error of the mean (SEMp) was reported for the control and treatment groups (Ceballos et al., 2009; Higgins and Green, 2011; Mederos et al., 2012) as follows:

$$S_p = SEM_p \times \sqrt{n_p} \quad (1)$$

Where n_p is the number of calves in the treatment and control groups.

In studies that reported only *P*-values, an estimate of a common standard deviation was computed using the *t*-statistic and assuming the data were normally distributed, using the formula (Ceballos et al., 2009; Mederos et al., 2012):

$$S_P = \frac{(x_2 - x_1)}{t(\alpha dfE)\sqrt{(1/n_2)+(1/n_1)}} \quad (2)$$

Where $x_2 - x_1$ represents the mean difference; $t(\alpha f dE)$ is the percentile from the reference distribution, and *n* is the sample size of each group.

When results were only graphically presented, the corresponding author was contacted by electronic mail and asked to provide the summary statistics. If no response was obtained or data were not provided, the mean or measure of dispersion or both were extracted by manual measurement using a ruler. Finally, as the cortisol data were collected in three different points, the summary data were retrieved, and the effect size was computed according to recommended methodological approaches (Borenstein et al., 2009).

Quality assessment

We used standardized methods to estimate the risk of bias of the individual studies included in the MA (Higgins and Green, 2011), with one minor modification. To evaluate the domain “blinding of outcome assessment”, we considered that the vocalization was at high risk of bias if blinding was not reported and at low risk of bias if blinding was reported (Dzikamunhenga et al., 2014). This is a subjective measure and more prone to poor reliability (Weary et al., 2006). Cortisol concentration and ADG were considered to be at low risk of bias regardless of the presence or absence of blinding. Quality assessment was performed by the first author.

Meta-analysis

Studies were included in the quantitative analysis when they reported sufficient data to estimate a mean difference (MD) between control and treatment groups and a 95% confidence interval (CI). Cortisol values were obtained from baseline to 20/30/40 min and up to 120 min was analysed, while for ADG we used data collected during the follow-up period reported by the authors. For cortsiol, the the term “30 min” will be used as a general descriptor for samples collected at 20/30/40 min, since the data were scarce for independent evaluation in each time. The random effect MA and meta-regressions were conducted given *a priori* assumption of between-study heterogeneity using the DerSimonian and Laird (1986) method. All statistical analyses were performed using statistical package Stata V 14.0 (StataCorp., Texas, USA).

Comparison groups. A separate MA was conducted using various subsets of data consisting of at least two individual studies that investigated similar treatments with the same outcome. As shown by many researchers, the MA comparison groups with small number of trials are possible and the results are reliable (Mederos et al., 2012; Falzon et al., 2014; Lean et al., 2014). Concurrently, each outcome was evaluated separately as a group using stratification by castration method and pain management and a pooled MD and 95% CI were generated. Cochran’s Q (chi-square test of heterogeneity) and I^2 (percentage of total variation across studies that is due to heterogeneity rather than chance) were calculated based on the castration method and outcome. The magnitude of I^2 was interpreted in the order of 25%, 50%, and 75% and considered as low, moderate, or high heterogeneity, respectively (Higgins et al., 2003).

Publication bias

Publication bias was visually and statistically assessed using a funnel plot, Begg's adjusted rank correlation, and Egger's regression asymmetry tests for each outcome. Bias was considered based on visual plot and whether at least one of the statistical methods was found to be significant ($P < 0.10$). If there was any evidence of publication bias, we used the "trim-and-fill" method to estimate the extent of the bias as suggested by Duval and Tweedie (2000).

Meta-regression analysis

Univariable random-effects models were performed to evaluate sources of between-study heterogeneity that may influence the cortisol level and ADG as response of subjects to treatment (Borenstein et al., 2009). The variables explored in the meta-regressions were (1) randomization (no or yes), (2) cluster control (no, yes, or not applicable), (3) confounders identified and controlled (no, yes, or not applicable), (4) manuscript publication year, (5) peer-reviewed (no or yes), (6) continent (North America, South America, Europe, Asia, or Oceania), (7) cattle group (*Bos taurus taurus*, *Bos taurus indicus*, hybrid/mixed, or not reported), (8) who performed the intervention (not reported, farm staff, or veterinarian), (9) application of medicine for pain relief (no or yes), (10) type of medicine (not applicable, analgesic-sedative, anaesthesia, anti-inflammatory, or multimodal therapy), (11) method of castration (surgical, non-surgical, or surgical *vs.* non-surgical castration methods), (12) cattle age (days), (13) intervention follow-up (i.e., it is the sum of the adaptation and the experimental periods), and (14) sample size. As explained above, univariable analysis was performed due to the small number of observations available for each outcome of interest that precluded the development of multivariable analysis.

Cumulative meta-analysis and influential studies

A cumulative MA was conducted to evaluate the pooled estimate of the treatment effect each time that a new potential study was published. Those analyses are most often used to display the pattern of the evidence over time (Borenstein et al., 2009). Sensitivity analyses were performed to determine whether certain studies had a substantial effect on the MD by manually replacing and removing one study at a time and evaluating whether the effect had changed by ±30%.

Results

Study selection

Our search identified 1,267 citations, from which 102 full-text publications were assessed for eligibility, and 69 were excluded after methodological soundness and data extraction (Fig. 1). Of the remaining, 9 publications did not have enough data to perform the quantitative analysis (see Table A.2), and 20 reports on castration were included in the SR-MA (Table 1).

Numerical data were obtained from two of 20 contacted authors who presented their results graphically or without sufficient data (one from the USA and one from Uruguay). The treatment groups evaluated in this study were: surgical castration (n = 19 studies), non-surgical castration (n = 17), and surgical *vs.* non-surgical castration methods (n = 14). Relevant pain mitigation included: two studies analysing analgesic-sedative, seven evaluating anaesthesia, six evaluating anti-inflammatory and four evaluating multimodal therapy. The total number and the average age (days) of cattle included in this MA were, respectively, 402 and 134 for cortisol concentration; 1,648 and 214 for ADG; and 32 and 150 for vocalization.

In total, 18 publications[1] were included in this SR-MA that comprised 23 studies and 156 unique treatment comparisons. The results of the main characteristics of the included studies are presented in Table A.3.

Risk of bias

None of the studies provided sufficient details about the blinding of personnel and the risk of performance bias was unclear. The risk of detection bias was considered relevant only for vocalization, and none of the studies used to blind outcome assessor from knowledge of which intervention a participant received, leading to high risk of detection bias. The approach to describe the completeness of outcome data for each main outcome showed a high risk of bias in two studies that evaluated cortisol concentration (Petherick et al., 2012). Both studies showed missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across the intervention group. Several studies failed to give enough detail to assess the potential risk of bias as presented in Table A.4 and Table A.5.

Meta-analysis

One hundred sixty three trials from 26 studies were included to perform MA on cortisol concentration and ADG data. The vocalization score was the least investigated outcome, and data were presented in a manner that was not usable in the quantitative MA. There were no exclusions due to lack of randomization procedures or lack of adjusting for clustering and confounders. The number of publications, studies, trials, and type of outcome measurements available for the statistical analyses are presented in Table 2.

---

[1] One publication can report more than one study, and each study is composed by one or more trials (comparisons).

Effect of castration on cortisol concentration. The variation in the overall cortisol mean difference attributable to the heterogeneity was high (I^2 = 56.7%).

Non-surgical castration: Eight studies (n = 20 trials) which evaluated non-surgical castration were included and when no stratification by control group or by the type of pain management were performed, the overall MD was 0.108 nmol/L (95% CI: -0.305, 0.522) with high between study heterogeneity (I^2 = 80.2%; P < 0.001). A stratified analysis on trials considered castrated animals without drug administration, at 30 min and 120, revealed no significant effect on cortisol concentration and high between study heterogeneity when compared with non-castrated animals (Table 3). The multimodal therapy yielded a non-significant decrease in cortisol concentration 30 min after procedure with a moderate between study heterogeneity (n = 2 trials; I^2 = 36.2%).

Surgical castration: Combining data from the eight studies (n= 30 trials) that evaluated surgical castration presented cortisol MD of 0.122 nmol/L (95% CI: -0.104, 0.349) with moderate between study heterogeneity (I^2 = 28.2%; P= 0.077). In the stratified analysis, cattle submitted to surgical castration without pain mitigation compared to uncastrated did not show an effect on cortisol concentration, at 30 min and at 120 min, with no between study heterogeneity and moderate between study heterogeneity, respectively (Fig. 2). Studies where anaesthesia was used to perform castration did not affect the cortisol level at 120 min in comparison to surgical procedure without drug administration with no between study heterogeneity (Table 3).

Non-surgical *vs.* Surgical castration methods: There was no consistent evidence of an overall effect on the cortisol concentration (MD = 0.080 nmol/L; 95% CI: -0.153, 0.314) (n = 17 trials) with low between study heterogeneity (I^2 = 1.3%). Regardless of the time of cortisol measurement, 30 or 120 min, the stratified analyses showed no and low between study heterogeneity, respectively, and no strong evidence on difference in cortisol level when castration was performed with no pain mitigation in both groups (Table 3).

Effect of castration on ADG. There was high between studies heterogeneity (I^2 = 79.6%) for the included studies reporting ADG data.

Non-surgical castration: Pooled results from 13 studies (*n* = 27 trials) evaluating non-surgical intervention showed an increase on ADG by 0.411 g/d (95% CI: 0.009, 0.812; *P* = 0.045) with high between study heterogeneity (I^2 = 90.4%). Results from the stratified analysis presented a higher performance for non-surgical castration without drug administration when compared to the non-castrated group, with high between study heterogeneity. The use of anaesthesia and multimodal therapy had no effect on ADG when compared to uncastrated cattle, with high between study heterogeneity (Table 4).

Surgical castration: Pooled analyses across all 14 studies (*n* = 44 trials) that evaluated surgical castration showed no significant difference on ADG (*MD* = 0.133; 95% CI: -0.040, 0.306), with high between study heterogeneity (I^2 = 61.3%). Results from the stratified analysis on surgical castration with no pain mitigation reported a tendency to increase the ADG compared to uncastrated animals, with moderate between study heterogeneity. No differences were found in ADG when castration was performed with anaesthesia, anti-inflammatory, or multi-modal therapy (Table 4).

Non-surgical *vs.* Surgical castration: The comparison between non-surgical and surgical castration was reported in eleven studies (*n* = 24 trials). We observed no difference (*MD* = -0.033; 95% CI: -0.293, 0.228) with high between study heterogeneity (I^2 = 56.8%). Non-significant effect and moderate between study heterogeneity was found when both intervention, surgical and non-surgical castration, were performed without drug administration. In addition, the between study heterogeneity was high when anaesthesia, anti-inflammatory or multimodal therapy were used in the surgical group (Table 4).

Publication bias

The included studies in our MA are highly heterogeneous; therefore, the results should be carefully interpreted. There was some evidence of publication bias in studies measuring cortisol concentration, since the Begg's test was marginally significant and the random-effects "trim-and-fill" indicated that an additional 13 trials would have been necessary to remove this apparent publication bias (or other small-study effects) (Fig. 3). A symmetrical funnel plot and the statistical Egger's and Begg's tests suggested that publication bias was not likely to be present when evaluating ADG as an outcome.

Meta-regression analysis

Meta-regression results on cortisol concentration. Twelve studies (n = 67 trials) were submitted to univariate meta-regression. However, none of them contributed to explain the between study variation. Two of 14 (control of confounders, and peer-reviewed) were significantly associated with the trial effect size. Univariable meta-regression indicated that studies reporting that controlling for confounders had a predicted MD in cortisol level of 0.50 nmol/L lower than studies that did not report control for confounders (P =0.045). Meta-regression results also suggested that studies published in a non peer-reviewed journal (including conference proceedings, thesis, and government or research station report) had a marginally lower predicted value for MD (-0.34 nmol/L; P = 0.054) compared to studies published in indexed and scientific journals.

Meta-regression results on ADG. The univariable meta-regression was conducted on 20 studies (n = 96 trials). None of the variables showed a significant association with the outcome of interest.

Cumulative meta-analysis and influential studies

In the cumulative meta-analysis for cortisol concentration, there was clear evidence of change in the estimated point of the pooled treatments MD, from negative (*MD* = -0.012) to positive (*MD* =0.114), using collected data from 2009 until 2014. The sensitivity analysis showed that removing two studies (Coetzee et al., 2010; del Campo et al., 2014) decreased and increased the MD from 0.114 nmol/L to 0.069 nmol/L and 0.161 nmol/L, respectively.

No evidence for a chronological tendency was found for the ADG outcome. The pooled estimate for the effects of castration on ADG showed a reduction from 0.176 g/d to 0.114 g/d (P = 0.134; 95% CI -0.035, 0.263) by removing only one study (Whitlock et al., 2013).

Discussion

There is a clear consensus about the moral and ethical treatment of animals undergoing painful procedures. For cattle farmers castration of their animals is an unpleasant but necessary husbandry procedure, that improves beef quality with increased marbling and tenderness, while for the general public it is considered an unnecessarily painful procedure (AVMA, 2009; Stafford and Mellor, 2010).

Unfortunately, literature is discordant and not conclusive on what recommendation should be transferred to farmers and practitioners. Due to the relevance of this topic for the beef supply chain, i.e. from farmers to consumers, we intended to synthesize the research knowledge available on this topic, using a meta-analytic approach. Despite the large number of studies identified using the methodology described above, studies providing data in a suitable manner to allow for a broad quantitative analysis was lower than expected.

Most publications suitable to be included in this MA were published in the 2000s. The development of methods of recognition, assessment, and management of animal pain has increased in the last 15 years. Also, in the 2000s, the proper management of pain in foodproducing animals became a matter of increasing public concern and growing interest leading to new legislation worldwide (Weary et al., 2006) regarding the emergence of castration as a painful procedure. Among the 18 publications, only one was non-English and more than half were conducted in North America.

Although reporting guidelines for randomized controlled trials already have been published (Sargeant et al., 2005), we detected unsuccessful report data for sample size justification, random sequence generation, and blinding. As presented above, there was a variable risk of bias for the outcomes of the included studies, which is a common feature reported on many of the published meta-analysis on livestock (Falzon et al., 2014; Golder and Lean, 2016).

The effects of castration on vocalization

The quantitative synthesis of approaches for measuring vocalization was not suitable for this study. Although vocal response is potentially a more revealing source of information about animals' experience that other pain indicators, only counting general vocalizations rates without specifying their types is not sufficient for welfare assessment (Watts and Stookey, 2000; Manteuffel et al., 2004). Hence, other pain indicators have been used to quantify changes in animal behaviour following castration, i.e. escape behaviour, struggle, locomotion activity, lying time, kicks, chute behaviour, and feeding behaviour (Fell et al., 1986; González et al., 2010; Coetzee et al., 2012; Pieler et al., 2013; Petherick, 2011; del Campo et al., 2014; Moya et al., 2014).

In our MA, the potential for detection bias was high and the obvious problem is the inherent subjectivity. Training independent observers using specific criteria and, moreover, the use of automated measures of animal behaviour, can improve the scientific value of vocal response, mainly in welfare assessment (Watts and Stookey, 2000; Manteuffel et al., 2004; Viñuela-Fernandez et al., 2011).

The effect of castration on cortisol concentration

Acute pain is a response to an established inflammatory and metabolic process that activates the hypothalamic-pituitary-adrenal axis (Mellor et al., 2000). Therefore, changes in cortisol concentration appear to be particularly useful evaluating pain assessment, despite monitoring value being limited by the difficulty of measuring the system's reaction, as well as the inter-animal variations to the stress response (Moberg, 2000; Möstl and Palme, 2002). Variations in response may also be due to differences in the way in which a castration technique is conducted by different operators (Coetzee, 2013). Thus, this variability decreases our capacity to detect differences among groups, and a greater numbers of animals are required for experimental models (Mellor et al., 2000). Our data showed a great variability in a small sample size (mean = 33.5; minimum = 10; maximum = 60) in the included SR-MA studies.

Cortisol response has been widely used to assess well-being in farm animals, providing an indication of the overall noxiousness of the experience. However, comparisons between manuscripts can be difficult due to the previous experience, ability to learn adaptative, concurrent stressors, circadian rhythm, differences in collection sample methods performed by a different operator, pharmacokinetic model, delay between the time of pain relief administration and the onset of analgesic activity, and analytical methodology (Mellor et al., 2000; Möstl and Palme, 2002; Coetzee et al., 2010, 2011). The pharmacokinetics and pharmacodynamics of the main drugs used for sedation, anaesthesia, or analgesia affect the pain management in cattle (Smith, 2013). Studies to evaluate possible circadian rhythm in cattle showed controversial results, as reported by a diurnal ultradian rhythm and a very weak circadian rhythm (Lefcourt et al., 1993) or by no diurnal variation in the endogenous cortisol secretion (Hudson et al., 1975). Also, it has been proposed that low cortisol responses may appear in individuals with high pain threshold or when the physiological effect of castration procedure was easily observed (Stafford and Mellor, 2005). However, the reasons why some individuals may have different responses are still unclear.

The evidences in our meta-analysis suggest that the surgical and non-surgical procedures without drug administration, when compared to uncastrated animals, did not increase cortisol levels as expected. One probable explanation for this observation is the abrupt change of the cortisol measurement after an animal intervention that may influence its final conclusive interpretation. Nineteen bibliographic references were screened and analysed by Bretschneider (2005), who showed that castration caused a fast and maximum adrenal corticoid secretion 12 min after the surgical procedure. Second, the explanation is that sampling blood from animals using catheters with minimal non-aversive handling did not stress the animals (Schwartzkopf-Genswein et al., 2005). Thus, the majority of included studies analysed cortisol in blood by single invasive collection samples (Petherick, 2012; del Campo et al., 2014). Third, there were no significant differences between non-castrated and castrated animals in physiological parameters (Coetzee et al., 2008; Petherick, 2012; del Campo et al., 2014). Cortisol concentration may be increased in response to the stress of animal handling itself and as an invasive method, then hardly difficult to distinguish between non-threatening stress and distress (Moberg, 2000). As mentioned, the absence of variation in cortisol responses can be affected by animals' internal and external characteristics.

Furthermore, a similar pattern of cortisol levels in surgical *vs.* non-surgical castration was observed in our study, in agreement to Petherick (2012) and del Campo et al. (2014). As concluded by Stafford et al. (2002), all methods cause an immediate and significant rise in cortisol concentration, but the *ceiling effect* of cortisol responses can lead to underestimate the adverse effects of the most invasive treatments (Mellor et al., 2000). On the other hand, researchers found an increase in cortisol concentration in surgical (Fell et al., 1986) or non-surgical castration (Petherick, 2011). Special attention is needed when interpreting individual differences, cattle age, and different castration techniques that could influence the cortisol results (Stafford et al., 2002; del Campo et al., 2014).

Despite the multimodal analgesic approach being more effective in mitigating pain associated with castration than a single analgesic agent (Coetzee, 2013), no effect of multimodal therapy in decreasing cortisol concentration during non-surgical castration was observed in the first 30 min. As highlighted by Coetzee (2011), the main challenges to provide an effective analgesia are the delay between the time of drug administration and the onset of analgesia activity, and the route or method of analgesic drug administration. Thus, we should consider the timing of administration (30 and 1 min before start of the procedure) and the drug administration route (epidural + i.v. jugular and i.m.+ local), as well as the control groups (one maintained intact and another submitted to non-surgical castration) and the strategy used for the pain relief (analgesic-sedative and anti-inflammatory or anaesthesia) of the included studies in this MA. Both studies used saliva samples to measure the cortisol concentration which may be ineffective as an indicator of immediate or chronic pain (González et al., 2010; Pieler et al., 2013). Moreover, the optimum balance of analgesic efficacy can be achieved by the combination of anaesthesia with anti-inflammatory (Coetzee, 2011, 2013), that could virtually eliminate the cortisol response during the first 8 h, and by inference, the pain and distress (Stafford et al., 2002).

With respect to anaesthesia, we found no evidence that this pain strategy reduces cortisol level after 120 min of surgical castration. Lidocaine, a short-acting local anaesthetic, was used in these studies, which is effective for approximately 45-90 min and reduces the acute distress associated with castration (Coetzee, 2013; Stafford and Mellor, 2005;). Studies detected that anaesthesia can attenuate serum cortisol response (del Campo et al., 2014; Stafford et al., 2002), with no difference in the integrated cortisol response during 60-150 min post-castration in comparison to uncastrated animals (Coetzee et al., 2010). A topical anaesthetic spray can be used to reduce pain for up to 24 h as a practical and affordable approach to beef cattle farm management (Lomax and Windsor, 2013).

Publication bias in the literature is likely to be reflected in the MA approaches (Borenstein et al., 2009). In this case, visual assessment and adjusted rank correlation test indicated some evidence of the presence of some bias. Funnel-plot asymmetries may also have resulted from clinical heterogeneity among studies (e.g. poor methodological design) (Lean et al., 2009). Inadequate quality of primary research has also been reported to yield larger effects (Egger et al., 2001). Meta-regression analysis suggested that studies from non-peer reviewed studies or those without control of confounders change the cortisol response.

The distinct result pattern observed in the cumulative MA, i.e. over time there was clear evidence that cortisol concentration decreased in castrated cattle, which might be related to the public concern about the welfare of farm animals and to the use of pain mitigation strategies. The changes observed in the effect size can be the result of the increase in the interest in pain caused by routine husbandry practices (Stafford and Mellor, 2005), as well as the improvement in study quality.

The average effect size changed after the removal of two and one studies, respectively. Coetzee et al. (2010) was the only study showing no clustering or group hierarchy, using a relatively small sample size (n= 22 animals), and high precision of the estimate was obtained directly from the graph published. A study performed by del Campo et al. (2014) was conducted in South America (Uruguay) and the animals were the youngest (7 days of age), but the intervention protocol was not described in sufficient detail.

The results of this MA complement and extend previous research describing the effect of castration in cortisol levels. However, results described in the literature are discordant, and additional studies are required.

The effect of castration on ADG

Production parameters do not reflect the pain experienced by cattle (Stafford and Mellor, 2005) at the moment of animal castration. Then, castration technique may not be as important from a growth rate standpoint, but can have negative effects on the feed intake and performance (Molony et al., 1995; Pang et al., 2008; Gonzaléz et al., 2010; Moya et al., 2014), mainly in the intensified production systems. In addition, the lower body weight gain in castrated males was possibly due to the decrease of testosterone (Fisher et al., 2001; Pang et al., 2008). However, assessment of these parameters is critical if research on animal welfare is to be of relevance for livestock producers (Coetzee et al., 2011) and more research is required to determine the relationship between castration and feed intake, growth rate, and feed efficiency (González et al., 2010).

The effect of castration in ADG had the largest number of trials for our MA. A single study was responsible for reducing the effect size and provides a non-significant change in ADG after castration (Whitlock et al., 2013). This influential study was published in conference proceedings. The abstract format does not allow precise, informative presentation of the methodology used and we cannot rely on contacting the manuscript authors (Egger et al., 2001), thus meaning that we did not have access to the final data for a more precise analysis.

In the univariate meta-regression analysis we explored the influence of the follow-up period on ADG, and the results showed no effect, i.e. timing that the ADG was measured in the included studies did not influence our result. As shown by many researchers, the differences in performance between non-castrated and castrated males cattle are mainly manifested after puberty at an average age of 10 months (Barber and Almquist, 1975; Lunstra et al., 1978) or when testosterone concentration peaked at 15 months of age (Gerrard et al., 1987). Concentrations of serum LH and testosterone increased linearly with advance age around the time of puberty (Lunstra et al., 1978) and, then, this is the period that occur the biggest contribution of testicular tissue on animal growth.

Field (1971) concluded after literature review that bulls gained weight 17% faster and were 13% more efficient in converting feed in live weight than steers. The decrease in ADG after castration in the first two weeks of age (Pang et al., 2008; Warnock et al., 2012) and the reduction in body growth rate (Knight et al., 2000; Fisher et al., 2001; González et al., 2010) can be attenuated at 28 days (Coetzee et al., 2012), 30 days (Knight et al., 2000) and 42 days post-castration (Warnock et al., 2012).

Despite the above reports, we did not find differences in growth performance favouring the non-castrated group in our MA. Inadequate nutrition (Bailey and Hironaka, 1969; Martin et al., 1978), as well as the more aggressive behaviour (Martí, 2012), can prevent bulls from expressing their greater productive potential for weight gain. Animal body weight can also be related to age at the time of intervention, hormonal status of the control group, castration method, feed intake, feeding activity, feeding program, and the level of performance achieved (Pang et al., 2008; González et al., 2010; Martí, 2012; Warnock et al., 2012). Furthermore, relevant information, i.e. feed behaviour and physiology, were not available in our database. Thus, caution in drawing final conclusions is crucial because the live weights in the short period of observation (minimum = 27 days; maximum = 217 days; mean = 87 days) are difficult to quantify, mainly when the difference in ADG between castrated and uncastrated cattle groups is below 1 kg/day.

Although del Campo et al. (2014) showed greater ADG for non-surgical than surgical castration, we found no strong evidence when both groups were compared. Therefore, the age at which male calves were banded or surgically castrated did not affect the weight at 217 days (Baker et al., 2000; Bretschneider, 2005), as well as on gain:feed when castration was performed after 7 months of age (Warnock et al., 2012; Reppening et al., 2013). The castration method may not influence the growth rate in the long term, indicating that beef cattle were able to compensate and recover from the castration technique intervention (Warnock et al., 2012; Pieler et al., 2013).

In agreement with our results, Newton and O'Connor (2013) showed that there was little evidence of castration effect on ADG regardless of the type of pain management. It may be that the cattle that experienced distress after the application of medication suffer changes in social status that lead to permanent changes in behaviour. It would be interesting to obtain more information on pain mitigation, i.e. the route of administration, period of exposure, and optimum dose (González et al., 2010; Coetzee et al., 2012).

The present MA has several limitations. First, the approach in reporting outcomes often limited our ability to summarize the data, as there was incomplete reporting of summary measures. However, an attempt was made by contacting researchers in the field, as suggested by Lean et al. (2009). We had excluded six full-text studies on castration because they were in German, Japanese, and Bulgarian as explained in the methodology section. Finally, in the absence of robust and specific direct and indirect measures associated with pain, the choice of parameters about welfare and its relationship with castration may be challenging for precise analysis.

Conclusions

In summary, this is the first SR-MA that summarized the available literature on the effects of castration on cortisol, ADG, and vocalization in beef cattle. There was limited evidence that the use of pain relief mitigated pain responses to castration, as well as which castration method was less painful. That lack of effect might be due to insufficient doses or inadequate duration of action of the drugs used, or due to low capacity of cortisol and ADG to detect pain caused by castration in beef cattle. The challenge in animal science studies is to provide complete and accurate details of the methodology using standardized guidelines available in the published literature.

Acknowledgements

The authors thank Bárbara Bremm and Claudia Medeiros Camargo for help during the screening step and all the experts who provided feedback during the systematic review-meta-analysis process. This study was supported by The Brazilian Council of Scientific and Technological Development (CNPq) and by Coordination for the Improvement of Higher Education Personnel (CAPES).

References

American Veterinary Medical Association, 2014. Literature review on welfare implications of castration of cattle (July 15, 2014). (https://www.avma.org/KB/Resources/LiteratureReviews/Documents/castration-cattle-bgnd.pdf. Accessed April 2nd, 2016).

Baker, J.F., Stricklan, J.E., Vann, R.C., 2000. Effect of castration on weight gain of beef calves. The Bovine Practitioner 34, 124-126.

Bailey, C.B., Hironaka, R., 1969. Growth and carcass characteristics of bulls, steers, and partial castrates kept on range for the first year of life and then fattened. Can. J. Anim. Sci. 49, 37-44.

Barber, K.A., Almlquist, J.O., 1975. Growth and feed efficiency and their relationship to puberal traits of Charolais bulls. J. Anim. Sci. 40, 288-301.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. Introduction to meta-analysis. John Wiley and Sons, Ltd., The Atrium, Chichester, UK.

Bretschneider, G., 2005. Effects of age and method of castration on performance and stress response of beef male cattle: a review. Livest. Prod. Sci. 97, 89-100.

Ceballos, A., Sánchez, J., Stryhn, H., Montgomery, J.B., Barkema, H.W., Wichte, J.J., 2009. Meta-analysis of the effect of oral selenium supplementation on milk selenium concentration in cattle. J. Dairy Sci. 92, 324–342.

Coetzee, J.F., 2011. A review of pain assessment techniques and pharmacological approaches to pain relief after bovine castration: practical implications for cattle production within the United States. Appl. Anim. Behav. Sci. 135, 192-213.

Coetzee, J.F., 2013. Assessment and management of pain associated with castration in cattle. Vet. Clin. Food Anim. 29,75-101.

Coetzee, J.F., Edwards, L.N., Mosher, R.A., Bello, N.M., O'Connor, A.M., Wang, B., KuKanich, B., Blasi, D.A., 2012. Effect of oral meloxicam on health and performance of beef steers relative to bulls castrated on arrival at the feedlot. J. Anim. Sci. 90, 1026-1039.

Coetzee, J.F., Gehring, R., Taurus-Sang, J., Anderson, D.E., 2010. Effect of sub-anesthesic xylazine and ketamine ('ketamine stun') administered to calves immediately prior to castration. Vet. Anaesth. Analg. 37, 566-578.

Coetzee, J.F., Lubbers, B.V., Toerber, S.R., Gehring, R., Thomson, D.U., White, B.J., Apley, M.D., 2008. Plasma concentration of substance P and cortisol in beef calves after castration or simulated castration. Am. J. Vet. Res. 69, 751-762.

del Campo, M., Hernández, S., Anchaño, M., Giorello, S., Soares de Lima, J.M., Botero, S., Ferrón, M., Freitas, G., Albernaz, F., Rodríguez, A., Piñero, J., Mondragón, J., Martínez, J., Freitas, G., Larronda, J., Lemor, R., Márquez, H., 2014. Bienestar animal en bovinos. Resultados experimentales: castración de terneros. INIA Uruguay 217, 493-527.

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control.Clin. Trials 7, 177-188.

Dzikamunhenga, R.S., Anthony, R., Coetzee, J., Gould, S., Johnson, A., Karriker, L., McKean, J., Millman, S.T., Niekamp, S.R., O'Connor, A.M., 2014. Pain management in the neonatal piglet during routine management procedure. Part 1: a systematic review of randomized and non-randomized intervention studies. Anim. Health Res. Rev. 15, 14-38.

Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463.

Egger, M., Smith, G.D., Altman, D.G., 2001. Systematic reviews in health care, second ed. MBJ Publishing Group, London, UK.

Faulkner, D.B., Eurell, T., Tranquilli, W.J., Ott, R.S., Ohl, M.W., Cmarik, G.F., Zinn, G., 1992. Performance and health of weanling bulls after butorphanol and xylazine administration after castration. J. Anim. Sci. 70, 2970-2974.

Falzon, L.C., O'Neill, T.J., Menzies, P.I., Peregrine, A.S., Jones-Bitton, A., vanLeeuwen, J., Mederos, A., 2014. A systematic review and meta-analysis of factors associated with anthelmintic resistance in sheep. Prev. Vet. Med. 117, 388-402.

Fell, L.R., Wells, R., Shutt, D.A., 1986. Stress in calves castrated surgically or by the application of rubber rings. Aust. Vet. J. 63, 16-18.

Field, R.A., 1971. Effect of castration on meat quality and quantity. J. Anim. Sci. 32, 849-858.

Fisher, A.D., Knigh, T.W., Cosgrove, G.P., Death, A.F, Anderson, C.B., Duganzich, D.M., Matthews, D.R., 2001. Effects of surgical or band castration on stress responses and behaviour of bulls. Aust. Vet. J. 79, 279-284.

Gerrard, D.E., Jones, S.J., Aberle, E.D., Lemenager, R.P., Diekman, M.A., Judge, M.D., 1987. Collagen stability, testosterone secretion and meat tenderness in growing bulls and steers. J. Anim. Sci. 65, 1236-1242.

González, L.A., Schwartzkopf-Genswein, K.S., Caulkett, N.A., Jansen, E., McAllister, T.A., Fierheller, E., Schaefer, A.L., Haley, D.B., Stookey, J.M., Hendrick, S., 2010. Pain mitigation after band castration of beef calves and its effects on performance, behaviour, *Escherichia coli*, and salivary cortisol. J. Anim. Sci. 88, 802-810.

Golder, H.M., Lean, I.J., 2016. A meta-analysis of lasalocid effects on rumen measures, beef and dairy performance, and carcass traits in cattle. J. Anim. Sci. 94, 306-326.

Higgins, J.P.T., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. http://www.cochrane-handbook.org.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analysis. BMJ 327, 557-560.

Hudson, S., Mullord, M., Whittlestone, W.G., Payne. E., 1975. Diurnal variations in blood cortisol in the dairy cow. J. Dairy Sci. 58, 30-33.

Knight, T.W., Cosgrove, G.P., Death, A.F., Anderson, C.B., Fisher, A.D., 2000. Effect of method of castrating bulls on their growth rate and liveweight. New Zeal. J. Agr. Res. 43, 187-192.

Lean, I J., Rabiee, A.R., Duffield, T.F., Dohoo, I.R., 2009. Invited review: use of meta-analysis in animal health and reproduction: methods and applications. J. Dairy Sci. 92, 3545–3565.

Lefcourt, A.M., Bitman, J., Kahl, S., Wood, D.L., 1993. Circadian and ultradian rhythms of peripheral cortisol concentrations in lactating dairy cows. J. Dairy Sci. 76, 2607-2612.

Lomax, S., Windsor, P.A., 2013. Topical anesthesia mitigates the pain of castration in beef calves. J. Anim. Sci. 91, 4945-4952.

Lunstra, D.D., Ford, J.J., Echternkamp, S.E., 1978. Puberty in beef bulls: hormone concentrations, growth, testicular development, sperm production and sexual aggressiveness in bulls of different bulls. J. Anim. Sci. 46, 1054-1062.

Manteuffel, G., Puppe, B., Schön, P.C., 2004. Vocalization of farm animals as a measure of welfare. Appl. Anim. Behav. Sci. 88, 163-182.

Martí, S., 2012. Physical castration and immunocastration of early-maturing bulls fed high concentrate diets: welfare, performance, and carcass and meat quality. PhD Thesis. Departament de Ciència Animal i dels Aliments. Universitat Autònoma de Barcelona. Barcelona.

Martin, T.G., Perry, T.W., Beeson, W.M., Mohler, M.T., 1978. Protein levels for bulls: comparison of three continuous dietary levels on growth and carcass traits. J. Anim. Sci. 47, 29-33.

Mederos, A., Waddell, L., Sánchez, J., Kelton, D., Peregrine, A.S., Menzies, P., VanLeeuwen, J., Rajic, A., 2012. A systematic review-meta-analysis of primary research investigating the effect of selected alternative treatments on gastrointestinal nematodes in sheep under field conditions. Prev. Vet. Med. 104, 1-14.

Mellor, D J., Cook, C.J., Stafford, K.J., 2000. Quantifying some responses to pain as a stressor, in: Moberg, G.P., Mernch, J.P., (Eds.), The biology of animal stress: basic principles and implications of animal welfare. CABI, London, UK, pp. 171-198.

Mintline, E.M., Varga, A., Banuelos, J., Walker, K.A., Hoar, B. , Drake, D., Weary, D.M., Coetzee, J.F., Stock, M.L., Tucker, C.B., 2014. Healing of surgical castration wounds: a description and an evaluation of flunixin. J. Anim. Sci. 92, 5655-5665.

Moberg, G.P. 2000. Biological response to stress: implications for animal welfare, in: Moberg, G.P., Mernch, J.P., (Eds.), The biology of animal stress: basic principles and implications of animal welfare. CABI, London, UK, pp. 1-21.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Open Med. 3, 123–130.

Molony, V., Kent, J.E., Robertson, I.S., 1995. Assessment of acute and chronic pain after different methods of castration of calves. Appl. Anim. Behav. Sci. 46, 33-48.

Möstl, E., Palme, R., 2002. Hormones as indicators of stress. Domest. Anim. Endocrinol. 23, 67-74.

Moya, D., González, L.A., Janzen, E., Caulkett, N.A., Fierheller, E., Schwartzkopf-Genswein, K.S., 2014. Effects of castration method and frequency of intramuscular injections of ketoprofen on behavioral and physiological indicators of pain in beef cattle. J. Anim. Sci. 92, 1686-1697.

Newton, H.P., O'Connor, A.M., 2013. The economics of pain management. Vet. Clin. North Am. Food Anim. Pract. 29, 229-250.

Pang, W.Y., Earley, B., Gath, V., Crowe, M.A., 2008. Effect of banding or burdizzo castration on plasma testosterone, acute-phase proteins, scrotal circumference, growth, and health of bulls. Livest. Sci. 117, 79-87.

Petherick, J.C., 2011. Pain management in castrated beef cattle – Final report. Meat and Livestock Australia Limited, Australia.

Petherick, J.C., 2012. Welfare outcomes of calves of two ages castrated by rings – Final report. Meat and Livestock Australia Limited, Australia.

Pieler, P., Peinhopf, W., Becher, A.C., Aurich, J.E., Rose-Meierhöfer, S., Erber, R., Möstl, E., Aurch, C., 2013. Physiological and behavioral stress parameters in calves in response to partial scrotal resection, orchidectomy, and Burdizzo castration. J. Dairy Sci. 96, 6378-6389.

Reppening, P.E., Ahola, J.K., Callan, R.J., Fox, J.T., French, J.T., Giles, R.L., Peel, R.K., Whittier, J.C., Engle, J.E., 2013. Effects of pain mitigation and method of castration on behavior and feedlot performance in cull beef cattle. J. Anim. Sci. 10, 4975-4983.

Roberts, S.L., Hughes, H.D., Burdick Sanchez, N.C., Carroll, J.A., Powell, J.G., Hubbell, D.S., Richeson, J.C, 2015. Effect of surgical castration with or without oral meloxicam on the acute inflammatory response in yearling beef bulls. J. Anim. Sci. 93, 4123-4131.

Sargeant, J.M., Amezcua, M.R., Rajic, A., Wadell, L., 2005. A guide to conducting systematic reviews in agri-food public health. Public Health Agency of Canada, Canada.

Schwartzkopf-Genswein, K.S., Booth-McLean, M.E., McAllister, M.A., Mears, G.J., 2005. Physiological and behavioural changes in Holstein calves during and after dehorning and castration. Can. J. Anim. Sci. 82:131-138.

Schwartzkopf-Genswein, K.S., Fierheller, E.E., Caulkett, N.A., Janzen, E.D., Pajor, E.A., González, L.A., Moya, D., 2012. Achieving pain control for routine management procedures in North American beef cattle. Anim. Front. 2, 52-57.

Smith, G., 2013. Extralabel use of anesthetic and analgesic compounds in cattle. Vet. Clin. North Am. Food Anim. Pract. 29, 29-45.

Stafford, K.J., Mellor, D.J., 2005. The welfare significance of the castration of cattle: a review. N Z Vet. J. 56, 271-278.

Stafford, K.J., 2007. Alleviating the pain caused by the castration of cattle. Vet J 173, 245-247.

Stafford, K.J., Mellor, D.J, 2010. Painful husbandry procedures in livestock and poultry, in: grandin, T. (Ed.), Improving animal welfare: a practical approach. London, UK, CABI, pp. 88-114

Stafford, K.J., Mellor, D.J., Todd, S.E., Bruce, R.A., Ward, R.N., 2002. Effects of anaesthesia or anaesthesia plus a non-steroidal anti-inflammatory drug on the acute cortisol response of calves to five different methods of castration. Res. Vet. Sci. 73, 61-70.

Thüer, S., S. Mellema, Doherr, M.G., Wechsler, B., Nuss, K., Steiner, A., 2007. Effect of anaesthesia on short- and long-term pain induced by two bloodless castration methods in calves. Vet. J. 173, 333-342.

Viñuela-Fernández, I., Jonnes, E., Welsh, E.M., Fleetwood-Walker, S.M., 2007. Pain mechanisms and their implication for the management of pain in farm and companion animals. Vet. J. 174, 227-239.

Viñuela-Fernández, I., Weary, D. M., Flecknell, P., 2011. Pain, in: Appleby, M.C., Mench, J.A., Olsson, I.A.S., Hughes, B.O. (Eds.), Animal Welfare. Wallingford, UK, CABI, pp. 64-77.

Warnock, T., Thrift, T., Irsik, M., Herson, M., Maddock, T., Lamb, C., Arthington, J., 2012. Effect of castration technique on beef calf performance, residual feed intake, and inflammatory response. J. Anim. Sci. 90, 2345-2352.

Watts, J.M., Stookey, J.M., 2000. Vocal behavior in cattle: the animal's commentary on its biological processes and welfare. Appl. Anim. Behav. Sci. 67, 15-33.

Weary, D.M., Niel, L., Flower, F.C., Fraser, D., 2006. Identifying and preventing pain in animals. Appl. Anim. Behav. Sci. 100, 64-76.

Whitlock, B., Krawczel, P., Carrol, J., Burdick Sanchez, N., Dailey, J., Daniel, J., Coetzee, J., 2013. Effect of meloxicam on gain and inflammatory response of calves castrated by banding post-weaning. J. Anim. Sci. 91 (E-Suppl. 2), 532 (Abstr.).

Table A.1. Population, intervention, and outcome search term strings used for the final search.

Acronym	Search string
Population	Bovine: refers to the subfamily Bovinae, which includes cattle, buffalo, and kudus.
	Beef cattle: are the domestic cattle to produce meat.
	Calf: as a young female or male bovine up to weaning.
	Herd: a group of animals that live or are kept together.
Intervention	Disbudding: refers to prevention of horn growth before it has become advanced.
	Dehorning: the amputation of horns at any stage after their growth of the early budding stage.
	Castration: is the process of removal, damage, or destruction of the testicles.
Outcome	Animal welfare or animal well-being: involves basic health and functioning, natural living and affective state.
	Animal pain: is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or is describable in terms of such damage.
	Animal stress: biological response elicited when an individual perceives a stressor to its homeostasis.
	Cortisol: widely used as a hormonal indicator of pain-induced distress caused by a range of husbandry practices in farm animals. In response to emotionally and physically noxious experiences, there is an increase in the activity of the hypothalamic-pituitary-adrenocortical system and, then, in the cortisol level.
	Behaviour: farm animal welfare behaviour has been used to assess the response

of calves, deer and lambs to painful husbandry procedures. Behaviours indicators, measured objectively or subjectively, can provide robust assessment tools for pain whereby they are clearly explained and validated.

Vocalization: vocalization is a good behavioural indicator of pain in farm animals. Hence, researchers are interested in using vocal behavior as a manner of evaluating animal welfare.

Table A.2. List of relevant publications excluded from the final dataset in the meta-analysis.

Reference	Country	Treatment	Analgesic regimen	Outcome parameter	Reason for exclusion
King et al., 1991	Canada	Surgical and non-surgical	NA	Cortisol and ADG	Number of animals per group not presented
Coetzee et al., 2007	USA	Surgical	Anti-inflammatory	Cortisol	No baseline value
Boesch et al., 2008	Switzerland	Non-surgical	Anaesthesia	Cortisol	Only median was presented
Currah et al., 2009	Canada	Surgical	Anaesthesia and anti-inflammatory	Vocalization	Numerical data not shown
González et al., 2009	Canada	Surgical and non-surgical	Anaesthesia	ADG	Only p-value was presented
Becker et al., 2012	Switzerland	Non-surgical	NA	Cortisol and ADG	Number of animals per group not presented
Brown et al., 2012	USA	Surgical	Anti-inflammatory	ADG	Number of animals per group not presented
Brown et al., 2013	USA	Surgical	Anti-inflammatory	ADG	Only p-value was presented

Daniel et al., 2013	USA	Non-surgical	Anti-inflammatory	ADG	Only p-value was presented
Moya et al., 2014	Canada	Surgical and non-surgical	Anti-inflammatory	Cortisol and ADG	Insufficient data for this study

NA: not applicable.

Table A.3. Descriptive characteristics of 20 publications reporting 26 studies included in the meta-analysis.

Variable	Description	Categories	Number of publications (studies)
Study design	Type of study design used	Control studies	20 (26)
Peer-reviewed	Type of literature the work was published	Peer-reviewed	16 (17)
		Conference proceedings	1 (1)
		Thesis	1 (1)
		Government or research station report	2 (7)
Treatment	Type of procedure evaluated	Surgical castration	13 (19)
		Non-surgical castration	13 (17)
		Non-surgical vs. Surgical castration	9 (14)
Data published	Year of study publication	1990-2000	2 (2)
		2001-2015	18 (24)
Medicament	It was used any class of medicament?	No	15 (21)
		Yes	11 (15)
Medicament	If was used any medicament	Analgesic-sedative	2 (2)

	to mitigate pain, which class?		
		Anaesthesia	5 (7)
		Anti-inflammatory	4 (6)
		Multimodal therapy	3 (4)
Cattle group	Cattle group in which interventions were evaluated	*Bos taurus taurus*	9 (11)
		Bos taurus indicus	0 (0)
		Hybrid / Mixed	11 (13)
		Not reported	2 (2)
Who performed	Who performed procedure	Farm staff	5 (6)
		Veterinarian	17 (14)
		Not reported	3 (3)
Outcome assessed	Parameter used to assess pain in calves	ADG	15 (20)
		Cortisol	10 (12)
Sample size	Size of total study population per study	n≤50	13 (16)
		n= 51-100	6 (8)
		n≥101	3 (3)
Continent		North America	12 (13)
		South America	1 (5)
		Europe	3 (3)
		Asia	0 (0)
		Oceania	4 (5)

Table A.4. Summary of assessment for methodological soundness and reporting of 20 publications reporting 26 studies including in this meta-analysis.

		Number of publications (studies)		
Variable	Assessment	ADG	Cortisol	Vocalization
Was the sample size justified?	Yes	1 (1)	1 (1)	0 (0)
	No	14 (19)	9 (11)	2 (2)
How were calves assigned to treatment groups?	Random[1]	0 (0)	2 (2)	1 (1)
	Reported random[2]	10 (11)	4 (4)	1 (1)
	Systematic[3]	1 (1)	0 (0)	0 (0)
	Convenience or unreported[4]	4 (8)	4 (6)	0 (0)
Was the intervention protocol described in sufficient detail to be replicated?	Yes	15 (16)	9 (10)	2 (2)
	No	1 (4)	1 (2)	0 (0)
	Reference paper	0 (0)	0 (0)	0 (0)
Did the author report that blinding was used to evaluate the outcome?	Yes	1 (1)	2 (2)	0 (0)
	No	14 (19)	8 (10)	2 (2)
Based on the study design was	Yes	15	8 (10)	1 (1)

clustering[5] accounted for appropriately in the analysis?		(20)		
	No	0 (0)	1 (1)	1 (1)
	Not applicable	0 (0)	1 (1)	0 (0)
Were identified confounders controlled for or tested?	Yes, analysis[6]	9 (10)	7 (7)	2 (2)
	Yes, inclusion/exclusion[7]	3 (3)	2 (3)	0 (0)
	Yes, matching[8]	0 (0)	0 (0)	0 (0)
	No[9]	1 (1)	0 (0)	0 (0)
	Not applicable[10]	2 (6)	2 (2)	0 (0)
Was the statistical analysis described adequately so it can be reproduced?	Yes	15 (20)	9 (11)	2 (2)
	No	0 (0)	1 (1)	0 (0)
	Reference paper	0 (0)	0 (0)	0 (0)
	Statistical analysis not done	0 (0)	0 (0)	0 (0)

[1] Computer or random number table, *a priori*, stratified random sample, cluster random sample.

[2] Author(s) report random, but randomization is not described.

[3] "n" samples obtained at x intervals or stratified by certain characteristics.

[4] Author indicated convenience sampling or sampling was not reported in the paper.

[5] Clustering was evaluated when repeated measures were reported.

[6] Author identified confounders and controlled for them in the analysis.

[7]Confounders were identified and included/excluded a priori.

[8]Confounders were controlled a priori by matching on certain characteristics.

[9] No adjustments were made for confounders/effect modifiers, etc., that were identified by the author.

[10]Confounders were not identified by the author or randomization was used to control for

Table A.5. Methodological quality assessment of risk of bias (classified as low, unclear, and high) of the 26 studies included in the meta-analysis in welfare animals from castrated beef cattle.

Reference	Sequence generation	Allocation concealment	Selective reporting	Outcome measurement	Blinding of personnel	Blinding of outcome assessment	Incomplete outcome data
Fell et al., 1986	Unclear	Unclear	High	Cortisol	Unclear	Low	Low
Faulkner et al., 1992	Low	Unclear	Low	ADG	Unclear	Low	Low
Baker et al., 2000	Low	Unclear	Low	ADG	Unclear	Low	Low
Fischer et al., 2001	Low	Unclear	High	ADG	Unclear	Low	Low
Thüer et al., 2007	Low	Low	Low	Cortisol	Unclear	Low	Low

Coetzee et al., 2008	Low	Low	Unclear	Cortisol	Unclear	Low	Low
				Vocalization	Unclear	High	Low
Pang et al., 2008	Low	Unclear	Low	ADG	Unclear	Low	Low
Coetzee et al., 2010	Low	Low	Unclear	Cortisol	Unclear	Low	Low
				Vocalization	Unclear	High	Low
González et al., 2010	Low	Unclear	High	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
Petherick et al., 2011	High	High	Low	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
Coetzee et al., 2012	Low	Low	Low	ADG	Unclear	Low	Low

Warnock et al., 2012	Low	Unclear	High	ADG	Unclear	Low	Low
Petherick et al., 2012	High	High	Low	Cortisol	Unclear	Low	High
Petherick et al., 2012	High	High	Low	Cortisol	Unclear	Low	High
Martí, 2012	Low	Unclear	High	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
Reppening et al., 2013	High	High	Low	ADG	Unclear	Low	Low
Pieler et al., 2013	Low	Unclear	Low	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
Whitlock et al.,	Low	Unclear	High	ADG	Unclear	Low	Low

2013							
del Campo et al., 2014	Unclear	Unclear	High	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
del Campo et al., 2014	Unclear	Unclear	High	ADG	Unclear	Low	Low
del Campo et al., 2014	Unclear	Unclear	High	Cortisol	Unclear	Low	Low
				ADG	Unclear	Low	Low
del Campo et al., 2014	Unclear	Unclear	High	ADG	Unclear	Low	Low
del Campo et al., 2014	Unclear	Unclear	High	ADG	Unclear	Low	Low
Mintline et al., 2014	Unclear	Unclear	Low	ADG	Low	Low	Low

Moya et al., 2014	Low	Unclear	High	ADG	Unclear	Low	Low
Moya et al., 2014	Low	Unclear	High	ADG	Unclear	Low	Low

ADG: average daily gain

Fig. 1. Flow diagram indicating the number of abstracts and publications included and excluded in each level. MA: meta-analysis. Adapted from PRISMA guidelines (Moher et al., 2009).

*Data from both procedures (castration and dehorning) are presented in the flow diagram to allow the researchers to update the same systematic review.

Fig. 2. Forest plot of studies that analysed the effect of surgical castration with no pain mitigation (on the right) in comparison to uncastrated (on the left) at 30 min (a) and to uncastrated or surgical castration without pain mitigation (on the left) at 120 min (b). The effect size (ES) is the mean difference between treated and control groups, expressed in cortisol concentration (nmol/L). *Note:* The size of the plotting symbol for the point estimate in each study is proportional to the weight that each trial contributes in the meta-analysis. The dashed line is the average effect of treatment obtained by the analysis, while the solid vertical line marks the value at which the treatment would have no effect. The overall estimate and the confidence interval are marked by a diamond (♦).

Fig. 3. Funnel plot obtained with the Duval and Tweedie's "trim-and-fill" linear random effect model measuring standard mean difference in cortisol concentration as an outcome. The circles represent the original point estimate for each study (MD) and the circles encased in a square represent the studies that the program imputed (n = 13) to create a symmetrical plot.