J. Dairy Sci. 96:2973–2987
http://dx.doi.org/10.3168/jds.2012-6154

Therapy of bovine endometritis with prostaglandin $F_{2\alpha}$: A meta-analysis

P. Haimerl, W. Heuwieser,[1] and S. Arlt
Clinic for Animal Reproduction, Faculty of Veterinary Medicine, Freie Universität Berlin, Koenigsweg 65, 14163 Berlin, Germany

ABSTRACT

The objective of the conducted meta-analysis was to assess the efficacy of the treatment of bovine endometritis with $PGF_{2\alpha}$ by statistical means. Postpartum uterine infections have a high prevalence and a very negative effect on reproductive performance in dairy cattle. Because of a wide discordance between research results, a meta-analysis of the efficacy of the treatment of bovine endometritis with $PGF_{2\alpha}$ was conducted. A comprehensive literature search was performed using online databases to reveal a total of 2,307 references. In addition, 5 articles were retrieved by reviewing citations. After applying specific exclusion criteria and evaluating specific evidence parameters, 5 publications, comprising 6 trials, were eligible for being analyzed by means of meta-analysis. Data for each trial were extracted and analyzed using meta-analysis software Review Manager (version 5.1; The Nordic Cochrane Centre, Copenhagen, Denmark). Estimated effect sizes of $PGF_{2\alpha}$ were calculated on calving to first service and calving to conception interval. Prostaglandin $F_{2\alpha}$ treatment of cows with chronic endometritis had a negative effect on both reproductive performance parameters. Heterogeneity was substantial for calving to first service and calving to conception interval [I^2 (measure of variation beyond chance) = 100 and 87%, respectively]; therefore, random-effects models were used. Sensitivity analysis as well as subgroup analysis showed that the performance of randomization was influential in modifying effect size of $PGF_{2\alpha}$ treatment. The funnel plot illustrated a publication bias toward smaller studies that reported a prolonged calving to conception interval after a $PGF_{2\alpha}$ treatment. We conclude that the investigation of this subject by means of meta-analysis did not reveal an improvement of reproductive performance of cows with endometritis after treatment with $PGF_{2\alpha}$. Furthermore, there is a shortage of comparable high quality studies investigating reproductive performance after $PGF_{2\alpha}$ treatment of cows with chronic endometritis.

Key words: meta-analysis, endometritis, prostaglandin $F_{2\alpha}$, dairy cow

INTRODUCTION

Postpartum uterine infections are a frequent disorder in dairy cattle with a prevalence of up to 57.7% (Sheldon, 2009). In addition, they are reported to have an immense negative effect on reproductive performance resulting in high opportunity costs for the farmers (Plaizier et al., 1998; LeBlanc et al., 2002; LeBlanc, 2008).

Clinical endometritis in cattle is defined as the presence of a purulent (>50% pus) uterine discharge detectable in the vagina 21 d or more postpartum or mucopurulent (approximately 50% pus, 50% mucus) discharge detectable in the vagina after 26 d postpartum (Sheldon et al., 2006).

The amount of literature addressing the treatment of endometritis is huge and this subject has been reviewed by several authors (Gilbert and Schwark, 1992; Olson, 1996; Azawi, 2008; Lefebvre and Stock, 2012). Nevertheless, the treatment of endometritis is still an issue of considerable controversy (Arlt et al., 2009; Dubuc et al., 2011). This may be due to the wide variety of therapies available for endometritis, including systemic or local antibiotics, $PGF_{2\alpha}$, and estradiol. Because the rate of self-cure is reported to range from 92% in wk 1 to 25% in wk 7 postpartum (Falkenberg and Heuwieser, 2005; Hirsbrunner et al., 2006), some authors question the necessity of any treatment at all. However, limited information exists on the proportions of cows that spontaneously recover (Dubuc et al., 2011).

The effect of a treatment with $PGF_{2\alpha}$ or its analogs within 40 d after calving on reproductive performance has been investigated in several studies (Haimerl et al., 2011). Remarkably, wide disparity exists between the results obtained (Burton and Lean, 1995). Young et al. (1984), for instance, reported a significant improvement in the first-service conception rates of 64 cows treated with $PGF_{2\alpha}$ compared with 64 untreated controls. Another study conducted by Macmillan et al. (1987)

Received September 13, 2012.
Accepted January 25, 2013.
[1]Corresponding author: w.heuwieser@fu-berlin.de

including 1,813 cows could not support these findings. According to a recent study conducted by Dubuc et al. (2011), an administration of $PGF_{2\alpha}$ at both 5 and 7 wk postpartum did not mitigate the effects of cytological endometritis or purulent vaginal discharge on reproductive performance. Thus, the authors postulate that clinical approaches to treatment of chronic postpartum reproductive tract infection and inflammation should be reassessed.

The unmanageable amount of medical information was criticized by Cochrane as early as 1972 (Cochrane, 1972). He also complained about the lack of reliable summaries of available findings, such as meta-analyses. Such situations lead to difficulties in making decisions on the basis of current and valid information (Cochrane, 1972). Meta-analyses are systematic summaries of a large collection of results from individual studies and statistical analysis of those results from multiple individual studies (Glass, 1976). According to Eisend (2004), the prototypical meta-analysis should undergo the following 5 steps. First, a clinical question should be clearly formulated, and then a systematic and comprehensive search for relevant literature must be conducted. In the third step, data from the included literature are extracted and evaluated according to predefined parameters. Then, the extracted data are combined by using statistical techniques to obtain a pooled estimate of the treatment effect (Barker and Carter, 2005). Finally, findings should be visualized (e.g., by forest plots) and critically interpreted. In conclusion, a meta-analysis is the statistical combination of at least 2 studies to generate an estimate of the magnitude of the effect of the intervention under investigation (Lam and Kennedy, 2005). Consequently, the results of meta-analyses are said to provide the greatest reliability when applied to the entire population (Arlt and Heuwieser, 2005). The principal aim of a meta-analysis is to provide an objective quantitative assessment of previously published data as well as to increase the precision of the estimate of a treatment effect by increasing the sample size and thus increasing the statistical power (Lean et al., 2009). Furthermore, meta-analyses can be conducted to identify and investigate heterogeneity in the results of the included studies, based on factors such as the design and sample variables. Finally, meta-analyses can be executed for the purpose of resolving conflicts among studies and developing new directions for research (L'Abbé et al., 1987; Henry and Wilson, 1992; Wilson and Henry, 1992).

Because of the frequent occurrence of postpartum uterine infections, the associated economic impact, and a wide discordance between research results concerning therapy, a meta-analysis of the efficacy of the treatment of bovine endometritis with $PGF_{2\alpha}$ was conducted.

MATERIALS AND METHODS

A comprehensive literature search was conducted on August 4, 2010, utilizing the search engine Vetseek (http://www.vetseek.info) and the databases Pubmed (http://www.pubmed.gov), Medline (http://www.medline.de), and Animal Production (http://www.ovid.com/site/catalog/DataBase/22.jsp) to identify literature related to the treatment of endometritis with prostaglandin in dairy cattle. The search terms "endometritis AND cattle" and "endometritis AND cattle AND prostaglandin" were used to include all articles addressing the treatment of bovine endometritis with $PGF_{2\alpha}$. In addition, we carried out a systematic review of citations in the retrieved papers.

Selection by Inclusion and Exclusion Criteria

Specific exclusion criteria were defined to exclude studies that were not written in English or German or did not focus on chronic endometritis; that is, occurring on and after 21 d after parturition (Sheldon et al., 2006). In addition, studies in which the animals received concomitant treatments with medications other than $PGF_{2\alpha}$ were excluded. Book chapters, case studies, review articles, and abstracts were excluded. Furthermore, publications describing etiological, epidemiological, microbiological, or nutritional results, clinical symptoms, or diagnostic procedures were rejected. Articles not meeting the inclusion criteria due to incorrect indexing and those not obtainable through the internet, bibliographies, or interlibrary lending services were excluded as well. If multiple publications were retrieved that described the same trial, those containing the least information were regarded as duplicates and excluded. Retrieval and management of references was performed with Endnote (version X3 for Windows, Thomson Reuters, New York, NY).

To examine quality and comparability, the remaining publications were evaluated according to various evidence parameters, utilizing an evaluation form developed by Arlt et al. (2010) and recently validated by Simoneit et al. (2011). Relevant criteria of the study design such as sample size, the involvement of control groups (i.e., untreated, placebo-treated, treated with another drug), blinding, and randomization were considered. Inclusion criteria to be considered for the meta-analytic investigation were the presence of an untreated control group and the calculation of calving to first service interval (**CFSI**) or calving to conception interval (**CCI**) as well as the respective standard deviations.

Statistical Analysis

Data for each trial meeting those criteria were extracted and analyzed using meta-analysis software Re-

view Manager (version 5.1, 2011; The Nordic Cochrane Centre, Copenhagen, Denmark).

Estimated effect sizes of $PGF_{2\alpha}$ were calculated on CFSI and CCI applying the effect size method (Hedges and Olkin, 1985). The effect size is the difference between the treatment and control groups for the number of days open, divided by the pooled standard deviation (Hedges and Olkin, 1985). A negative effect size indicates that a greater percentage of the treated cows have fewer days open than the untreated cows. Effect sizes were calculated for each study. In addition, an overall effect size, weighted by sample size, was calculated. Because the extracted variables were continuous, a weighted mean difference and 95% CI were calculated for each study outcome (Hedges and Olkin, 1985). Variation in experiment level effect size was assessed with a χ^2 test for heterogeneity (Duffield et al., 2008).

Degree of heterogeneity of results among trials was quantified using the I^2 statistic (Higgins et al., 2003). The I^2 statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance; I^2 was calculated as

$$I^2 = \frac{Q - (k-1)}{Q} \times 100\%,$$

where Q is the χ^2 heterogeneity statistic and k is the number of trials. Negative values of I^2 were made equal to zero; consequently, I^2 was between 0 and 100%. A value greater than 50% may be considered indicative of substantial heterogeneity (Duffield et al., 2008).

Because significant heterogeneity was found, results were reported using the random effects model and potential causes of the heterogeneity were sought. A random-effects meta-analysis assumes a normal distribution of study effects. Sources of heterogeneity of response were explored using subgroup analysis. Subgroup analysis was prespecified and conducted by considering aspects of study design such as the method of allocation to treatment and control groups as well blinding. In this context, subgroups A and B represented the randomized and the nonrandomized trials, respectively. Blinded trials were allocated to subgroup C, whereas subgroup D included those not blinded. In addition, subgroups were formed according to the presence of statistically significant effects and the applied $PGF_{2\alpha}$ derivative—dinoprost (subgroup E), cloprostenol (subgroup F), or tiaprost (subgroup G). Because the randomized trials were those showing statistically significant effects, subgroups A and B additionally reflected the trials showing significant and nonsignificant effects, respectively.

To investigate the possible effect of large weighting on the summary estimated effect obtained from the meta-analysis, a sensitivity analysis was performed considering the effect of a $PGF_{2\alpha}$ treatment on CFSI and CCI by eliminating the studies with the largest weights one by one.

Forest plots were used to visually display the estimated effect size ($\mathbf{Z}$), 95% CI, and study weights. The presence of publication bias was investigated graphically using funnel plots, in which the size of effect for each treatment and control group comparison was plotted against its standard error and the resulting plot observed for deficiencies in predicted funnel shape.

RESULTS

In total, 4,393 publications were retrieved (Vetseek, 2,369; PubMed, 570; Medline, 565; Animal Production, 889). After excluding duplicates (n = 2,086), 2,307 publications remained.

Specific exclusion criteria were defined to exclude studies that were not written in English or German (n = 905) or did not focus on chronic endometritis (i.e., equal to and after 21 d after parturition; Sheldon et al., 2006) (n = 725). In addition, studies in which the animals received concomitant treatments with medications other than $PGF_{2\alpha}$ were excluded (n = 358). Book chapters (n = 44), case studies (n = 36), review articles (n = 23), and abstracts (n = 1) were excluded. Furthermore, publications describing etiological, epidemiological, microbiological, or nutritional results, clinical symptoms, or diagnostic procedures were rejected (n = 177). Articles not meeting the inclusion criteria due to incorrect indexing, and those not obtainable through the internet, bibliographies, or interlibrary lending services (n = 32) were excluded as well. If multiple publications were retrieved describing the same trial, those containing the least information were regarded as duplicates and excluded (n = 6). In summary, 2,246 indexed articles had to be excluded according to the exclusion criteria, resulting in 61 remaining publications comprising 63 individual trials. Because 4 articles were retrieved through search by hand, a total of 65 publications, comprising 68 trials, met the inclusion criteria. After applying the inclusion criteria (i.e., control group, calculation of median and SD of CFSI and CCI), 6 studies were eligible for meta-analytic investigation. However, because 1 of the 6 studies only provided data concerning CCI, only 5 studies were available for the overall investigation of CFSI.

A summary of the studies used for the various analyses concerning reproductive performance after $PGF_{2\alpha}$

treatment is presented in Table 1. Six trials in 5 eligible studies had 2,596 cows available to assess the effect of a $PGF_{2\alpha}$ treatment in case of chronic endometritis on CCI. Because one of those trials did not calculate CFSI, only 5 trials in 4 studies with 2,510 cows were available to assess the effect of a $PGF_{2\alpha}$ treatment on CFSI. Statistical significance was considered to be 0.05.

Over all the trials analyzed, $PGF_{2\alpha}$ treatment increased the CFSI (Z = 2.12, 95% CI = 0.59 to 15.40; $P = 0.03$) and the CCI (Z = 12.35, 95% CI = 16.41 to 22.61; $P < 0.00001$) compared with the control group, respectively (Table 2).

The 95% CI of individual values (based on SD) varied considerably between studies. For 2 trials (Feldmann et al., 2005; Hirsbrunner et al., 2006), confidence intervals including zero were found, indicating no effect of a $PGF_{2\alpha}$ treatment in case of chronic endometritis. Concerning overall effect sizes, noticeable confidence intervals could be found for either or both reproductive performance parameters in the subgroups A, C, E, and F. In addition, all of those confidence intervals included zero.

Heterogeneity was substantial for CFSI (χ^2 = 1,898.34, df = 4, $P < 0.00001$, $I^2 = 100\%$) and CCI (χ^2 = 37.38, df = 5, $P < 0.00001$, $I^2 = 87\%$). Therefore, random effects models were used and sources of heterogeneity were explored with a subgroup analysis.

Forest plots illustrating the effect of $PGF_{2\alpha}$ on CFSI and CCI are presented in Figure 1. Confidence intervals of the single studies did not always overlap, which indicates further evidence of medium to high heterogeneity between the studies.

The results of all subgroups were heterogeneous for both CCI and CFSI ($P \leq 0.04$), except subgroups B and G for CCI ($P = 0.06$; Table 2). Concerning both outcomes, significant pooled estimates for subgroups including nonrandomized trials (B), nonblinded trials (D), and trials administering tiaprost (G) could be detected. Although expressed numerically in Table 2, Figure 2 shows how the pooled data of 3 (CFSI) or 4 (CCI) randomized studies provided smaller and nonsignificant effect (CFSI: Z = −0.35, $P = 0.73$; CCI: Z = 0.17, $P = 0.87$) than did the 2 studies that were not randomized (CFSI: Z = 8.25, $P < 0.00001$; CCI: Z = 34.36, $P < 0.00001$). Comparing heterogeneity of those 2 subgroups, significant differences concerning CFSI ($P = 0.01$) as well as CCI ($P = 0.04$) were detected. Comparing trials with ($P < 0.05$) and without statistically significant effects of $PGF_{2\alpha}$ treatment, the only trials providing statistical significance were those 2 that did not apply any method of randomization. Hence, the results found by comparing randomized and nonrandomized trials equate to those found through subgroup analysis based on the question of statistical signifi-

Table 1. Summary of studies used for meta-analysis of reproductive performance after therapy of chronic endometritis with $PGF_{2\alpha}$, considering different subgroups

Study	Total cows (no.)	Dose of $PGF_{2\alpha}$ (mg/cow)	Date of treatment 1	Date of treatment 2	Measured outcomes[1]	Included in subgroup[2]						
						A	B	C	D	E	F	G
Steffan et al. (1984)	153	25.0	37 d pp[3]	14 d after treatment 1	CCI, CR, SC	Yes	No	No	Yes	Yes	No	No
Mejía and Lacau-Mengido (2005); trial 1	678	0.75	30 to 50 d pp	20 d after treatment 1 (if necessary)	CCI, CFSI, PR, SC, FSCR	No	Yes	No	Yes	No	No	Yes
Mejía and Lacau-Mengido (2005); trial 2	1,308	0.75	30 to 50 d pp	20 d after treatment 1 (if necessary)	CCI, SC, CFSI, PR,	No	Yes	No	Yes	No	No	Yes
LeBlanc (2003)	316	0.5	20 to 33 d pp	14 d after treatment 1 (if necessary)	CCI, SC, CFSI, PR, FSCR	Yes	No	No	Yes	No	Yes	No
Hirsbrunner et al. (2006)	—	0.15	21 to 35 d pp	14 d after treatment 1	CCI, SC, CFSI,	Yes	No	Yes	No	No	Yes	No
Feldmann et al. (2005)	178	5.0	≥21 d pp	14 d after treatment 1 (if necessary)	CCI, SC, CFSI, PR, FSCR	Yes	No	No	Yes	Yes	No	No

[1]CCI = calving to conception interval; CR = cure rate; SC = services per conception; CFSI = calving to first service interval; PR = pregnancy rate; FSCR = first-service conception rate.

[2]A = randomized trials or trials showing a statistically significant effect; B = nonrandomized trials or trials showing no statistically significant effect; C = blinded trials; D = nonblinded trials; E = trials administering dinoprost; F = trials administering cloprostenol; G = trials administering tiaprost.

[3]Postpartum.

Table 2. Summary of effect sizes (Z), 95% CI, mean differences (MD), between-studies variance (τ^2), ratio of true heterogeneity to total observed variation (I^2), χ^2, and expected variation (df), and *P*-values for each outcome and subgroup

Group analyzed[1]	Z	*P*-value of Z	95% CI	MD	τ^2	I^2 (%)	χ^2	df	*P*-value of df
Calving to first service interval									
All trials	2.12	0.03	0.59, 15.40	8.00	60.88	100	1,898.34	4	<0.00001
A	−0.35	0.73	−23.61, 16.46	−3.58	284.77	97	72.18	2	<0.00001
B	8.25	<0.00001	17.87, 29.00	23.44	16.10	100	477.32	1	<0.00001
C	1.46	0.14	−1.38, 3.38	4.00	NA[2]	NA	NA	NA	NA
D	2.15	0.03	0.79, 17.32	9.05	60.02	100	1,857.68	3	<0.00001
E	0.86	0.39	−10.53, 27.13	8.3	NA	NA	NA	NA	NA
F	−0.68	0.50	−31.65, 15.39	−8.13	283.61	98	65.56	1	<0.00001
G	8.25	<0.00001	17.87, 29.00	8.25	16.10	100	477.32	1	<0.00001
Calving to conception interval									
All trials	12.35	<0.00001	16.41, 22.61	19.51	7.02	87	37.83	5	<0.00001
A	0.17	0.87	−17.80, 21.15	1.67	333.96	89	27.85	3	<0.00001
B	34.36	<0.00001	20.96, 23.50	22.23	0.63	73	3.66	1	0.06
C	−0.46	0.65	−15.91, 9.91	−3.00	NA	NA	NA	NA	NA
D	16.14	<0.00001	18.59, 23.73	21.16	4.23	83	23.62	4	<0.0001
E	−0.54	0.59	−38.31, 21.84	−8.24	362.94	77	4.35	1	0.04
F	0.82	0.41	−13.65, 33.28	9.81	264.11	92	12.06	1	0.0005
G	34.36	<0.00001	20.96, 23.50	22.23	0.63	73	3.66	1	0.06

[1]A = randomized trials or trials showing a statistically significant effect; B = nonrandomized trials or trials showing no statistically significant effect; C = blinded trials; D = nonblinded trials; E = trials administering dinoprost; F = trials administering cloprostenol; G = trials administering tiaprost.

[2]NA: test for heterogeneity was not applicable because only one study administering dinoprost offered data concerning calving to first service interval.

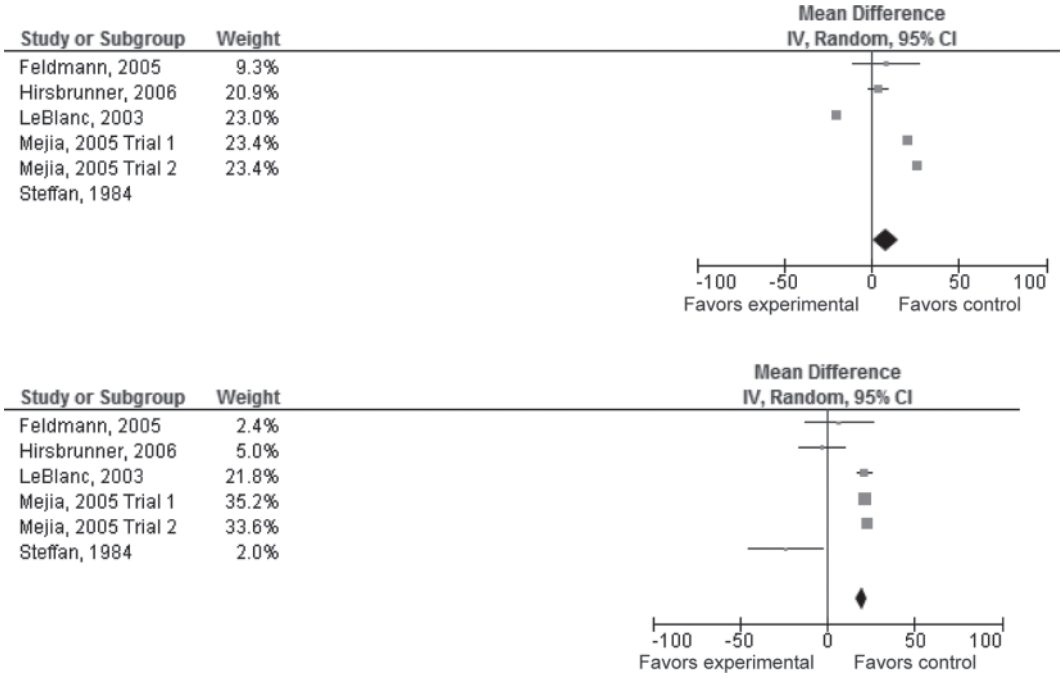

Figure 1. Forest plot of the effect of $PGF_{2\alpha}$ on calving to first service interval (top) and calving to conception interval (bottom) in dairy cows suffering from chronic endometritis. Each square represents the mean effect size for that study. The upper and lower limit of the line connected to the square represents the upper and lower 95% CI for the effect size. The size of the square reflects the relative weighting of the study to the overall effect size estimate with larger squares representing greater weight. The diamond at the bottom represents the 95% CI for the overall estimate. The solid vertical line represents a mean difference of zero or no effect. Squares located on the left side of this line represent studies showing an effect in the group treated with $PGF_{2\alpha}$, whereas squares located on the right side of this line indicate an effect found in the control group. Study or subgroup refers to the first author and year of the publication.

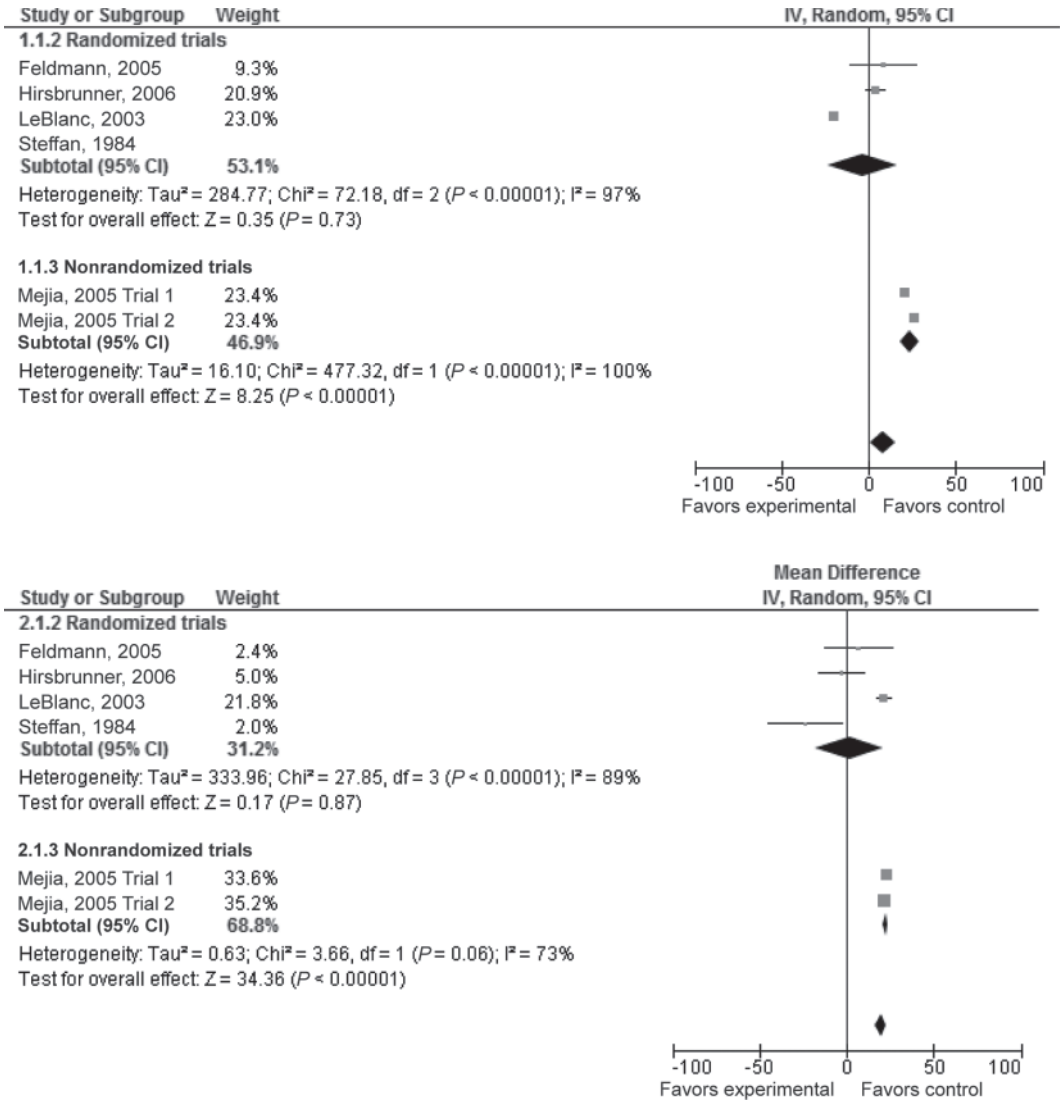

Figure 2. Forest plot of the effect of $PGF_{2\alpha}$ on calving to first service interval (top) and calving to conception interval (bottom) in dairy cows suffering from chronic endometritis. Subgroups were created according to the manner of allocation to treatment and control groups. Study or subgroup refers to the first author and year of the publication.

cance. Table 2 indicates that only the nonblinded trials provided a significant effect for both outcomes (CFSI: Z = 2.15, $P = 0.03$; CCI: Z = 16.14, $P < 0.00001$). Heterogeneity was only significantly different for CCI ($P = 0.0003$; Table 3). Grouping according to the $PGF_{2\alpha}$ derivative applied in the treatment group showed that only those 2 studies having applied tiaprost could find a statistically significant effect concerning both outcomes: CFSI (Z = 8.25, $P < 0.00001$; Figure 3) and CCI (Z = 34.36, $P < 0.00001$; Figure 4). Also, for both outcomes, except CCI associated with tiaprost treatment, significant differences concerning heterogeneity were found within each group that included more than one trial and therefore allowed an investigation of heterogeneity. When investigating subgroup differences, we noticed a significant difference concerning heterogeneity between the subgroups for CFSI ($P = 0.02$; Table 2).

Sensitivity analysis, performed by excluding studies stepwise, showed that excluding the 2 trials included in one study attributed the highest weight and showing consistently outlying results (Mejía and Lacau-Mengido, 2005) did not improve the homogeneity of either outcome ($P = 0.29$ for CFSI; $P = 0.08$ for CCI). However, in contrast to the significant effect sizes (CFSI: Z = 2.12, $P = 0.03$; CCI: Z = 12.35, $P < 0.00001$) favoring the control group, omitting those 2 trials resulted in a somewhat reduced effect (CFSI: Z = −0.35, $P = 0.73$; CCI: Z = 0.17, $P = 0.87$) and a

Table 3. Differences between the 3 subgroups concerning χ^2, the expected variation (df) and its *P*-value, and ratio of true heterogeneity to total observed variation (I^2) regarding calving to first service interval based on the 6 included trials (A to G)

Item	χ^2	df	*P*-value of df	I^2 (%)
Randomization (A vs. B)	6.48	1	0.01	84.6
Blinding (C vs. D)	1.01	1	0.32	0.9
Agent (E vs. F vs. G)	8.37	2	0.02	76.1

reduction of the variance (CFSI: 8.00 vs. −3.58; CCI: 19.51 vs. 1.67). Concerning CFSI, this procedure generated an effect favoring the treatment group instead of the control group (Figure 5). Because those 2 trials were the only ones not randomized, sensitivity analysis showed, in accordance with the results found through subgroup analysis, that randomization was influential in modifying the effect size of $PGF_{2\alpha}$ treatment (Figure 2) as well as the size of variance (Tables 3 and 4). Regarding CCI, an additional exclusion of the study by LeBlanc (2003), which had the third highest weight, led to a significant reduction of heterogeneity ($P = 0.001$; Figure 6). In addition, the effect size shifted from a previously significant effect ($Z = 12.35$, $P < 0.00001$) seen in the control group to a smaller nonsignificant effect ($Z = -0.74$, $P = 0.46$) found in the treatment group. Concerning CFSI, omitting the only trial that showed a positive effect of the $PGF_{2\alpha}$ treatment (LeBlanc, 2003) reduced heterogeneity significantly ($P = 0.03$) and increased the statistically significant effect size ($Z = 7.22$, $P < 0.00001$; Figure 7).

Especially concerning CCI, the presented funnel plots (Figures 8 and 9) suggest a publication bias toward studies with higher standard errors (i.e., usually smaller studies) that reported a positive effect on CCI after a $PGF_{2\alpha}$ treatment.

DISCUSSION

Prostaglandin $F_{2\alpha}$ has been recommended as the treatment of choice for chronic bovine endometritis by numerous authors (e.g., Lewis, 1997 and Azawi, 2008) and is, in some parts of the world (e.g., Argentina), the most common intervention used in veterinary practice (Mejía and Lacau-Mengido, 2005). Despite the reported inconsistency in recent literature concerning the effect (Haimerl et al., 2011), the presented significant negative effect of $PGF_{2\alpha}$; that is, an increase in

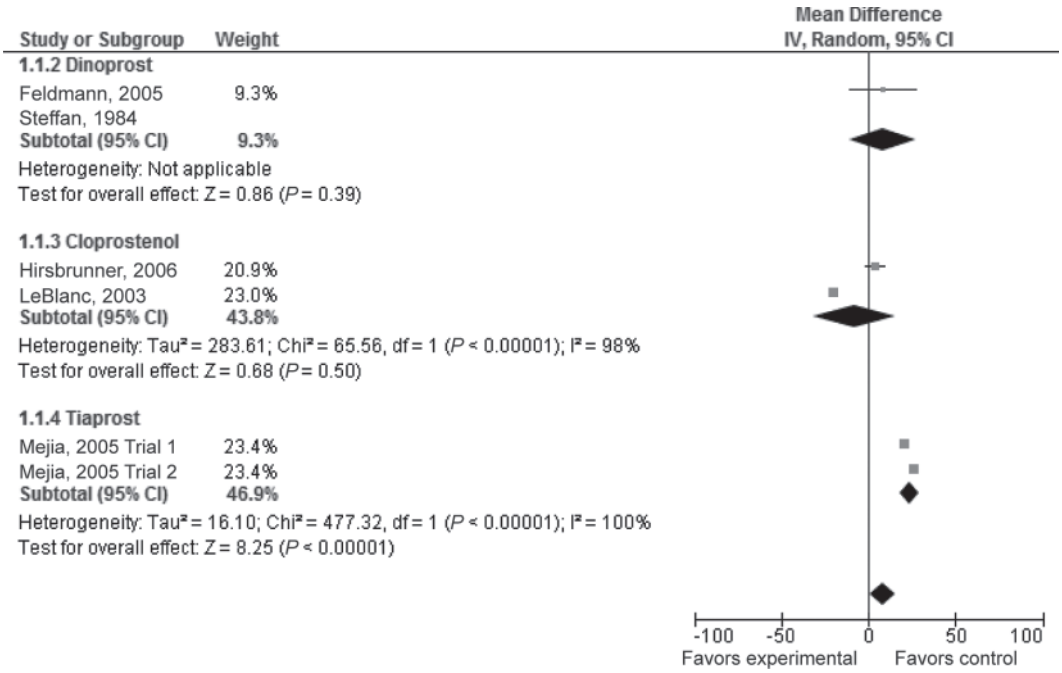

Figure 3. Forest plot of the effect of $PGF_{2\alpha}$ on calving to first service interval in dairy cows suffering from chronic endometritis. Subgroups were created according to the $PGF_{2\alpha}$ derivative applied. Study or subgroup refers to the first author and year of the publication.

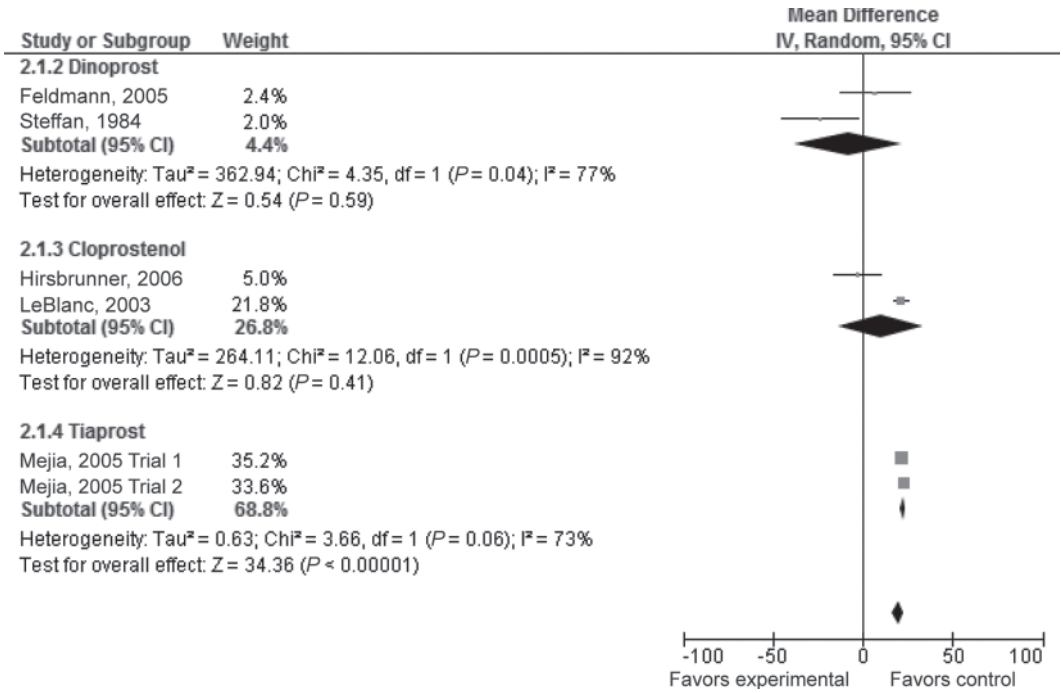

Figure 4. Forest plot of the effect of $PGF_{2\alpha}$ on calving to conception interval in dairy cows suffering from chronic endometritis. Subgroups were created according to the $PGF_{2\alpha}$ derivative applied. Study or subgroup refers to the first author and year of the publication.

CFSI and CCI, was surprising. Yet, the presence of significant heterogeneity in response was anticipated because of the considerable variations not only in the study design but also concerning the outcomes in the existing literature, as previously shown by Haimerl et al. (2011).

In addition to those findings, we revealed further variation concerning study design. One might question

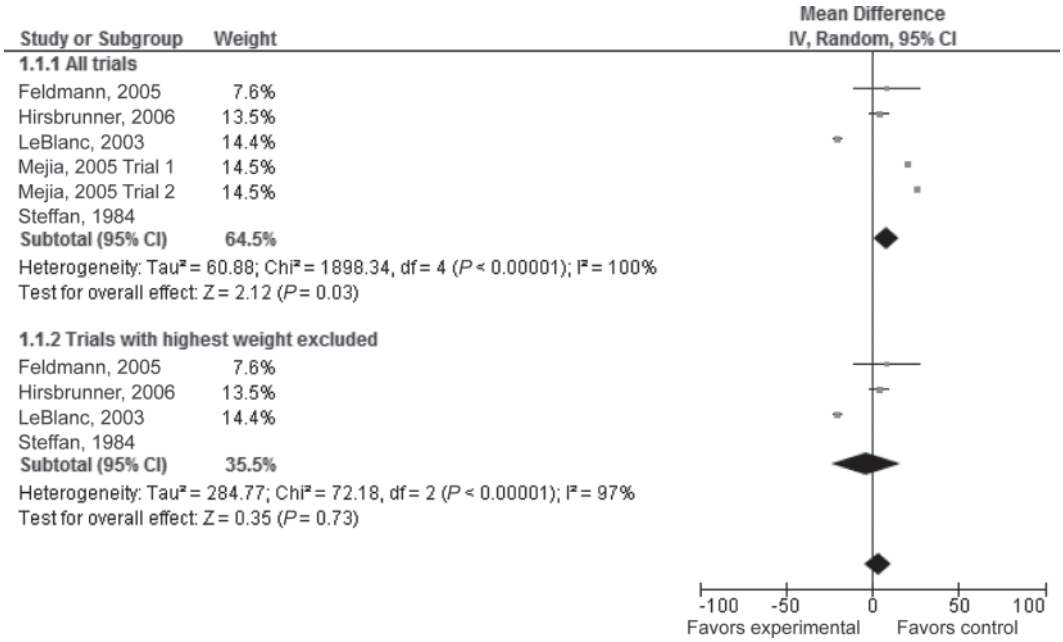

Figure 5. Forest plot of the effect of $PGF_{2\alpha}$ on calving to first service interval in dairy cows suffering from chronic endometritis. Results of sensitivity analysis after the 2 trials with the highest weight (Mejía and Lacau-Mengido, 2005) had been excluded. Study or subgroup refers to the first author and year of the publication.

Table 4. Differences between the 3 subgroups concerning χ^2, the expected variation (df) and its *P*-value, and ratio of true heterogeneity to total observed variation (I^2) regarding calving to conception interval based on the 6 included trials (A to G)

Item	χ^2	df	*P*-value of df	I^2 (%)
Randomization (A vs. B)	4.26	1	0.04	76.5
Blinding (C vs. D)	12.95	1	0.0003	92.3
Agent (E vs. F vs. G)	5.00	2	0.08	60.0

whether the collective analysis of 6 studies with 3 different $PGF_{2\alpha}$ derivatives of variable dosages is meaningful (Table 1). However, every study, except the one by Feldmann et al. (2005), applied the derivative according to the respective product information. Feldmann et al. (2005) based their study on the thesis by Tenhagen (2001) administering 25 mg of dinoprost but reported a dosage of only 5 mg of dinoprost. Therefore, we assume that this dose was reported erroneously. Differences between dinoprost, DL-cloprostenol, and D-cloprostenol concerning their effect on uterine motility have been described (Hirsbrunner et al., 1998). Notwithstanding, the main effect—luteolysis—followed by the induction of estrus and improving the immune response in cows with a responsive corpus luteum is present (Kasimanickam et al., 2005). Considering these pharmacological effects and assuming that the dosage reported by Feldmann et al. (2005) was 25 mg, every study applied the derivative at the dosage recommended by the manufacturer. Hence, one might hypothesize that, in every case, the drug used was effective.

Subgroup analysis based on the chosen groups did not identify one exclusive source of heterogeneity. Instead, within every subgroup, except subgroups B, E, and G concerning CCI, significant heterogeneity, if measurable, was still apparent. However, concerning CCI, in the subgroup consisting of nonrandomized trials, no significant heterogeneity could be found. Both the nonrandomized and randomized trials displayed significant heterogeneity when investigating CFSI. However, significant differences concerning heterogeneity between the 2 subgroups could be found for both reproductive performance parameters. Thus, one might hypothesize that the heterogeneity found in both subgroups has different sources.

Weights are assigned to each trial's result according to the precision of its treatment effect estimate. Trials that have more precise estimates are more heav-

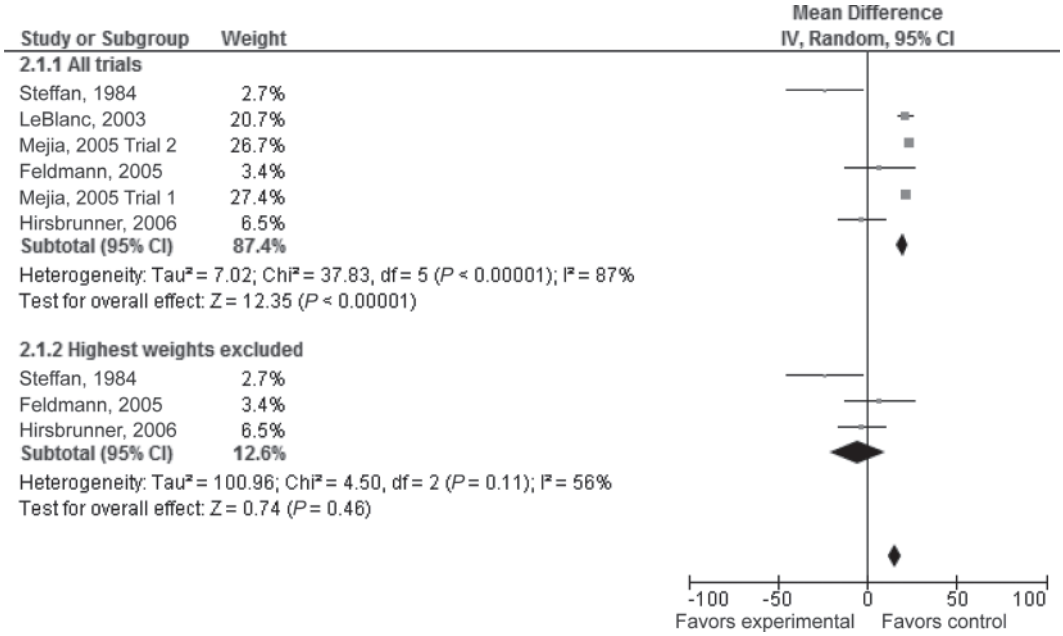

Figure 6. Forest plot of the effect of $PGF_{2\alpha}$ on calving to conception interval in dairy cows suffering from chronic endometritis. Results of sensitivity analysis after the 3 trials with the highest weight (LeBlanc, 2003; Mejía and Lacau-Mengido, 2005) had been excluded. Study or subgroup refers to the first author and year of the publication.

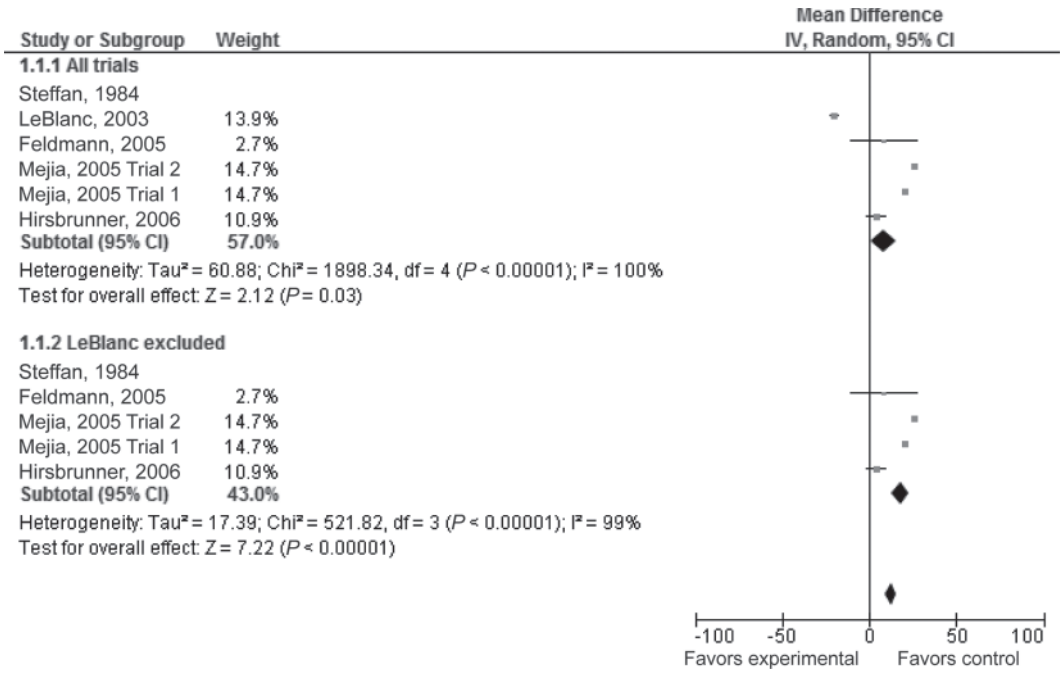

Figure 7. Forest plot of the effect of $PGF_{2\alpha}$ on calving to first service interval in dairy cows suffering from chronic endometritis. Results of sensitivity analysis after the only trial showing an effect in the treatment group (LeBlanc, 2003) had been excluded. Study or subgroup refers to the first author and year of the publication.

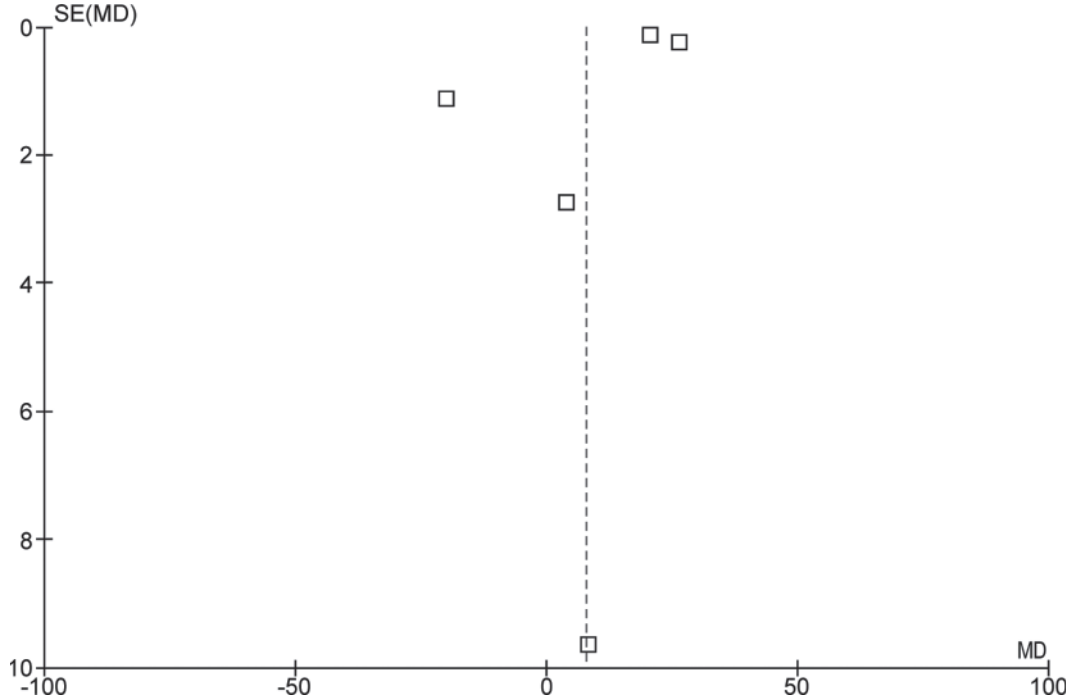

Figure 8. Funnel plot of $PGF_{2\alpha}$ effect on calving to first service interval for assessing publication bias. MD = mean difference; SE of MD = standard error of the standardized mean difference. The vertical dashed line represents the mean effect size. Publication bias may be present if an unequal number of studies (particularly smaller weight studies) are plotted on one side of the line.

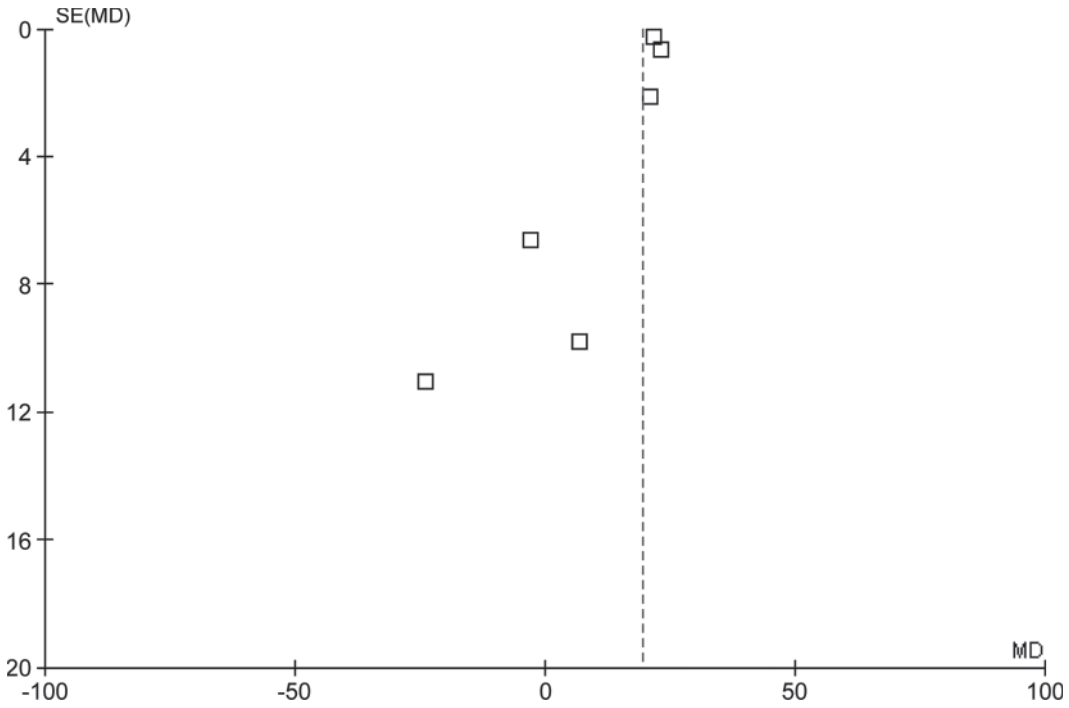

Figure 9. Funnel plot of $PGF_{2\alpha}$ effect on calving to conception interval for assessing publication bias. MD = mean difference; SE of MD = standard error of the standardized mean difference. The vertical dashed line represents the mean effect size. Publication bias may be present if an unequal number of studies (particularly smaller weight studies) are plotted on one side of the line.

ily weighted. Thus, one might expect that those trials having been assigned the highest weights should be those of highest quality and therefore provide the most reliable results. High-quality trials use blinding and random assignment of patients (Cleophas and Zwinderman, 2007). The precision of the estimates, however, is primarily determined by the number of probands. The 2 trials by Mejía and Lacau-Mengido (2005), for example, which were attributed the highest weight, did not apply any method of randomization but included a large number of cows (trial 1, n = 678; trial 2, n = 1,308). Nevertheless, the meta-analysis literature contains multiple examples of simple data pooling without weighting, which has long been known to be even less appropriate than simple weighting by sample size (Barker and Carter, 2005).

Some researchers (Lam and Kennedy, 2005; Willich, 2006) suggested that low-quality studies tend to show higher treatment effects. Our results provide supporting evidence for this observation. However, our analysis comprises too few trials to formulate valid conclusions concerning the relationship between sample size and quality. Also, other reviews (Kunz and Oxman, 1998) do not support this suggestion.

According to Cleophas and Zwinderman (2007), an I^2 value >50% is often used as a cutoff for heterogeneity. The high I^2 values of CFSI and CCI of 100 and 87%, respectively, not only show high inconsistency between the studies but also indicate that most of the variability across studies is due to heterogeneity rather than to chance. Compared with our data, the meta-analysis conducted by Duffield et al. (2008) on the effect of monensin on selected plasma, serum, or blood parameters in lactating dairy cattle revealed considerably lower I^2 values, with the highest being 54% for cholesterol.

Significant heterogeneity suggests that the studies are not measuring a common population effect and that the differences in the study parameters are likely responsible for the varying treatment effect (Oxman et al., 1994). As mentioned by Lam and Kennedy (2005), different schools of thought exist regarding how much homogeneity is required for appropriate comparisons. Lam and Kennedy (2005) also indicate that there is some danger of over-interpreting heterogeneity. Hence, heterogeneity may occur by chance and is an important possibility when a clinical explanation is not found or when the heterogeneity is clinically irrelevant. However, a great deal of uniformity among the results of independently performed studies is not necessarily good (Riegelman, 2005). Hence, a low heterogeneity might indicate consistency in bias rather than consistency in real effects. In contrast, Fourichon et al. (2000) emphasized that an overall summary estimate may not be relevant if it is based on heterogeneous studies. Concerning the

presented meta-analysis, more clinical trials should be conducted and included in an updated analysis to learn more about the sources of heterogeneity.

Subgroup and sensitivity analysis led to interesting findings concerning the summary estimate. Although the investigation of all trials showed statistically significant overall effect sizes concerning CFSI and CCI, significance could not be sustained in the subgroups A (randomized trials), C (blinded trials), E (dinoprost), and F (cloprostenol). Concerning the randomized and blinded trials, this effect may again be explained by the assumption made by Willich (2006)—that trials revealing a stringent study design tend to display lower treatment effects. To support this statement, Figure 2 illustrates how the pooled data of 3 (CFSI) or 4 (CCI) randomized studies provided a smaller and nonsignificant effect than did 2 studies that did not conduct randomization. Surprisingly, in the case of CFSI, excluding nonrandomized trials resulted in a positive effect of the treatment instead of control group. Obviously, the summary result is mainly determined by the nonrandomized studies.

To determine the precision of the effect estimates, the 95% CI of individual values (based on SD) was calculated. Concerning overall effect sizes, wide confidence intervals were found for CFSI or CCI in the subgroups A (CFSI, −23.61, 16.46; CCI, −17.80, 21.15), C (CFSI, −1.38, 3.38; CCI, −15.91, 9.91), E (CFSI, −10.53, 27.13; CCI, −38.31, 21.84) and F (CFSI, −31.65, 15.39; CCI, −13.65, 33.28). These findings are surprising as they imply that trials with a sound study design did not generate a precise effect estimate. Also, in comparison to other meta-analyses, the confidence intervals calculated in our study were wider. Fourichon et al. (2000), for example, investigated the effect of various diseases on reproduction in dairy cows based on results from 70 papers. Of all the diseases considered, confidence intervals were mainly narrow, except for days open in all cows concerning the effect of retained placenta (3.0, 48.8) and cystic ovaries (−22.8, 47.0), respectively. The meta-analysis conducted by Duffield et al. (2008) on the effect of monensin, however, revealed large variations regarding confidence intervals, ranging from −149.5, −78.3 for BHBA to −0.019, 0.037 for calcium.

However, confidence intervals for meta-analyses do not depend only on the precision of the individual study estimates, but also on the number of studies combined. Because the width of the confidence interval usually decreases as more studies are added to a meta-analysis (Borenstein et al., 2009), the limited number of trials included, especially in the single subgroups, contributed to the magnitude of the confidence intervals. This could account for the mainly narrow confidence intervals found by Fourichon et al. (2000) because that meta-analysis included 70 papers. However, meta-analytic investigations conducted by Burton and Lean (1995) revealed considerably smaller 95% CI (−0.44, −0.01) concerning CCI after a PGF treatment of cows with an abnormal puerperium, although the number of trials included was similar to that in our analysis (n = 7). In addition, for random-effects models, as used in this meta-analysis, precision will decrease with increasing heterogeneity and confidence intervals will widen correspondingly. Furthermore, all of those confidence intervals include zero, which leads to the assumption that the mean effect size does not differ significantly from zero (Eisend, 2004). This, in turn, indicates that $PGF_{2\alpha}$ treatment has no effect in the case of chronic endometritis.

The 95% CI of individual values varied considerably between studies, which is further evidence for considerable heterogeneity. Because I^2 reflects the extent of overlap of confidence intervals, which is not dependent on the actual location or the spread of the true effects (Borenstein et al., 2009), the high values for I^2 are consistent with the limited overlapping of the single confidence intervals. Beyond that, for 2 trials (Feldmann et al., 2005; Hirsbrunner et al., 2006), confidence intervals including zero were found.

Grouping according to the $PGF_{2\alpha}$ derivative showed that only those 2 studies having applied tiaprost could find a statistically significant effect concerning both outcomes. This was not surprising, however, because this subgroup consisted exclusively of the 2 trials executed by Mejía and Lacau-Mengido (2005), which were those 2 not randomized and solely showing significant effect estimates. Therefore, it remains unclear if the statistical significance of the effect estimate is due to the derivative applied or the differences in study design.

Another essential objective of a meta-analysis is to search for potential bias, which may be introduced by deficient literature search or selection (Barker and Carter, 2005). As shown by several authors (Egger and Smith, 1998; Montori et al., 2000; Song et al., 2000), one of the most frequently occurring events is the absence of studies with "negative" outcomes, which is due to a number of reasons. "Negative" trials, or those that fail to show superiority of the experimental treatment, tend to be left unpublished (publication bias) or published more slowly (pipeline bias) and are less frequently retrieved by the searching process (retrieval bias). Another problem is that "positive" trials are more likely to be published more than once, not always in a readily identifiable way. Publication bias can have an important effect on the outcomes of meta-analyses (Scifres et al., 2009). Measuring the effects of publication bias quantitatively can be difficult. Nevertheless, various approaches exist. The most widely applied method for

detecting publication bias in a meta-analysis is the funnel plot, in which the effect size for individual trials is plotted against a measure of the precision of the treatment effect estimate, such as the size of the trial (Egger et al., 1997). In the present meta-analysis, the measure of precision was the standard error. When a funnel plot is used to study an unbiased sample of trials, the observed effect sizes should range symmetrically around the true effect size, which will be most accurately estimated by the largest trials, thus giving a symmetrical, funnel-shaped plot (Barker and Carter, 2005). In the conducted meta-analysis, however, such a funnel shape was not detected for either CFSI or CCI. Especially with regard to CCI, a strong asymmetry was found toward studies with higher standard error, reporting a positive effect on CCI after a $PGF_{2\alpha}$ treatment. Such an asymmetry is often caused by small trials not providing much evidence of efficacy and therefore being infrequently published or retrieved, leading to larger estimates of treatment effect being observed in smaller trials (Barker and Carter, 2005). However, difficulties exist in assessing the shape of a funnel plot when the number of trials included is small (Tang and Liu, 2000; Sterne and Egger, 2001).

In addition, small trials tend to overestimate treatment effects because of methodological differences (either design flaws or more rigorous implementation of treatment), causing a false appearance of publication bias when none actually exists (Schulz et al., 1995; Stuck et al., 1998). This "small-study effect" (Kjaergard et al., 2001) leads to the paradox that increasing success in locating small, unpublished trials may actually bias a meta-analysis toward an overestimation of treatment effect (Barker and Carter, 2005). Several researchers (Sterne et al., 2000; Sutton et al., 2000) have estimated that approximately 25 to 50% of meta-analyses may be affected by publication bias and that treatment effects were overestimated by up to 47% as a result.

Beyond that, the shape of the funnel plot is extremely sensitive to the method used to express treatment effect and precision. Thus, applying the fixed-effects method in the present meta-analysis might have led to different results. However, a fixed-effects analysis is only appropriate when the trials are so similar in design and execution that one can assume they address the same research hypothesis, using the same methods and treatments in the same population, and when the formal test of homogeneity demonstrates a lack of heterogeneity between the trial results (Barker and Carter, 2005). Thus, due to the considerable heterogeneity found in this study, a random-effects analysis was conducted assuming that the true effect size varied from study to study and the summary effect was our estimate of the mean of the distribution of effect sizes (Borenstein et al., 2009). With reference to heterogeneity, Terrin et al. (2003) and Tang and Liu (2000) demonstrated that an asymmetrical funnel plot is only related to publication bias if the trials included are homogeneous. Hence, the results found concerning publication bias have to be considered carefully.

The last step in performing a meta-analysis is to assess its robustness with the help of a sensitivity analysis; that is, checking the stability of the results in relation to the different ways the analysis could reasonably be performed (Olkin, 1999). Thus, sensitivity analysis strengthens the estimation and interpretation of effects obtained from several studies (Fourichon et al., 2000), and conclusions derived from meta-analysis that are robust to multiple sensitivity analyses gain credibility in application to general practice (Barker and Carter, 2005). Sensitivity analyses can be performed in several ways (Barker and Carter, 2005). By excluding the 2 nonrandomized trails due to the highest weights (Mejía and Lacau-Mengido, 2005), we simultaneously investigated the possible effect of large-weight and low-quality studies on the summary estimated effect obtained from the meta-analysis. This approach led to a considerable reduction of effect estimates, accompanied by the loss of statistical significance. Remarkably, omitting those 2 trials resulted in an effect on CFSI seen in treatment instead of the control group. These results clearly illustrate that the inclusion criteria applied are an important factor in determining the summary result. According to Cleophas and Zwinderman (2007), this observation leads to the conclusion that the present meta-analysis lacks robustness and to the necessity of more clinical trials to obtain more reliable results.

Meta-analysis methods have been developed to calculate the summary estimates of several studies to integrate the available information from different source studies and to examine the potential origin of heterogeneity of the results. However, the findings of a meta-analysis must be considered with caution because these methods have been developed to summarize the results of randomized, controlled trials. Thus, extending its use to observational studies must account for the variability of study designs and of populations used (Fourichon et al., 2000). Therefore, one might favor the strict exclusion of nonrandomized trials. However, as pointed out by Khan et al. (1996), it may be worthwhile to not completely reject the studies with lower methodology because they can be used for assessing sensitivity. In addition, not many meta-analyses have been carried out in the field of veterinary medicine, for several reasons (Holmes, 2004; Arlt et al., 2005), one of which is a shortage of randomized, controlled studies in veterinary medicine (Arlt and Heuwieser, 2005). Therefore, if insufficient randomized, controlled trials

are available, trials of lower evidence levels should be integrated to enhance the power of the meta-analysis.

CONCLUSIONS

Our meta-analysis did not reveal an improvement of reproductive performance of cows with endometritis after treatment with $PGF_{2\alpha}$. Therefore, we encourage reconsideration of $PGF_{2\alpha}$ as a routine treatment for cows with chronic endometritis. This is particularly important because a blanket treatment of $PGF_{2\alpha}$ could be perceived by the public as unjustified hormone use. Furthermore, we conclude that there is a shortage of comparable and high-quality studies investigating reproductive performance after $PGF_{2\alpha}$ treatment of cows with chronic endometritis. Additional confirmatory trials are necessary to shed more light on the contradictory results published, to assess the practical value of the treatment, and to identify the sources of heterogeneity. The latter may reveal important aspects concerning study quality and applicability of the results in livestock research.

REFERENCES

Arlt, S., V. Dicty, and W. Heuwieser. 2010. Evidence-based medicine in canine reproduction: Quality of current available literature. Reprod. Domest. Anim. 45:1052–1058.

Arlt, S., M. Drillich, J. Kluth, and W. Heuwieser. 2005. With a critical viewpoint. Prakt. Tierarzt. 86:206–207.

Arlt, S., and W. Heuwieser. 2005. Evidence based veterinary medicine. Dtsch. Tierarztl. Wochenschr. 112:146–148.

Arlt, S., W. Padberg, M. Drillich, and W. Heuwieser. 2009. Efficacy of homeopathic remedies as prophylaxis of bovine endometritis. J. Dairy Sci. 92:4945–4953.

Azawi, O. I. 2008. Postpartum uterine infection in cattle. Anim. Reprod. Sci. 105:187–208.

Barker, F. G. 2nd, and B. S. Carter. 2005. Synthesizing medical evidence: Systematic reviews and metaanalyses. Neurosurg. Focus 19:E5.

Borenstein, M., L. V. Hedges, J. P. T. Higgins, and H. R. Rothstein. 2009. Subgroup analyses. Pages 149–186 in Introduction to Meta-Analysis. John Wiley & Sons, Chichester, UK.

Burton, N. R., and I. J. Lean. 1995. Investigation by meta-analysis of the effect of prostaglandin $F_{2\alpha}$ administered post partum on the reproductive performance of dairy cattle. Vet. Rec. 136:90–94.

Cleophas, T. J., and A. H. Zwinderman. 2007. Meta-analysis. Circulation 115:2870–2875.

Cochrane, A. L. 1972. Effectiveness and Efficiency: Random Reflections on Health Services. Royal Soc. Med. Press., London, UK.

Dubuc, J., T. F. Duffield, K. E. Leslie, J. S. Walton, and S. J. LeBlanc. 2011. Randomized clinical trial of antibiotic and prostaglandin treatments for uterine health and reproductive performance in dairy cows. J. Dairy Sci. 94:1325–1338.

Duffield, T. F., A. R. Rabiee, and I. J. Lean. 2008. A meta-analysis of the impact of monensin in lactating dairy cattle. Part 1. Metabolic effects. J. Dairy Sci. 91:1334–1346.

Egger, M., G. Davey Smith, M. Schneider, and C. Minder. 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315:629–634.

Egger, M., and G. D. Smith. 1998. Bias in location and selection of studies. BMJ 316:61–66.

Eisend, M. 2004. Metaanalyse–Einführung und kritische Diskussion. Page 44 in Wirtschaftswissenschaft. Freie Universität Berlin, Berlin, Germany.

Falkenberg, U., and W. Heuwieser. 2005. Influence of time of initiation of a prostaglandin $F_{2\alpha}$ protocol in dairy cows with puerperal endometritis. Dtsch. Tierarztl. Wochenschr. 112:252–256. (In German)

Feldmann, M., S. G. E. Tenhagen, and M. Hoedemaker. 2005. Treatment of chronic bovine endometritis and factors for treatment success. Dtsch. Tierarztl. Wochenschr. 112:10–16. (In German).

Fourichon, C., H. Seegers, and X. Malher. 2000. Effect of disease on reproduction in the dairy cow: A meta-analysis. Theriogenology 53:1729–1759.

Gilbert, R. O., and W. S. Schwark. 1992. Pharmacologic considerations in the management of peripartum conditions in the cow. Vet. Clin. North Am. Food Anim. Pract. 8:29–56.

Glass, G. V. 1976. Primary, secondary, and meta-analysis of research. Educ. Res. 10:3–8.

Haimerl, P., S. Arlt, and W. Heuwieser. 2011. Evidence-based medicine: Quality and comparability of clinical trials investigating the efficacy of prostaglandin $F_{2\alpha}$ for the treatment of bovine endometritis. J. Dairy Res. 79:287–296.

Hedges, L. V., and I. Olkin. 1985. Statistical Methods for Meta-Analysis. Academic Press, New York, NY.

Henry, D. A., and A. Wilson. 1992. Meta-analysis. Part 1: An assessment of its aims, validity and reliability. Med. J. Aust. 156:31–38.

Higgins, J. P., S. G. Thompson, J. J. Deeks, and D. G. Altman. 2003. Measuring inconsistency in meta-analyses. BMJ 327:557–560.

Hirsbrunner, G., H. W. Burkhardt, and A. Steiner. 2006. Effects of a single administration of prostaglandin $F_{2\alpha}$, or a combination of prostaglandin $F_{2\alpha}$ and prostaglandin E2, or placebo on fertility variables in dairy cows 3–5 weeks postpartum, a randomized, double-blind clinical trial. Reprod. Biol. Endocrinol. 4:65.

Hirsbrunner, G., U. Kupfer, H. Burkhardt, and A. Steiner. 1998. Effect of different prostaglandins on intrauterine pressure and uterine motility during diestrus in experimental cows. Theriogenology 50:445–455.

Holmes, M. 2004. Evidence-based veterinary medicine. 2. Identifying information needs and finding the evidence. In Pract. 26:96–102.

Kasimanickam, R., T. F. Duffield, R. A. Foster, C. J. Gartley, K. E. Leslie, J. S. Walton, and W. H. Johnson. 2005. The effect of a single administration of cephapirin or cloprostenol on the reproductive performance of dairy cows with subclinical endometritis. Theriogenology 63:818–830.

Khan, K. S., S. Daya, and A. Jadad. 1996. The importance of quality of primary studies in producing unbiased systematic reviews. Arch. Intern. Med. 156:661–666.

Kjaergard, L. L., J. Villumsen, and C. Gluud. 2001. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann. Intern. Med. 135:982–989.

Kunz, R., and A. D. Oxman. 1998. The unpredictability paradox: Review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 317:1185–1190.

L'Abbé, K. A., A. S. Detsky, and K. O'Rourke. 1987. Meta-analysis in clinical research. Ann. Intern. Med. 107:224–233.

Lam, R. W., and S. H. Kennedy. 2005. Using meta-analysis to evaluate evidence: Practical tips and traps. Can. J. Psychiatry 50:167–174.

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

LeBlanc, S. 2003. Field study of the diagnosis and treatment of clinical endometritis in dairy cattle. Cattle Pract. 11:255–261.

LeBlanc, S. J. 2008. Postpartum uterine disease and dairy herd reproductive performance: A review. Vet. J. 176:102–114.

LeBlanc, S. J., T. F. Duffield, K. E. Leslie, K. G. Bateman, G. P. Keefe, J. S. Walton, and W. H. Johnson. 2002. Defining and diagnosing postpartum clinical endometritis and its impact on reproductive performance in dairy cows. J. Dairy Sci. 85:2223–2236.

Lefebvre, R. C., and A. E. Stock. 2012. Therapeutic efficiency of antibiotics and prostaglandin $F_{2\alpha}$ in postpartum dairy cows with clini-

cal endometritis: An evidence-based evaluation. Vet. Clin. North Am. Food Anim. Pract. 28:79–96.
Lewis, G. S. 1997. Uterine health and disorders. J. Dairy Sci. 80:984–994.
Macmillan, K. L., A. M. Day, V. K. Taufa, H. V. Henderson, and P. A. Allison. 1987. Some effects of injecting a prostaglandin F2-alpha (Lutalyse) during the post partum period on the subsequent fertility of dairy cows. Pages 65–68 in Proc. 47st Annual NZSAP Conf. New Zealand Society of Animal Production (NZSAP), Palmerston North, New Zealand.
Mejía, M. E., and I. M. Lacau-Mengido. 2005. Endometritis treatment with a $PGF_{2\alpha}$ analog does not improve reproductive performance in a large dairy herd in Argentina. Theriogenology 63:1266–1276.
Montori, V. M., M. Smieja, and G. H. Guyatt. 2000. Publication bias: A brief review for clinicians. Mayo Clin. Proc. 75:1284–1288.
Olkin, I. 1999. Diagnostic statistical procedures in medical meta-analyses. Stat. Med. 18:2331–2341.
Olson, J. D. 1996. Metritis/endometritis: Medically sound treatments. Pages 8–14 in 29th Annual Conference of the American Association of Bovine Practitioners, San Diego, CA. Am. Assoc. Bovine Pract., Auburn, AL.
Oxman, A. D., D. J. Cook, and G. H. Guyatt. 1994. Users' guides to the medical literature. VI. How to use an overview. Evidence-based medicine working group. JAMA 272:1367–1371.
Plaizier, J. C., K. D. Lissemore, D. Kelton, and G. J. King. 1998. Evaluation of overall reproductive performance of dairy herds. J. Dairy Sci. 81:1848–1854.
Riegelman, R. K. 2005. Meta-analysis. Pages 99–115 in Studying a Study and Testing a Test. R. K. Riegelman, ed. Lippincott Williams & Wilkins, Philadelphia, PA.
Schulz, K. F., I. Chalmers, R. J. Hayes, and D. G. Altman. 1995. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273:408–412.
Scifres, C. M., J. D. Iams, M. Klebanoff, and G. A. Macones. 2009. Metaanalysis vs. large clinical trials: Which should guide our management? Am. J. Obstet. Gynecol. 200:484.e1–484.e5. http://dx.doi.org/10.1016/j.ajog.2008.09.873.
Sheldon, I. M. 2009. Defining postpartum uterine disease and the mechanisms of infection and immunity in the female reproductive tract in cattle. Biol. Reprod. 81:1025–1032.
Sheldon, I. M., G. S. Lewis, S. LeBlanc, and R. O. Gilbert. 2006. Defining postpartum uterine disease in cattle. Theriogenology 65:1516–1530.
Simoneit, C., W. Heuwieser, and S. Arlt. 2011. Evidence-based medicine in bovine, equine and canine reproduction: Quality of current literature. Theriogenology 76:1042–1050.
Song, F., A. J. Eastwood, S. Gilbody, L. Duley, and A. J. Sutton. 2000. Publication and related biases. Health Technol. Assess. 4:1–115.
Steffan, J., M. Agric, S. Adriamanga, and M. Thibier. 1984. Treatment of metritis with antibiotics or prostaglandin $F_{2\alpha}$ and influence of ovarian cyclicity in dairy cows. Am. J. Vet. Res. 45:1090–1094.
Sterne, J. A. C., and M. Egger. 2001. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. J. Clin. Epidemiol. 54:1046–1055.
Sterne, J. A. C., D. Gavaghan, and M. Egger. 2000. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. J. Clin. Epidemiol. 53:1119–1129.
Stuck, A. E., L. Z. Rubenstein, and D. Wieland. 1998. Bias in meta-analysis detected by a simple, graphical test. Asymmetry detected in funnel plot was probably due to true heterogeneity. BMJ 316:469–471.
Sutton, A. J., S. J. Duval, R. L. Tweedie, K. R. Abrams, and D. R. Jones. 2000. Empirical assessment of effect of publication bias on meta-analyses. BMJ 320:1574–1577.
Tang, J. L., and J. L. Liu. 2000. Misleading funnel plot for detection of bias in meta-analysis. J. Clin. Epidemiol. 53:477–484.
Tenhagen, G. E. S. 2001. Vergleich verschiedener Behandlungen der chronischen Endometritis des Rindes unter Berücksichtigung von Faktoren des Behandlungserfolges. Klinik für Geburtshilfe und Gynäkologie des Rindes. Tierärztliche Hochschule Hannover, Hannover, Germany.
Terrin, N., C. H. Schmid, J. L. Griffith, R. B. D'Agostino, and H. P. Selker. 2003. External validity of predictive models: A comparison of logistic regression, classification trees, and neural networks. J. Clin. Epidemiol. 56:721–729.
Willich, S. 2006. Randomisierte kontrollierte Studien—Pragmatische Ansätze erforderlich. Dtsch. Arztebl. 103:2524–2529.
Wilson, A., and D. A. Henry. 1992. Meta-analysis. Part 2: Assessing the quality of published meta-analyses. Med. J. Aust. 156:173–174, 177–180, 184–177.
Young, I. M., D. B. Anderson, and R. W. Plenderleith. 1984. Increased conception rate in dairy cows after early post partum administration of prostaglandin $F_{2\alpha}$ THAM. Vet. Rec. 115:429–431.