

# POPULATION MEDICINE NEWS

Epidemiology, Preventive Medicine, Public Health  
Production Medicine, Computer Applications in Vet Med

October 14, 1991 4(24)

## Critical Evaluation of the Scientific Literature

### Case studies in allocation: the good, the bad, and the ugly

*We are sometimes lulled into believing that the "refereed" literature has been carefully scrutinized for scientific blunders such as grossly unsuitable allocation schemes. Not only is such faith misplaced, it can lead us to make bad practice decisions.*

Considerable debate has centered on the efficacy of prostaglandin products in the treatment of postpartum reproductive disease in dairy cattle. The debate was brought to the fore by a recent article by Gianvill and Dobson (1) in which a beneficial effect was not observed for including prostaglandin injection in the treatment scheme for cows diagnosed with reproductive disease between days 14 and 28 postpartum. Whether we feel comfortable with this finding or not, we are obliged to evaluate the article critically, including, as our present interest dictates, the allocation scheme.

We first turn our attention to the section in the Methods in which the allocation scheme is described: "In all, 180 animals were allocated alternately to treatment with 25 mg dinoprost tromethamine (Lutalyse; Upjohn) intramuscularly 14 to 28 days after calving, or to the control group." This we recognize as systematic (i.e., every-other-cow) allocation which adequately randomizes the animals into the treatment groups if faithfully followed.

The candidate cows were defined as so: "The cows were selected on the basis of having a post parturient condition recognized as being capable of delaying conception..." We are told in the results that 61 % of the 180 cows in the trial were diagnosed as having endometritis and that 22.2% had experienced assisted deliveries. This heterogeneity of study animals causes us some concern since we can well imagine unequal aggregation of, say, pyometra into one group or the other. We are also concerned that animals of a variety of parities on 4 different farms were included in the study since we recognize such diversity as fertile ground for potential confounding (eg., if mostly heifers had been treated at Farm C and mostly older cows at Farm B). Our concerns are allayed by the section in the Methods: "In order to take a valid statistical analysis of the effectiveness of treatment, the cows were then paired. Each pair of cows consisted of a treated cow and an untreated cow, on the same farm and of the same parity, which had experienced the same clinical problem." This we recognize as partial restriction.

Given that a critical reader cannot have blind faith in the researcher's faithfulness in following a specified allocation regimen, we must look for clues. First, the allocation scheme set out would be expected to yield equal numbers of treated cows and controls overall and on each of the 4 farms. Table 1 was reconstructed from the data presented in article (a typographical error presented a minor challenge). The numbers are completely consistent with the allocation scheme set forth. We are not yet totally satisfied, however.

Next we look for descriptive data on cases and controls to see if the animals are comparable. The only such information provided is progesterone concen-

*Three questions the critical reviewer should ask about the allocation scheme.*

*1. Is the scheme spelled out in the Methods beyond the empty phrase: "the animals were randomly allocated to treatment group"?*

*2. Are the treatment group sizes equal or approximately so (unless departures are explicitly justified in the scheme)?*

*3. Do the data presented demonstrate that the animals in the different groups were equivalent except for the treatment?*

trations at time of treatment (Table 2). Overall, the % of cows with high progesterone at time of treatment matched perfectly between prostaglandin treated cows and controls. Had we seen a large divergence, we would have concluded that the systematic allocation scheme was not followed faithfully (eg., pyometras might have been preferentially placed in the treated group and cycling cows in the control group). The deviations from perfect balance within herds is well within the realm of chance given that cows were not pair matched on this criterion (you

Table 1

Group sizes reconstructed from Gianvill and Dobson (1).

Herd	Treated*	Control
W	31	31
E	21	21
C	14	14
B	24	24

\* "treated" group received prostaglandin along with other routine treatments such as intrauterine antibiotic infusions; "control" received other treatments only.

Table 2

Cows with high progesterone concentrations at time of treatment in Gianvill and Dobson study. (1)

Herd	Treated	Control
W	19/31 (61%)*	15/31 (48%)
E	12/21 (57%)	8/21 (38%)
C	9/14 (64%)	13/13# (100%)
B	16/24 (66%)	19/24 (79%)
Total	56/90 (62%)	55/89 (62%)

\* No. cows with progesterone concentrations > 2 ng/ml divided by total number of cows. (% with high progesterone)  
# one sample evidently lost from Herd C.

### Table 3

Results from Gianvill and Dobson: (1)

	Treated	Control
Calving-1st serv. (days)	71.9	73.5
Calving-conc. (days)	103.4	104.8
Conception rate: (%) to		
1st service	41.0	40.0
2nd service	45.0	52.0
3rd service	41.0	43.0
Services/conception	2.18	2.14

can use the binomial distribution to confirm this if you are skeptical).

We would have preferred more ambitious statistical analyses than the paired t-tests used by Gianvill and Dobson. Stratified analyses using herd, clinical problem, and progesterone concentration as strata were warranted since the high progesterone cows were not perfectly balanced within herds. However, given the marked similarity of fertility statistics between treated and control groups (Table 3), we doubt that stratified analysis would have been revealing.

We have not discussed issues such as blinding, the potential for measurement bias, or the physiologic basis (or lack of) for treatment at 14 to 28 days, but based on the allocation scheme, we are hard pressed to assail Gianvill and Dobson's conclusion: "that the administration of prostaglandin between 14 and 28 days after calving had no beneficial effect upon the reproductive performance of 'problem cows' from herds in which re-breeding occurred more than 70 days after calving." (1)

One study which has been cited to justify prostaglandin use for routine treatment of postpartum endometritis/metritis is that reported by Steffan. (2) Consider the allocation scheme he laid out in his Methods: "cows affected by metritis were randomly distributed to 1 of the following 3 groups: (1) Antibiotics therapy (n=53)... (2) Prostaglandin F<sub>2α</sub> therapy (n=61 cows)... (3) Control (placebo; n=39)...". Though this paper has been cited as authoritative (3), even a cursory examination should confront one with the inconsistency of "random distribution" with group sizes of 61, 53, and 39. "Random distribution" implies that the 153 study cows should have been equally allocated which would have resulted in group sizes of 51 cows. Even an open-ended probability allocation of 0.33 would not likely (P = .02) have resulted

in a control group of 39 cows. We can imagine numerous reasons for the deviance: cows destined to be culls were used as controls, owner permission was required to keep a cow in the control group, severe cases were withdrawn from the study if allocated to the control group, the control group was an afterthought added after the

other groups were underway, etc. Though we don't know which, if any, of these biases was at work, we are compelled to assume that some sort of allocation bias occurred; and thus, as critical readers, we must add Steffan's paper to that great stack, ~~teetering~~ by now from the accumulated weight of decades of poorly run bovine fertility studies, and labeled INTERESTING BUT NOT VERY USEFUL FOR DECISION-MAKING.

Next consider a paper by Murray, et al (4) which reports a study limited to "cases of endometritis" identified and treated by 7 practitioners. According to the Methods, the cows were "randomly allocated" to 1 of 3 treatment groups: alfaprostal (a synthetic prostaglandin related to PGF<sub>2α</sub>) injection, intrauterine infusion of an antibiotic, or both alfaprostal injection and antibiotic infusion. The group sizes of 114, 94, and 98 are a little disturbing but within the realm of possibility for a randomized trial with an open-ended probability allocation scheme. However, when we look for clues about the characteristics of cows assigned to the 3 groups, we find a very serious problem. As shown in Table 4, the allocation scheme was light-years away from random. Cows judged to have severe endometritis were preferentially placed in the group receiving both antibiotic infusion and alfaprostal injection. Although an informal stratified analysis partly removed this one extreme source of bias, we are forced to wonder what other allocation criteria were used by the 7 clinicians. Any treatment effects that may exist were hopelessly buried in a morass of confounding produced by whimsical allocation criteria. Thus the critical reader is forced to file this paper in the great stack with its fecund relatives.

The road to a final decision on the efficacy of prostaglandin for postpartum reproductive disease will probably be filled with more unsuitably allocated studies, and their appearance in the

### Table 4

Data from Fig 1 of Murray, et al, showing extreme departure from "random allocation" scheme claimed in Methods: (4)

Treatment group	Severity of endometritis		
	Severe	Moderate	Mild
1. Alfaprostal injection	20(18%)*	62(54%)	32(28%)
2. Antibiotic infusion	15(16%)	54(57%)	25(27%)
3. Both 1 and 2	49(50%)	45(46%)	4(4%)

\* number of cows (percent within treatment group)

refered literature should convince us that we can't count on the academic priesthood to interpret the professional literature for us. The Steffan and Murray papers were published in reputable veterinary journals (Am J Vet Res and Vet Rec), yet the reviewers allowed inclusion of "randomly distributed" and "randomly allocated" even when the terms appear to have been blatantly inappropriate. Both papers have been or will be cited as authoritative references in review papers. The only way to avoid being led to bad decisions is to critically review scientific papers for yourself.

#### References:

- Gianvill SF, Dobson H; Effect of prostaglandin treatment on the fertility of problem cows. Vet Rec 123:374-376, 1991).
- Steffan J, Adriamanga S, Thiber M; Treatment of metritis with antibiotics or prostaglandin F<sub>2α</sub> and influence of ovarian cyclicity in dairy cows. Am J Vet Res 45:1090-1094; 1984.
- Wichtel JJ; When and why prostaglandins are used in postpartum cows. Vet Med, June 1991.
- Murray RD, Allison JD, Gard RP; Bovine endometritis: comparative efficacy of alfaprostal and intrauterine therapies, and other factors influencing clinical success; Vet Rec 127:86-90, 1990.

POPULATION MEDICINE NEWS is produced by the Washington State University Field Disease Investigation Unit, Pullman, WA 99164-6610 (509-335-0711) and is supported in part by the WSU Cooperative Extension Service. Viewpoints expressed are not necessarily those of the University or its Administration. This newsletter may be copied by anyone. Editors: Dale Hancock and Susan Waller.