Acute Phase Response to Surgery of Varying Intensity in Horses

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Veterinary Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>VSU-08-090.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article - Research</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Fibrinogen, horse, Inflammation, Serum amyloid A, Surgical trauma, White blood cell count</td>
</tr>
</tbody>
</table>
ABSTRACT

Objective – To evaluate the postoperative inflammatory response to elective surgery of varying intensity.

Study Design – Prospective longitudinal study.

Animals – Horses referred to two hospitals for arthroscopic removal of a unilateral osteochondrotic lesion in the tibiotalar joint (small surgical trauma, n = 11), correction of recurrent laryngeal neuropathy by a combination of laryngoplasty and ventriculectomy (intermediate surgical trauma, n = 10) or removal of an ovarian tumour by laparotomy (large surgical trauma, n = 5).

Methods – White blood cell counts and concentrations of serum amyloid A (SAA), fibrinogen and iron were assessed in blood samples obtained before and 1, 2, 3, 5, 7, 9, and 11 days after surgery. Horses underwent thorough clinical examination every day throughout the study period. Differences in levels of the inflammatory markers between the three surgical groups were analysed using repeated measurements ANOVA.

Results – Postoperative concentrations of SAA and fibrinogen were significantly higher in horses that underwent laparotomy and ovariectomy than in horses undergoing the combined laryngoplasty and ventriculectomy procedure or arthroscopy. Iron concentrations decreased to lower levels after intermediate and large surgical trauma than after small surgical trauma. White blood cell count did not differ between the three groups.

Conclusions – Levels of SAA, fibrinogen and iron reflected the intensity of the surgical trauma, whereas the white blood cell count did not.

Clinical Relevance – Postoperative measurements of SAA, fibrinogen and iron may in the future be used for comparing surgical trauma inflicted by new surgical techniques to that of already established techniques. Moreover, knowledge of the normal postoperative acute phase response is
essential, if acute phase reactants are to be used for monitoring occurrence of postoperative infections.
INTRODUCTION

When attempting to optimise surgical procedures, it is necessary to identify markers useful for evaluating the intensity of the trauma caused by the procedure. The surgical stress response is a set of well-described hormonal, metabolic, and inflammatory reactions occurring after surgery, which are directed at allowing the body to adapt to the trauma and injured tissues to heal. The intensity of the surgical trauma has been defined as 'the extent to which the factors that disrupt homeostasis is present'.

If uncontrolled, hypermetabolism and catabolism occurring postoperatively may lead to erosion of body mass and physiological reserve, which subsequently results in prolonged convalescence. Therefore, attempts to minimise surgical trauma through improvement of surgical techniques have been and are currently going on in human as well as veterinary medicine. Minimally invasive techniques such as endoscopic surgery have been shown to cause less postoperative inflammation compared to the corresponding open surgical techniques. This has been suggested to be the underlying pathophysiological reason for the observed improvement of short-term outcome measures such as length of hospital stay, pain, fatigue, and postoperative morbidity with minimally invasive surgery, even though some controversy exists (reviewed by Kehlet).

Several biomarkers have been proposed to reflect intensity of tissue injury in humans and animals. Interest has focused around endocrine metabolic factors and inflammatory mediators. While most research has indicated little or no difference in levels of endocrine metabolic factors released in response to open and to minimally invasive surgery (reviewed by Kehlet), several studies have shown that levels of inflammatory markers released in response to surgery do reflect the magnitude of the surgical trauma imposed on the patient. Concentrations of the cytokine interleukin-6 are higher in the postoperative period following conventional open surgery than after the corresponding
minimally invasive procedure. Interleukin-6 thus seems to be a sensitive indicator of the intensity of surgical trauma. Interleukin-6 is the main inducer of hepatic synthesis of acute phase proteins, which are proteins that are produced by hepatocytes and released to the blood stream in response to all inflammatory stimuli – including surgery – that cause tissue injury. Studies in humans have shown that levels of acute phase proteins are correlated to the intensity of the surgical trauma and the resulting cytokine response. 

In horses, several acute phase proteins exists. Among these, fibrinogen is the most well-known, and it has been used for decades to diagnose the presence of inflammatory conditions in horses and for monitoring changes in disease activity. Within recent years, interest has focused on serum amyloid A (SAA). Serum amyloid A is a major acute phase protein in horses, as concentrations increase quickly and with large amplitude in response to inflammation. These characteristics has been suggested to make SAA particularly well suited for diagnostic and monitoring purposes in horses.

Iron is a so-called negative acute phase reactant in horses and other species. During inflammation iron is sequestered in mononuclear phagocytes with the purpose of limiting iron available for microbial growth. In humans, serum concentrations of iron seem to reflect the intensity of the surgical trauma, and a recent paper by Borges et al. showed that iron was a very sensitive indicator of systemic inflammation.

The aim of the present study was to evaluate the inflammatory response to surgery in horses by measuring white blood cell counts (WBC) and concentrations of fibrinogen, SAA and iron before and after elective surgery of varying intensity and to determine whether the four inflammatory markers reflected intensity of the surgical trauma.
MATERIALS AND METHODS

Horses and Samples

The study was carried out as a prospective, longitudinal study. It included 26 horses, which underwent elective surgery at one of two equine hospitals (Hospital 1 and 2). Surgeries were either arthroscopic removal of a unilateral osteochondrotic lesion of the cranial intermedial ridge of the distal tibia (group 1, n = 11), correction of recurrent laryngeal neuropathy by a combination of laryngoplasty and ventriculectomy (group 2, n = 10) or removal of an ovarian tumour by laparotomy (group 3, n = 5 [midline incision n = 1, flank incision n = 4]) (Table 1). Groups were selected with the purpose of having three distinct levels of tissue injury represented and in order to include only surgeries performed under total anaesthesia. Arthroscopy was hypothesised to cause the smallest surgical trauma and laparotomy and ovariectomy the largest, with laryngoplasty and ventriculectomy causing a surgical trauma intermediate between the two other procedures. We opted for elective surgeries with fairly standardised surgical procedures and perioperative care in order to make groups as comparable as possible. Moreover, emergency surgeries with possible preoperative alterations in inflammatory, endocrine and metabolic parameters were not included, as we hypothesised that such preoperative changes might influence the postoperative inflammatory responses.

Horses underwent a standardised clinical examination before surgery and every day until the end of the study on day 11 postoperatively. Several clinical parameters were recorded: general appearance, appetite, rectal temperature, pulse, respiratory frequency, faeces, local signs of inflammation (heat, swelling, pain, and discharge) in the surgical wounds, as well as occurrence of disease unrelated to the surgical procedure. Further follow-up examinations were not performed.
Horses were only included in the study if they were judged healthy by a preoperative clinical examination and if their haematological and blood biochemical indices were within reference range prior to surgery. Exclusion criteria were development of infection in operation wounds (as determined by excessive swelling, pain and purulent exudate) or development of disease postoperatively. The study initially included 27 horses, but one horse from group 2 had to be excluded due to a slight fever and excessive swelling of and discharge from the surgical wound on day 1-3 after surgery. This horse is not included in table 1 and was also omitted from statistical calculations and results.

Blood samples were obtained by venipuncture of the jugular vein before (day 0) and on day 1, 2, 3, 5, 7, 9, and 11 after surgery. Blood was collected in tubes (Becton Dickinson Vacutainer Systems Europe, Meylan, France) containing sodium-EDTA for determination of WBC, tubes containing sodium-citrate for determination of plasma fibrinogen, and tubes with no additive for preparation of serum samples for analysis of SAA and iron concentrations. Citrated plasma was prepared by centrifugation at 2500 g for 15 minutes. Serum was prepared by letting blood samples coagulate for approximately 6 hours before centrifugation at 2500 g for 15 minutes. Analyses were carried out immediately (WBC) or serum/plasma was stored at -20 °C until analysis (SAA, fibrinogen, and iron).

Surgical Procedures and Perioperative Care

Four skilled equine surgeons (who had all worked exclusively with equine surgical diseases for 10 to 30 years) performed the surgeries. Surgical procedures were standardised between the two hospitals and carried out according to the procedures described by McIlwraith et al\textsuperscript{6} (arthroscopy), Fjeldborg\textsuperscript{12} (laryngoplasty and ventriculectomy), or Emberton\textsuperscript{16} (laparotomy and ovariectomy). All
surgeries were carried out in general anaesthesia. All horses in group 1 had one small osteochondrotic fragment removed from the intermedial ridge of tibia, and based on arthroscopic findings all horses were classified as having little or no synovitis.

All horses received tetanus prophylaxis and perioperative antibiotics at the operating surgeon’s discretion. All horses received 1.1 mg/kg flunixin meglumin b.i.d. for 3 days (Finadyne, Schering-Plough Animal Health, Farum, Denmark). Horses were box rested throughout the study period.

Laboratory Analyses

White blood cell counts were obtained using an automatic cell counter (ADVIA 120, Bayer A/S, Lyngby, Denmark). Serum SAA concentrations were determined by a previously described immunoturbidometric method (LZ test SAA, EIKEN Chemical Co., Tokyo, Japan). Fibrinogen concentrations were determined by the Clauss method in an automated coagulometric analyser (ACL 9000, Instrumentation Laboratory, Barcelona, Spain). Serum iron concentrations were determined by colorimetric spectrophotometry (ADVIA 1650, Bayer A/S, Lyngby, Denmark).

Statistical Analyses

The response parameters (WBC, SAA, fibrinogen, and iron) were analysed using repeated measures analysis of variance (ANOVA). Response parameters were transformed when necessary to meet model requirements (SAA transformed log(Y+0.1), fibrinogen Y**0.25, WBC log(y) and iron sqrt(Y)). The explanatory variables day relative to surgery, group, interaction between day and group, hospital, age and the baseline (day 0) value of the response parameters were included in the analysis.
model. Baseline (day 0) was included as a (continuous) covariate in the model to account for the
known variations in levels of (some of) the inflammatory markers in healthy horses. Akaike's information criteria were applied to choose the best fitting variance-covariance structure.

Models were checked by inspection of residual plots. Non-significant effects in the systematic part of the models were successively removed until only significant terms remained. Differences in least squares means estimates identified through the repeated measures ANOVA were used to determine significant differences between the three groups and the sampling days (and the interaction between group and day), where these showed overall significant differences. The data shown in Figure 1 is based on raw data averages and standard deviations, while comparisons within and between groups are based on the final statistical model. P < 0.05 was considered significant.
RESULTS

Horses remained clinically healthy throughout the study period. There were small spikes in rectal temperatures in the three groups on day 1 or 2, but with the exception of group 3, in which the average rectal temperature was 38.2 ºC on day 1, rectal temperatures, pulse and respiratory frequencies stayed within normal limits in the three groups for the entire duration of the study. One horse in group 3 developed mild swelling (day 1-11) and pain (day 1-2) of the laparotomy wound, but these signs of inflammation were not accompanied by fever or discharge, and the horse remained bright and alert during the study period.

The repeated measures ANOVA showed that levels of all four inflammatory markers changed in response to the surgical procedure (Table 2). White blood cell counts were significantly higher than preoperative baseline levels on day 1 (P < 0.01) and significantly below preoperative baseline levels on day 5 (P < 0.01). After day 5, WBC returned to baseline levels (Figure 1). Concentrations of SAA increased significantly from preoperative baseline levels in all three groups on day 1 and returned to baseline within 5 (group 1), 7 (group 2) and 11 (group 3) days after surgery (Figure 1). Fibrinogen concentrations increased significantly from preoperative baseline levels on day 1 (P < 0.01) in all three groups, and concentrations remained significantly elevated for the duration of the study (Figure 1). Serum iron concentrations decreased significantly (P < 0.0001) on day 1 in all three groups, but already on day 2 levels had returned to baseline concentrations (Figure 1). Levels of all four inflammatory markers were similar in the groups prior to surgery (P > 0.05).

Postoperative levels of WBC and fibrinogen, but not SAA and iron, depended significantly on the horses' baseline (day 0) levels of the markers (Table 2). Concentrations of SAA, fibrinogen and iron depended on the intensity of the surgery, whereas WBC did not differ between groups (Table 2). Concentrations of SAA were significantly higher in
group 2 than in group 1 on day 1-5 (P < 0.001) and tended to be higher in group 3 than in group 1 on the same days (P = 0.06). Likewise, concentrations of SAA tended to be higher in group 3 than in group 2 (P = 0.06) (Figure 1). Fibrinogen concentrations were higher in group 2 (P = 0.01) and 3 (P = 0.02) than in group 1, but did not differ significantly between group 2 and 3 (P = 0.36). Iron concentrations were lower in group 2 (P = 0.02) and 3 (P = 0.01) than in group 1, but did not differ significantly between group 2 and 3 (P = 0.81) (Figure 1).

Levels of the inflammatory markers did not differ between hospitals. Serum iron concentrations were affected by age, with concentrations increasing with increasing age. White blood cell counts and SAA and fibrinogen concentrations were not influenced by age of the horses (Table 2).
DISCUSSION

The results of the present study showed that all three surgical procedures caused significant inflammatory responses with alterations in inflammatory markers lasting for one or more days postoperatively. Moreover, the study showed that concentrations of the acute phase reactants SAA, fibrinogen and iron reflected intensity of the surgical trauma, while WBC did not.

As expected, patterns of responses differed between the four markers, with concentrations of SAA changing faster than concentrations of fibrinogen in response to the surgical trauma.9 Iron was a negative acute phase reactant, levels of which decreased quickly in response to the surgically-induced inflammation. This is consistent with findings in previous studies on localised inflammation in horses induced experimentally or surgically, as well as in hospitalised horses suffering from systemic inflammatory disease. Knowledge about the normal postoperative acute phase response may enable use of acute phase proteins and other markers of inflammation for detection of postoperative complications. Under normal circumstances the postoperative acute phase protein response has a rise-and-fall pattern in horses. Results from the present and previous studies have shown that SAA concentrations peak 2-3 days after surgery and return to preoperative levels within 7-10 days after surgery, whereas fibrinogen concentrations peak on day 3-6 after surgery and stay elevated for more than 11-15 days postoperatively. Deviations from this pattern may indicate that postoperative complications such as infection have occurred. A study by Jacobsen et al. showed that in horses developing clinical signs of excessive inflammation or infection after castration, SAA concentrations were persistently high and iron concentrations persistently low, whereas in horses with uncomplicated postoperative recovery, levels of these two inflammatory markers returned to preoperative levels within 8 days. Rectal temperatures, WBC and fibrinogen concentrations on the other hand were not useful for monitoring postoperative recovery.
Similarly, several studies in humans have suggested that higher or persistently elevated levels of cytokines, acute phase proteins, and other biomarkers of inflammation indicate presence of postoperative complications and may be used to predict short- and long-term outcome in surgical patients. Haematological and blood biochemical parameters may thus be used to support the clinical assessment of recovery and serve as an aid in early detection and institution of adequate therapy of postoperative complications.

Postoperative levels of WBC and fibrinogen depended significantly on levels determined preoperatively, while levels of SAA and iron did not. Horses with preoperative WBC and/or fibrinogen concentrations in the upper end of the reference interval thus had higher levels of these inflammatory markers postoperatively than horses with preoperative levels in the lower end of the reference interval. The SAA and iron responses did not depend on baseline levels in the horses, and a measured concentration of these two markers are thus easily interpreted, as their levels in a particular horse under healthy conditions do not affect their levels in the same horse during inflammatory disease.

The correlation between intensity of surgical trauma and postoperative inflammatory response has not previously been investigated in horses, but several studies in humans have found significant effects of magnitude of surgical trauma on serum levels of C-reactive protein, a major acute phase protein in humans, or on levels of interleukin-6, the main inducer of hepatic acute phase protein synthesis. In accordance with the findings of the present study, several previous studies in humans and horses have also shown that serum iron concentrations reflect magnitude of the underlying inflammatory response. When the surgical stress response can be graded, it becomes possible also to modulate it, for example through rational evaluation and improvement of surgical techniques; this may in turn improve postoperative outcome measures such as weight loss, fatigue, fever, time to wound healing and other physical variables with direct link to the acute phase
response. The clinical relevance of a reduced postoperative inflammatory response is currently not clear, but minimising surgical stress is generally perceived as beneficial. Surgery causes immunosuppression, and a reduction in the surgery-induced acute phase response has been suggested to play a role in improved recovery seen after minimally invasive procedures compared to the corresponding open procedures.1

White blood cell counts did not reflect intensity of the surgical trauma. While one study in dogs detected higher WBC after major surgery than after minor,2 human studies have consistently shown that white blood cell counts increase to similar levels after open and minimally invasive surgery.1,2,5 WBC thus seems to be a poor marker of the intensity of surgical trauma.

The inflammatory parameters measured in the present study may have been influenced by several factors other than the intensity of the surgical trauma. First of all, the age of the horses could possibly influence the inflammatory response. The results of the present study showed that WBC, SAA and fibrinogen levels were unaffected by age. In contrast, iron levels depended on the age of the horse, with iron concentrations increasing with increasing age. Iron stores and status is usually considered to increase with advancing age, but a study in clinically healthy horses (age 3 to 21 years) did not detect any age or gender difference in serum iron concentration. Concentrations did, however, differ between breeds, and diurnal variation was also demonstrated.12 Foals also display large changes in iron indices within their first months of life,14 and levels should thus be interpreted with some caution in this group. Several factors can thus influence iron concentrations in healthy individuals, but such concentrations differences between groups of healthy horses are probably of minor importance relative to the large concentration changes induced by inflammation.13,15

Secondly, anaesthetics and other drugs administered pre- or postoperatively might have influenced the inflammatory response. However, levels of the inflammatory parameters did not differ between hospitals, which suggests that the anaesthesia and treatment protocols of the two hospitals did not...
influence the postoperative inflammatory responses significantly. Two small studies (each with only two horses) have previously shown that general anesthesia alone has no effect on SAA and fibrinogen levels, and an equally small study in humans also indicated that duration of anesthesia did not influence postoperative interleukin-6 levels. In humans undergoing elective inguinal hernia repair, postoperative acute phase protein levels did not depend on whether the procedure was performed under regional or general anesthesia, and anesthesia thus seems to influence postoperative inflammation to a very limited degree or not at all.

Thirdly, the skill of the four involved surgeons could influence the results, as a more experienced surgeon could be hypothesised to cause smaller trauma during surgery resulting in smaller postoperative inflammation. It was not possible to control for the effect of surgeon in the statistical analysis, but based on the fact that all surgeons had more than 10 years of experience from specialised equine surgical practise, they were all considered very skilled.

Fourthly, nature of the surgical procedure might influence the postoperative inflammatory response. Not only the length of the surgical incision and the degree of tissue disruption, but also the type(s) of tissue involved in the procedure may play a role in determining the magnitude of postoperative inflammation. Abdominal and – to an even greater extent – thoracic surgery seems to cause release of higher levels of interleukin-6 and acute phase proteins than musculoskeletal surgery. Peritoneal cells produce several types of cytokines including interleukin-6, and laparotomy has been shown to cause large increases in cytokine concentrations in peritoneal fluid, thus possibly affecting magnitude of the postoperative acute phase response.

In addition to the above mentioned factors, the results of the present study may also have been affected by the modest sample size. Twenty-six horses were included, and this number may have been insufficient for detection of small differences between groups. Small studies carry a risk of making type II errors, i.e., accepting a null hypothesis of no difference, where a difference is truly
present in the study population. Type II error might for example explain why the observed
difference in SAA levels between group 2 and 3 did not quite reach statistically significant levels.

We and others have previously suggested that SAA may be particularly suited for real-time
monitoring of disease activity in horses.\textsuperscript{10,24,26,37} The results of the present study corroborate these
suggestions by showing 1) that SAA concentrations reflected the magnitude of underlying tissue
injury, 2) that postinjury levels did not depend on preinjury levels, 3) that amplitude of the response
was very large, which facilitates differentiation between healthy and diseased (average
postoperative peak concentration of SAA was 444 times higher than preoperative levels, as
compared to 1.6 times higher and 2.5 times lower for fibrinogen and iron, respectively), and 4) that
levels changed in close parallel to injury and recovery (as opposed to fibrinogen levels, which
increased more slowly and took more than 11 days to start returning to preinjury levels).

In conclusion, the acute phase reactants SAA, fibrinogen and iron reflected intensity of the surgical
trauma, and these three inflammatory markers may thus be used for evaluating differences in
trauma caused by different surgical techniques. Moreover, the results of the present study
corroborate previous studies, which have suggested that SAA is a particularly useful marker of
inflammation in horses.
ACKNOWLEDGEMENTS

REDACTED
REFERENCES


Figure 1. Total white blood cell count (A) and concentrations of serum amyloid A (B), fibrinogen (C), and iron (D) before and after surgery in horses that underwent surgery of varying intensity (arthroscopy [group 1] = ■, laryngoplasty and ventriculectomy [group 2] = ●, laparotomy and ovariectomy [group 3] = ▼). Data shown are average ± standard deviations. The day 0 sample was obtained before surgery, and blood samples were collected every or every other day postoperatively for 11 days.
Table 1. Horses included in the study

<table>
<thead>
<tr>
<th>Surgical procedure</th>
<th>No. of horses</th>
<th>Age (years)</th>
<th>Gender (n)</th>
<th>Race (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthroscopy (group 1)</td>
<td>9</td>
<td>4 (1-8)</td>
<td>Gelding (6)</td>
<td>Danish warmblood (6), standardbred trotter (3), stallion (1), oldenburg (1), mixed breed (1), unknown (1)</td>
</tr>
<tr>
<td>Laryngoplasty and ventriculectomy (group 2)</td>
<td>3</td>
<td>8.5 (3-15)</td>
<td>Gelding (7)</td>
<td>Danish warmblood (7), trakehner (1), Russian trotter</td>
</tr>
<tr>
<td>Laparotomy and ovariectomy  (group 3)</td>
<td>4</td>
<td>9 (5-16)</td>
<td>Mare (5)</td>
<td>Danish warmblood (1), standartbred trotter (1), connemara (1), knapstrup (1), unknown (1)</td>
</tr>
</tbody>
</table>
Table 2. Significance of effects of the explanatory variables day, group(small, intermediate or large surgical trauma), interaction between day and group, preoperative baseline levels of the response parameters, age and hospital on the white blood cell, serum amyloid A, fibrinogen and iron responses to surgery

<table>
<thead>
<tr>
<th>Response parameter</th>
<th>Day</th>
<th>Group</th>
<th>Day*group</th>
<th>Baseline (day 0)</th>
<th>Age</th>
<th>Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cell count</td>
<td>P = 0.0001</td>
<td>NS</td>
<td>NS</td>
<td>P &lt; 0.01</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Serum amyloid A</td>
<td>P &lt; 0.0001</td>
<td>P &lt; 0.01</td>
<td>P = 0.01</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>P &lt; 0.0001</td>
<td>P = 0.02</td>
<td>NS</td>
<td>P &lt; 0.0001</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Iron</td>
<td>P &lt; 0.0001</td>
<td>P = 0.02</td>
<td>NS</td>
<td>NS</td>
<td>P = 0.03</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS = non-significant
Figure 1. Total white blood cell count (A) and concentrations of serum amyloid A (B), fibrinogen (C), and iron (D) before and after surgery in horses that underwent surgery of varying intensity (arthroscopy [group 1] = ■, laryngoplasty and ventriculectomy [group 2] = ●, laparotomy and ovariectomy [group 3] = ▲). Data shown are average ± standard deviations. The day 0 sample was obtained before surgery, and blood samples were collected every other day postoperatively for 11 days.

179x138mm (600 x 600 DPI)
RESPONSE TO REVIEWERS’ COMMENTS

Thank you very much for the very thorough review of our work and for relevant comments. Below is the response to each reviewer’s comments and to those of the associate editor. Changes in the manuscript have been highlighted by using the track changes mode in Word.

Reviewer #1

Line 12: Changed as requested.
Line 63: Reference made to the Borges paper as suggested, as this paper is indeed highly relevant.
Lines 93-94: Changed as suggested.
Line 96: Assumption correct. Information added as requested.
Line 118: Changed as suggested.
Line 186: Description strengthened as suggested.
Line 245: Statistical analyses were repeated in order to include age as an explanatory variable. The results show that by including age of the horses in the statistical analysis, the difference in iron levels is not related to the hospital (and subsequent sample handling), but rather to the age of the horses (many of the younger horses underwent surgery at hospital 1). The manuscript has been amended accordingly.
Discrepancies: The optimal situation would of course have been if more horses could have been included (particularly in the group with the largest surgical trauma) and if surgeries could have been randomised among hospitals and surgeons. However, randomisation was not possible in the current study, as it included client-owned (as opposed to experimental) horses. We thus had to go ahead and include all horses that fulfilled all our criteria (owned by clients willing to let their horse enter the study, completely healthy by clinical and haematological and blood-biochemical examination, possible to handle with multiple blood samplings etc) presented to the hospitals during the study period. Moreover, we had to accept that the horse owners wanted surgeries performed at a particular hospital.
Collecting clinical data over time presents a herculean task, and the number of horses included in the current study represents the absolute maximum manageable during the study period. Similar studies carried out in humans also have modest number of individuals (n ranges between 16 and 61, most have 30-40 individuals divided into 2-5 groups).
We feel that the low number of horses and the unequal distribution among hospitals did not significantly influence our results. The main problem in studies with low numbers of study subjects is the increased risk of making a type II error (i.e. accepting a null hypothesis of no difference, where a difference is truly present in the study population), but in this study we indeed demonstrated group differences despite the low number of horses, and it thus seems that the number of horses included was sufficient for our purpose.
However, we agree that some findings could be related to the low number of horses in the study (an example: that the group-differences in preoperative WBC were not statistically significant). An explanation of the possible problems arising from the low number of horses has been added to the Discussion.

Reviewer #2
1. We agree that the data might also suit a table format. However, as exact values are not of importance for the conclusions (as we are not attempting to define cut off values etc.), we feel that a visual presentation gives the quickest overview of the data. We will leave the decision on presentation to the editorial office. We are of course willing to change figure 1 into a table, if the editors prefer that solution.

Reviewer #3

1. Sample size: see comment re. sample size stated above (under response to Reviewer #1). As stated above, we have added a paragraph about type II error to let the reader know how the small sample size may have affected results. Surgeons’ ability: unfortunately formalised surgical training (such as ECVS diplomacy) has only recently become available in our country. Therefore, no formal degrees can be added to explain the surgeons’ ability. However, we have added descriptive information, which hopefully suffices (Materials and Methods, under Surgical Procedures and Perioperative care). Moreover, a paragraph discussing the possible influence of the surgeon has been added to the Discussion. Surgical procedures: see comment below.

2. It is correct that levels of inflammatory markers do not differ between group 2 and 3 on a 95% level. SAA levels did, however, differ on a 94 % level, as explained in the Results section. This might very well be a result of the low number of horses in group 3, as explained in the added paragraph about the possible type II error (see comments for Reviewer #1).

3. All horses were very similar in that they had a small osteochondrotic fragment located on the intermediary ridge of tibia. All had no or negligible synovitis. This information has been added to the manuscript.

4. We agree that inclusion of both midline and flank approaches are not optimal. However, due to the low number of cases in this group, we felt that it was necessary to include all cases presented to us during the study period. Our hypothesis was that length of incision (which was much larger than in the other two groups) and possible incision in peritoneum was more significant than site. In fact, the one horse operated through a midline approach had intermediate SAA peak value, while one horse receiving flank incision had a very low peak value, two had intermediate, and one a high SAA peak value. According to our inclusion criteria none of the horses developed postoperative complications or morbidity.

5. None of the horses showed systemic signs of inflammation prior to surgery (based on clinical examination, haematology and blood biochemistry). This information has been added to the manuscript. We agree that local inflammation could have been present, but to a very low degree judged to be negligible for the purpose of the study (as it could not be detected by a thorough clinical examination or by blood analyses). As described above, only very mild synovitis was found in the arthroscopy group.

6. It is not clear to us, whether the reviewer means better choices for performing the three types of surgeries chosen for this project or whether other surgical groups might have been better. With regard to the first point, the chosen procedures reflect those used in the hospitals at the time of data collection (or rather: the procedures, which the two hospitals could agree upon using for the project). With regard to the second point, the choice of groups was based on several factors: they had to be elective conditions with negligible or no preoperative inflammation, they had to be sufficiently frequent to allow data collection to occur within a reasonable time frame, they had to be surgeries that both the involved
hospitals and all involved surgeons were familiar with and felt comfortable performing etc. We were also keen to find procedures mimicking those used in previous human studies, as we knew in advance that it would be difficult to find equine studies on the subject (for comparing our results with those of previous studies). Such choices may of course always be discussed, and when performing clinical studies on client-owned horses, the optimal solution may not always be practically feasible.

7. We are encouraged by the fact that this reviewer supports our interest in increased use of acute phase proteins in the equine clinic. And we certainly agree with the shortcomings of the study design pointed out by the reviewer. The reasons underlying the choices re. study design, the relatively low sample size, and surgical procedures are explained above. It is unfortunately not possible for us to re-do the entire study with different surgical procedures (if we understand the reviewer correctly, this is what is suggested, as a changes in surgical procedures would not allow us to use the current data at all, as groups would not be comparable). We feel that this is the first study of its kind in the equine scientific literature, and one of its merits is therefore novelty. Further studies expanding and substantiating the findings of the present study are therefore highly relevant. We would be happy to add more detailed descriptions of what we did and explain why we did it, if the reviewer will let us know specifically where such information is needed.

Reviewer #4

Paragraph 1: We agree with the reviewer that a small study such as ours may have problems with low statistical power. As described under our comments for Reviewer #1, we have added a paragraph to the Discussion about the risk of making type II errors in order to make this absolutely clear to the readers. Group-specific information about the response parameters are depicted in figure 1.

Paragraph 2: As stated in the manuscript, baseline (day 0) values of the markers were included as explanatory variable in the statistical analysis because levels of WBC and fibrinogen (and to some extent iron) varies markedly between healthy horses (which is why these parameters have wide reference intervals). For example, one horse may a normal fibrinogen level of 2.2 g/L (and concentrations will remain at this level as long as the horse is healthy), and after an inflammatory stimulus this concentration increases to 5 g/L. Another horse may have 4 g/L fibrinogen in plasma as long as it remains healthy, and concentrations may increase to 8, 9 or even higher levels, when the horse receives the same inflammatory stimulus as the first horse. Knowing this, we felt that preoperative values of the markers were important determinates of the levels the markers would reach postoperatively (see the inserted figure, which shows levels of WBC and SAA before and after an inflammatory stimulus in two horses). And indeed, the statistical analysis showed that postoperative
levels of WBC and fibrinogen depended on the levels measured prior to surgery. However, the reviewer's comments shows that this is inadequately explained in the manuscript, and additional descriptions and interpretations of this finding have therefore been added in Materials and Methods (under Statistical Analyses) and in the Discussion. The title of table 2 has also been changed in an attempt to make it absolutely clear that day 0 levels are used as explanatory variables.

Baseline comparisons between groups have been performed (there were no preoperative differences in levels of the four markers between the three groups); this information has been added to the Results.

Paragraph 3: As explained above, day 0 was included only as an explanatory variable.

Paragraph 4: We agree completely with the reviewer, please see our response to reviewer #3.

Paragraph 5: We have now included age in the statistical analyses, which have changed the results (please see our response to Reviewer #1). The entire manuscript has been changed accordingly.

Paragraph 6: This information has been added to Materials and Methods (under Statistical Analyses).

Paragraph 7: Number of significant digits reduced in Results section and Table 2

Paragraph 8: Taking preoperative blood samples for several days is not possible in a study using client-owned horses. We asked the horse owners to let the horses stay at the hospitals for the entire duration of the postoperative study period (11 days), and this is much longer than normal. So adding extra days would not have been possible at all (at least the number of horse owners agreeing to let their horse participate would drop dramatically). Moreover, it is well known that there are some day-to-day changes (and for some parameters even diurnal changes) in levels of haematological and clinical-chemical parameters. However, these changes are very small as compared to those induced by inflammation (for an example see the inserted figure showing day-to-day changes in SAA, fibrinogen, WBC and iron in 15 horses), so for this study these day-to-day changes are of little importance.

**Associate editor**

We appreciate the positive comments and generally agree with editor re. the shortcomings of the study. However, we feel that we have adequately answered the reviewers' comments and amended the manuscript accordingly.
1. The small number of horses: this is mainly a problem when statistical significance is not demonstrated, as this could be a result of a type II error. An explanation of this has been added to the manuscript in order to let the readers decide for themselves.

2. Adding additional explanatory factors to the statistical model: we have included age in the analyses and revised the manuscript according to the new results. As Reviewer #4 describes, it is not possible to also include surgeon as an explanatory factor. As the participating surgeons had several years of experience, we feel that they are comparable. Other factors that might have contributed also could be breed, gender, season, previous disease history etc. We included the explanatory factors that were judged to be most relevant, and added age, as this was pointed out as a possibly relevant factor by the reviewers. Testing all the possibly relevant factors (and all their possible interaction terms) would require statistical power that cannot be obtained in clinical studies (as the required n would go up very drastically).

3. Statistical methods: we have made an attempt at explaining the statistical methods in more detail and present the results as clear as possible in the revised manuscript. The possible difference in baseline levels has been tested (there is no difference among groups). We are not sure what the editor means by the significant P-values in Table 2, but these show the postoperative response alone, as day 0 levels of the markers were included as an explanatory (and not an outcome) variable. This seems to have been inadequately described, as Reviewer #4 also remarks on this point, and we have thus tried to explain it better.

Summary: we feel that the study was adequately designed to test our hypothesis. We used a longitudinal study design and applied the appropriate statistics. With regard to the suboptimal execution, we are not sure exactly which parts of the study that the editor feels were poorly executed. We will of course be happy to answer any specific comments. As described in our answer to Reviewer #1 the data collection for this study was lengthy and cumbersome. We have no further funding for the study, and we are therefore not in a position to add more horses (in addition, this would be a problem, as surgical techniques have changed in the meantime, which would make horses within groups less comparable). As also described above, the number of cases match similar studies performed in humans.