



## Evaluation of Cosmetic Results of Surgical Wound Closure in Dogs

**Rachel Wendy Williams** DVM<sup>1\*</sup>

**Wanda Gordon-Evans** DVM, PhD, DACVS, DACVSMR<sup>2</sup>

**Jeffery Biskup** DVM, DACVS<sup>3</sup>

**Dominique Griffon** DMV, MS, PhD, DACVS, DECVS<sup>1</sup>

<sup>1</sup> Western University of Health Sciences College of Veterinary Medicine, East 2nd Street Pomona CA, CA 91766, United States

<sup>2</sup> University of Minnesota College of Veterinary Medicine, 1365 Gortner Ave, St Paul, MN 55108, USA

<sup>3</sup> University of Tennessee College of Veterinary Medicine, 2407 River Dr, Knoxville, TN 37996, USA

\* Corresponding Author ([rchlwilliams9@gmail.com](mailto:rchlwilliams9@gmail.com))

---

ISSN: 2396-9776

Published: 15 Feb 2018

in: Vol 3, Issue 1

DOI: <http://dx.doi.org/10.18849/ve.v3i1.123>

Reviewed by: Kristina Kiefer (DVM, PhD) and Victoria Arbona (DVM, MS)

---



## ABSTRACT

**Objective:** To evaluate the correlation between wound cosmesis and pet owner satisfaction, to determine the agreement among vet surgeons, and pet owners evaluating a surgical wound with a visual assessment score (VAS), and to determine the agreement between the VAS, a semi-quantitative score of wound inflammation, and wound width.

**Background:** Perception of post-surgical cosmesis by human patients has been found to influence their satisfaction and perceived adequacy of vet surgeons. Due to the trend of owners anthropomorphising pets, this logic can be extended to veterinary patients. Also, there is a lack of consistent, reliable methods to evaluate cosmesis have been developed, creating the need for a scoring system that is accurate and reproducible.

**Evidentiary value:** This was a prospective cohort study with one hundred and seven patients. This study may not change day-to-day practices, but it will bring to light for practitioners the discordance between pet owners and vet surgeons concerning attractiveness of an incision as well as overall satisfaction with a procedure.

**Methods:** Photographs of surgical wounds were taken immediately, 2 weeks, and 8 weeks after surgery in dogs. Owners were asked to evaluate satisfaction with the procedure and attractiveness of the incision using the VAS. Photos were evaluated for cosmetic outcome by pet owner and vet surgeon evaluators with different scoring systems. The reliability of the scoring systems was evaluated using intraclass correlations and kappa statistics as appropriate.

**Results:** Owners' evaluation of cosmetic outcome correlated with their overall satisfaction. ( $r^2=0.37$ , where  $r^2$  is the coefficient of determination, and can be used to determine what proportion of variance in one variable is predictable from the other variable) There was generally poor reliability of the subjective scoring between both vet surgeons and pet owners, and the less subjective scoring systems.

**Conclusion:** The cosmetic outcome as judged by the owner accounted for 37% of the variability in satisfaction, that is to say that 37% of the total variation in satisfaction could be explained by the cosmetic outcome as evaluated by the owner. This suggests that wound cosmesis significantly influences clients' satisfaction. However, the other methods of rating cosmesis evaluated in this study were unreliable.

**Application:** The results of this study are relevant to all veterinarians in clinical practice, from general practice to tertiary referral centers. This study should be kept in mind when speaking with owners to establish pre-surgical expectations as well as mediating conflict due to a difference in opinion with regards to a surgical outcome.

## INTRODUCTION

In human surgery, patient's perception of outcome is affected by the cosmetic appearance of the scar. There can be marked psychosocial changes related to the appearance of the surgical site over multiple surgical disciplines (Duncan et al 2006; Exner et al 2012; Flanagan 2009; Kim et al 2015). In addition, it has been shown that cosmetic outcomes can affect patient-perceived adequacy of their surgeon in human medicine (Boulding et al 2011).

Although it is unlikely that animals are psychologically affected by appearance as humans are, owners may anthropomorphise especially as pets become increasingly treated like family members and thus leaving

owners less satisfied with less cosmetic incisions (Boni 2008). Owners may also use the incisional appearance to judge the quality of the surgery or the ability of the surgeon. Cosmesis of the incision is an obvious outcome that owners are able to easily see and understand at home, as compared to other factors such as hemostasis and the ability to acquire margins such as in the case of a mass removal. Although this is logical, it has not been evaluated scientifically in veterinary medicine.

A variety of validated scoring systems have been developed for use in humans with different types of surgery, ranging from deformity correction to cancer reconstruction. These scoring systems range from a simple, self-reported visual analog score to a complicated psychological screening test evaluating the effect of cosmetic results on quality of life (Duncan et al 2006; Holland et al 2001; Kim et al 2015). Evaluators in these studies included pet owners (Duncan et al 2006); vet surgeons (Holland et al 2001); and psychologists (Kim et al 2015). In spite of the multitude of scoring systems designed for multiple types of cosmetic outcomes, the authors were unable to find scoring systems that could be extrapolated to veterinary surgery. No scoring system for surgical wounds has been validated in animals. In recent publications dealing with cosmesis, scoring systems were either absent, or were substituted by parameters like inflammation (Gallegos 2007; Maninchedda 2015). Additionally, no evidence of validity or reproducibility was disclosed in these reports (Etter et al 2013; Peeters 2011; Sylvestre 2002). In this study, the authors chose to use the visual analog score that had been previously described (Duncan et al 2006) due to its demonstrated validity, consistency, and reliability.

The objectives of this study were to evaluate the correlation between wound cosmesis and owner's satisfaction with the surgery, to determine the inter-rater agreement, among vet surgeons, and pet owners evaluating a surgical wound with a visual assessment score (VAS), and to determine the agreement between the VAS, a composite score of wound inflammation and wound width.

We hypothesised that cosmesis would influence owner satisfaction and that the agreement in wound assessment would be fair between owners and vet surgeons, and good between vet surgeons. We hypothesised that the VAS would correlate with the composite scores. Finally, we hypothesised that the scar width and the less subjective scores would have the highest inter-rater reliability and correlation with owner opinion.

## METHODS & MATERIALS

### Animals

Dogs were recruited from Wisconsin Veterinary Referral Center or Banfield Pet Hospital of Oviedo with owners' consent. Animals were included if they underwent soft tissue or orthopaedic surgery between May 2014-July 2014. Surgery was performed by vet surgeons as well as general practitioners. Fractious patients and those subjected to radiation therapy were excluded. Signalment, location, and type of surgery were recorded. Owners were asked at 2 and/or 8 weeks to evaluate the incision and satisfaction with the surgery. Photographs were taken post-operatively, at 2 weeks, and at 8 weeks if possible for evaluation by the pet owners and vet surgeons.

### VAS Score

Owners used a visual analog score to rate the appearance of the surgical wound. A 10cm line was marked where the level of cosmetic outcome fell. The 0cm point was defined as the most unacceptable cosmetic outcome, whereas the 10cm point represented the cosmetically pleasing outcome. A similar line and instructions were provided to rate owner' satisfaction with the surgery.

### **Digital image analyses**

Digital photographs were taken immediately, 2 weeks and 8 weeks after surgery. The photographs were taken with a camera placed perpendicular to the incision. A ruler was placed on the skin adjacent to, but not interfering with the incision. The photos were assigned a numerical identifier and provided to masked evaluators. Each image was provided to each evaluator in a random order, without any identification or indication of timing relative to surgery.

### **Composite score**

Three of the pet owners were recruited, 3 vet surgeons (JB,WGE, DG) independently evaluated all photographs using the VAS used by pet owners. Additionally, the vet surgeons evaluated each image with a numerical rating scale (Sylvestre et al 2002) to assess 4 aspects of the wounds: swelling, erythema, discharge, and dehiscence. These factors were adopted from a publication comparing different suture patterns to close celiotomies after canine ovariohysterectomies (Sylvestre et al 2002). These 4 scores were added together and the mean was found to calculate a composite score for each wound, at each time point, and for each surgeon.

### **Quantitative image analysis:**

After completion of the subjective and semi-quantitative scorings, the same 3 board certified vet surgeons (JB, WGE, DG) evaluated each image with image processing software (Image J, NIH, USA). Each image was calibrated with the ruler placed next to the incision. The width of abnormal tissue surrounding the beginning, the middle, and the end of the wound was measured (Table 1) and the 3 values were averaged.

### **Data analysis**

Correlation of owner cosmetic and satisfaction VAS scores were performed using a regression analysis. The inter-rater reliability of the pet owner and vet surgeon VAS scores, composite score, and average width of incisions were evaluated using intraclass correlations (ICC). Each individual category within the composite score was evaluated with a kappa statistic. Significance was set at  $p < 0.05$ . If the inter-rater reliability was significant, a regression was used to determine the correlation with owner opinion and satisfaction.

ICCs were considered poor for values less than 0.40, fair for values between 0.40 and 0.59, good for values between 0.60 and 0.74, and excellent for values between 0.75 and 1 (Hallgren 2012). Kappa values 0-0.20 indicated slight agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 substantial agreement, and 0.81-1 almost perfect agreement (Landis 1977).

**Table 1:** Mean ± Standard Deviations of the VAS, Composite Score, and Wound Width

		Mean score from pet owners (n=3)	Standard deviation of pet owners (n=3)	Mean score from vet surgeons (n=3)	Standard deviation of vet surgeons(n=3)
VAS	PO (N=85)	81.08	20.03	79.14	11.6
	Week 2 (N=41)	91.09	11.32	83.24	13.65
	Week 8 (N=14)	95.82	7.65	94.47	8.6
Composite Scores	PO (N=90)	N/A	N/A	1.09	0.49
	Week 2 (N=41)	N/A	N/A	0.64	0.62
	Week 8 (N=14)	N/A	N/A	0.48	0.32
Scar Width	PO (N=90)	N/A	N/A	5.01	3.53
	Week 2 (N=42)	N/A	N/A	4.25	4.07
	Week 8 (N=14)	N/A	N/A	1.6	3.28

## RESULTS

One hundred and seven patients were included in this study. The average age of patients included in this study was 5.88 years ranging from 0.4 to 14.7 years. Of those 119 patients, 64 of them underwent soft tissue surgery while the remaining 55 were orthopaedic procedures. Of the orthopaedic procedures, 26 of them were tibial plateau leveling osteotomies (TPLO) which accounted for 24% of all procedures.

Owner satisfaction and owner cosmetic score was correlated significantly, with an  $r^2$ -value of 37% ( $p < 0.0001$ ). In general, the majority of the incisions were rated above the 5cm mark with only a few raters assigning VAS scores below 5cm. At the first recheck, 56 of the reported VAS scores by owners were rated above the 5cm mark, while only two were scored below that, indicating that most owners evaluated the incision as having an appearance closer to excellent. This could demonstrate a limitation in the study by which there was not a wide enough spread with regard to the appearance of incisions. The intraclass correlation of the pet owners and vet surgeon groups, composite score, and average scar width are presented in Table 2. Overall, only the 2 week composite score and the average scar width at 8 weeks were moderately reliable. When evaluating the individual components of the composite score (Table 3), reliability was generally poor. The kappa statistic could not be calculated at 8 weeks because of low numbers of cases evaluated.

**Table 2:** Intraclass Coefficients of VAS, Composite Score, and Average Width to Evaluate Inter-rater Reliability

	pet owners (N=3)	Vet surgeons (N=3)		
	VAS	VAS	Composite Score	Wound Width (mm)
Post-operative (N=85)	0.03	0.41	0.04	0.41
2 Weeks (N=41)	0	0.5	0.74	0.58
8 Weeks (N=14)	0.28	0.41	0.25	0.73

**Table 3:** Kappa Statistic for the Components of the Composite Score to Evaluate Inter-rater Reliability between Vet surgeons. This statistic is used to demonstrate the reliability of non-parametric data and compares only 2 observers at a time. Therefore, 1x2 signifies observer 1 compared to observer 2 and 2x3 signifies observer 2 compared to observer 3 etc. If  $p < 0.05$ , no significant finding (NSF) was listed.

	0 weeks	2 weeks
Swelling	1x2: NSF	1x2: 0.28
	1x3: NSF	1x3: 0.21
	2x3: 0.20	2x3: NSF
Erythema	1x2: NSF	1x2: 0.34
	1x3: NSF	1x3: 0.31
	2x3: 0.24	2x3: 0.53
Dehiscence	1x2: NSF	1x2: 0.53
	1x3: NSF	1x3: 0.42
	2x3: NSF	2x3: 0.47
Discharge	1x2: NSF	1x2: NSF
	1x3: 0.50	1x3: NSF
	2x3: NSF	2x3: NSF

## DISCUSSION

Owners' opinion of cosmesis accounted for 37% of the variation in satisfaction. Although this is not a high correlation, it is a clinically significant portion of overall satisfaction. Over 1/3 of the owners' satisfaction was due to the cosmetic appearance of the surgery site. Other factors affecting satisfaction, such as resolution of disease, interactions with staff, or prior expectations for outcome were not evaluated in this study. However, some types of surgery may be more or less dependent on cosmesis in terms of owners' satisfaction. For example, almost half of the procedures were orthopaedic and within that group TPLOs were overrepresented. It is possible that factors such as return to function, or lameness could have influenced the satisfaction

response provided by owners. Additionally TPLO incisions may be more noticeable on a day-to-day basis than an abdominal incision. Alternatively, owner satisfaction may have been higher in patients that underwent oncologic surgery. In such cases, the owner may give higher priority to the success of disease than appearance of the wound, than owners of pets undergoing elective procedures. For example, cases where large malignancies were removed may be found very satisfactory in spite of unsightly scars because the disease has been addressed successfully. Our inability to assess these variables could be considered limitations of this study, and further investigation is warranted to determine that relationship.

The reliability of the VAS, composite score, and scar width measurements was generally disappointing. In general, the intraclass coefficient is an indicator of inter-rater agreement with 1 indicating perfect agreement. The highest intraclass coefficient (ICC) in this study was equal to 0.74 and characterised the composite scores assigned 2 weeks after surgery. The higher level of agreement likely reflects the fairly uniform levels of inflammation between wounds at this stage. The relative uniformity facilitates the evaluation compared to that at the time of surgery. In general, most of the characteristic signs of inflammation take at least a day to appear (Flanagan 2009). The subsequent decrease in ICC at 8 weeks may have been influenced by the decrease in number of follow-ups available at this stage. Many owners did not follow up at referring veterinary hospitals or were not medically required to return after suture removal. The second highest agreement between observers consisted of the scar width measured at 8 weeks, with an ICC equal to 0.73. This level of agreement is not surprising, since incisions should be more uniform by this time. In addition, subjective factors, such as swelling and inflammation likely affected early postoperative measurements but would have likely been less of a factor by 8 weeks, as they should have resolved by that point in time. Moving forward, this would indicate that use of the composite score would facilitate evaluation by multiple veterinarians, such as in the situation where a patient is seen by a specialist but follow-ups are conducted by the primary veterinarian, as compared to using the VAS, or something similar, that has not demonstrated reproducibility. This data also suggests a use for the composite score in further studies evaluating wound healing.

The poor agreement amongst evaluators using the VAS may reflect interindividual variations in assessing different aspects of the incision. A small misalignment between skin edges may be perceived acceptable by a surgeon, but may be considered as a poor outcome by another. Additionally, the majority of the incisions in our study were cosmetically acceptable, and tended to score favorably via VAS. The agreement between observers would have likely been improved if our study included more cases with poor cosmesis. Wounds would have been better distributed along the entire spectrum of possible cosmetic results. The poor agreement between pet owners and vet surgeons when using the VAS emphasises the importance of direct follow-up, as opposed to follow-up via phone.

The kappa statistic reflects the difference between how much agreement is observed and how much should be present based on random chance. Based on this test, we found some agreement amongst the observers for the components of the composite score. When looking at the traditional kappa statistic, 0 indicates no agreement above what would be expected by random chance, while 0.01-0.20 is slight chance, 0.21-0.40 is fair, 0.41-0.60 is moderate, 0.61-0.8 is substantial, and 0.81-1.00 is almost perfect agreement. The kappa statistics calculated here ranged from fair to moderate with no scores reported above 0.53. The best scores were reported for erythema and dehiscence. Unlike the ICC which took all three evaluators into consideration, the kappa statistic only evaluates 2 evaluators at a time. Although some of the evaluators were in high agreement, this finding was not consistent among all of them. Interobserver agreement was best when evaluating dehiscence, but the level of agreement between all evaluators remained fair to moderate regardless.

The ideal scoring system for cosmesis would be objective, accurate, and highly reliable. In general, evidence is stronger when objective methods are used because of the reduction in variation and bias introduced by subjectivity. For example, gait analysis is an accepted objective proxy for joint pain and is considered the gold standard for lameness evaluation whereas subjective lameness scoring is less reliable (Donnell 2015 Waxman et al 2008). Developing objective measures to characterise inherently subjective traits is difficult. Cosmesis is much like pain in that it is subjective by definition, and therefore hard to reliably quantify. In a study comparing the subjective assessment of colic, there was a high degree of interobserver variance, leading to the conclusion that it was inaccurate (Keppie et al 2008). It is not surprising that a subjective method of

determining cosmetic outcomes of surgery, as used in this study, would perform similarly to other subjective assessments such as those mentioned earlier.

There are multiple ways to improve the reliability of scoring systems devised in this study. Image capture of the incisions may be improved with standardised lighting and focus. Digital imagery and analysis also improves rapidly and it is possible that 3-dimensional analysis would improve results. Swelling and scars that are raised were difficult to discern from flat scars in this study. Training can also improve agreement amongst evaluators. However, cosmesis is an inherently subjective evaluation that is influenced by multiple factors and scoring systems will likely always need refinement and improvement.

Limitations that were noted in this study were the lack of spread when it came to the reported scores by owners. As mentioned earlier in this paper, the majority of the owner-reported scores of the appearance of their incisions were above the 5 cm mark. A wider spread in the data could have strengthened the significance of the findings addressed here. Another limitation that was encountered was the lack of follow up at 8 weeks, which likely affected the ICC for the composite scores at that time frame. Given the severe reduction in cases at the 8 week time frame, this would have allowed for outliers to shift calculations far more than if there had been a greater number or wider variety in cases. The final limitation that was noted in this study was the use of 2-dimensional images for this study. Images were analyzed from a uniplanar approach, which may not have been sufficient at the 8 week time from when scar tissue may protrude from the surface of the patient and be unable to be visualized using the methods detailed here. Future studies may consider the use of photographs from multiple angles or a standardized video platform.

## CONFLICT OF INTEREST

**Acknowledgments:** The authors thank Banfield Pet Hospital of Oviedo for their technical assistance.

**Funding:** Supported in part by the Steve Henriksen Extramural Grant.

**Competing interests:** No competing interests.

## REFERENCES

1. Boni, S. (2008). Anthropomorphism: how it affects the human-canine bond. *Journal of Applied Companion Animal Behavior*, 2(1), 16-21.
2. Boulding, W. Glickman S. Manary M. et al. (2011). Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. *American Journal of Managed Care*, 17(1), 41-48.
3. Donnell, J. Frisbie, D. King, M. et al. (2015). Comparison of subjective lameness evaluation, force platforms and an inertial-sensor system to identify mild lameness in an equine osteoarthritis model *The Veterinary Journal*, 206(2), 136-142. <http://dx.doi.org/10.1016/j.tvjl.2015.08.004>
4. Duncan, J. Bond, J. Mason, T. et al. (2006). Visual analogue scale scoring and ranking: a suitable and sensitive method for assessing scar quality? *Plastic and Reconstructive Surgery*, 118(4), 909-918. <http://dx.doi.org/10.1097/01.prs.0000232378.88776.b0>
5. Etter, S. Ragetly, G. Bennett, R. et al. (2013). Effect of using triclosan-impregnated suture for incisional closure on surgical site infection and inflammation following tibial plateau leveling osteotomy in dogs. *Journal of the American Veterinary Medical Association*, 242(3), 355-358. <http://dx.doi.org/10.2460/javma.242.3.355>
6. Exner, R. Krois, W. Mittlböck, M. et al. (2012). Objectively measured breast symmetry has no influence on quality of life in breast cancer patients. *European Journal of Surgical Oncology*, 38(2), 130-136. <http://dx.doi.org/10.1016/j.ejso.2011.10.012>
7. Flanagan M. (2009). The physiology of wound healing. *Journal of Wound Care*, 9(6), 299-300. <http://dx.doi.org/10.12968/jowc.2000.9.6.25994>



8. Gallegos, J. Schmiedt, C. McNulty J. (2007). Cosmetic rostral nasal reconstruction after nasal planum and premaxilla resection: technique and results in two dogs. *Veterinary Surgery*, 36(7), 669-674. <http://dx.doi.org/10.1111/j.1532-950X.2007.00319.x>
9. Hallgren, K. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Quantitative Methods for Psychology*, 8(1), 23-34. <http://dx.doi.org/10.20982/tqmp.08.1.p023>
10. Holland, J. Smith, G. Ross, F. et al. (2001). HOSE: an objective scoring system for evaluating the results of hypospadias surgery. *BJU International*, 88(3), 255-258. <http://dx.doi.org/10.1046/j.1464-410x.2001.02280.x>
11. Keppie, N. Rosenstein, D. Holcombe, S. et al. (2008). Objective radiographic assessment of abdominal sand accumulation in horses. *Veterinary Radiology and Ultrasound*, 49(2), 122-128. <http://dx.doi.org/10.1111/j.1740-8261.2008.00337.x>
12. Kim, M. Kim, T. Moon, H. et al. (2015). Effect of cosmetic outcome on quality of life after breast cancer surgery. *European Journal of Surgical Oncology*, 41(3), 426-432. <http://dx.doi.org/10.1016/j.ejso.2014.12.002>
13. Landis, J. Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <http://dx.doi.org/10.2307/2529310>
14. Maninchedda, U. Nottrott, K. Gangl, M. (2015). Enucleation in equids with and without the use of an orbital suture meshwork implant: a retrospective study of 37 cases. *Veterinary Record*, 177(14), 367. <http://dx.doi.org/10.1136/vr.103305>
15. Peeters, M. Kirpensteijn, J. (2011). Comparison of surgical variables and short-term postoperative complications in healthy dogs undergoing ovariohysterectomy or ovariectomy. *Journal of the American Veterinary Medical Association*, 238(2), 189-194. <http://dx.doi.org/10.2460/javma.238.2.189>
16. Sylvestre, A. Wilson, J. Hare, J. (2002). A comparison of 2 different suture patterns for skin closure of canine ovariohysterectomy. *Canadian Veterinary Journal*, 43(9), 699-702.
17. Waxman, A. Robinson, D. Evans, R. et al. (2008). Relationship between objective and subjective assessment of limb function in normal dogs with an experimentally induced lameness. *Veterinary Surgery*, 37(3), 241-246. <http://dx.doi.org/10.1111/j.1532-950X.2008.00372.x>



---

### Intellectual Property Rights

Authors of Articles submitted to RCVS Knowledge for publication will retain copyright in their work, and will be required to grant to RCVS Knowledge a non-exclusive license of the rights of copyright in the materials including but not limited to the right to publish, re-publish, transmit, sell, distribute and otherwise use the materials in all languages and all media throughout the world, and to license or permit others to do so.

---

Veterinary Evidence and EBVM Network are RCVS Knowledge initiatives. For more information please contact us at [editor@veterinaryevidence.org](mailto:editor@veterinaryevidence.org).

RCVS Knowledge is the independent charity associated with the Royal College of Veterinary Surgeons (RCVS). Our ambition is to become a global intermediary for evidence based veterinary knowledge by providing access to information that is of immediate value to practicing veterinary professionals and directly contributes to evidence based clinical decision-making.

**[www.veterinaryevidence.org](http://www.veterinaryevidence.org)**

RCVS Knowledge is a registered Charity No. 230886.  
Registered as a Company limited by guarantee in England and Wales No. 598443.

Registered Office:  
Belgravia House  
62-64 Horseferry Road  
London SW1P 2AF



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

---