



Published in final edited form as:

*J Invest Surg.* 2017 April ; 30(2): 88–94. doi:10.1080/08941939.2016.1229367.

## A Novel Method for Evaluating Postoperative Adhesions in Rats

Geoffrey M. Bove<sup>1</sup>, Susan L. Chapelle<sup>1,2</sup>, Eleanor Boyle<sup>3,4</sup>, David J. Mokler<sup>1</sup>, and Jan Hartvigsen<sup>3,5</sup>

<sup>1</sup>Department of Biomedical Sciences, University of New England College of Osteopathic Medicine, Biddeford, ME, USA

<sup>2</sup>Squamish Integrated Health, Squamish, BC, Canada

<sup>3</sup>Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

<sup>4</sup>University of Toronto, Dalla Lana School of Public Health, Toronto, ON, Canada

<sup>5</sup>Nordic Institute of Chiropractic and Clinical Biomechanics, University of Southern Denmark, Odense M, Denmark

### Abstract

**Purpose/Aim:** Postoperative adhesions remain an undesirable and commonly symptomatic side effect of abdominopelvic surgeries. Animal models of postoperative adhesions typically yield heterogeneous adhesions throughout the abdominal cavity and are not easily quantified. Here we present a novel method of postoperative adhesion assessment and report its reliability and measurement error.

**Materials and Methods:** A model of cecal abrasion with partial sidewall attachment was performed on female rats. After 1, 2, 4, or 7 days of recovery, the rats were euthanized and their abdominopelvic cavities were systematically evaluated for postoperative adhesions. The necropsy was recorded through the surgical microscope. Four raters were trained to use a ballot to capture key factors of the adhesions as they viewed the recordings. Their ratings were compared for measurement error and reliability (using Bland-Altman plots and intraclass correlation coefficients, respectively) and for the ability to discriminate differences in experimental groups. A subset of the data was analyzed to determine practical utility.

**Results:** The rating system was shown to have low measurement error and high inter-rater reliability for all parameters measured. Applied practically, the system was able to discriminate groups in a manner that was expected.

---

Address correspondence to Geoffrey M. Bove, DC, PhD, Department of Biomedical Sciences, University of New England College of Osteopathic Medicine, 11 Hills Beach Rd., Stella Maris 102, Biddeford ME 04005, USA. gbove@une.edu.

#### AUTHOR CONTRIBUTIONS

GMB and SLC designed the rating methods, EB, JH, and GMB designed and performed the data analysis, and all authors contributed to the writing of the manuscript.

*Declaration of interest.* The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

**Conclusions:** We have developed and validated a rating system for postoperative adhesions and shown that it can detect group differences. This method can be used to quantify postoperative adhesions in rodent models.

### Keywords

postoperative adhesions; reliability; gastrointestinal surgery; rat; bowel obstruction; infertility

---

## INTRODUCTION

Intraperitoneal postoperative adhesions (PAs) form following the majority of abdominal and pelvic surgeries and are often symptomatic [1]. A vast number of substances have been tested in humans and animal models as preventatives [2, 3], but no treatment has been found to consistently result in prevention. Therefore, research using animal models continues with the goal of solving this prevalent side effect of surgery and could benefit from a more standardized method of quantifying adhesions throughout the abdominopelvic cavity.

There are numerous methods used to induce PAs in animals [4–6]. All involve some type of trauma to the internal organs, the parietal peritoneum, or both, and may include leaving a source of persistent inflammation in the gut, such as a site of ischemia, chromic gut suture, or even gallstones [7].

Quantification of PAs in animal models has proven to be difficult for a number of reasons. Methods such as the cecal sidewall model [8] and the formation of ischemic buttons [9] allow quantification of the number and extent of the adhesions to these sites. However, these quantification methods do not characterize PAs in other locations in the gut, for which there remains no “gold standard.” The most common methods are subjective and have not been well characterized. During necropsy, one or more observers, preferably blinded to the treatment group, rate the adhesions for various parameters including identification of the structures that are adhered, and the extent, tenacity, and vascularity of each adhesion. Scales typically consider no adhesions as 0, and the most severe-looking adhesions as 4 or 5 [10–15]. More recent articles seeking to rate adhesions cite these initial methods, and apply them in whole or with slight modifications. Parameters are often added together for the final score, such as “extent” and “tenacity,” although they are not necessarily independent. Only a few reports used more than one observer [7, 16–18] or used consensus scoring [19]. Only one study reported the assessment of examiner agreement [7]. In sum, the results from different studies cannot be readily compared to each other using current methods.

The aim of this article is to describe a novel method of grading PAs within the entire abdominal cavity in rats and report the intraexaminer reliability and measurement error of the grading scale.

## MATERIALS AND METHODS

All studies were performed using methods approved by University of New England’s Institutional Animal Care and Use Committee and in compliance with the Guide for the Care and Use of Laboratory Animals (National Research Council, USA). Data are reported

from a total of 72 female Sprague Dawley rats (Charles River, USA) weighing 200–225 g. These rats were being used for a different experiment studying the effect of an intervention on PAs, which will be reported elsewhere.

We developed a variation of previous surgical methods to suit our studies of PAs, which will be described in detail elsewhere. Following anesthesia with isoflurane (1.75%–2.25% in pure oxygen) and appropriate surgical preparation, the abdominal wall was incised just right of the linea alba for 2.5 cm. The cecum and 5–6 cm of small intestine were exteriorized. The cecum was abraded over its entire surface using a circular brush [19], causing visceral peritoneal disruption and petechial hemorrhaging. The intestines were placed back in the abdominal cavity, and the left abdominal wall was retracted over a custom brace. A 2 cm1 cm rectangle of the peritoneum was outlined with × a scalpel, and the parietal peritoneum with transversalis fascia and part of the transversus abdominis muscle were undermined with small scissors, avoiding the epigastric vessels, and removed. The cecum was retrieved and sutured above and below the abdominal wall lesion using two 4–0 nylon sutures, along the midline of the lesion (5 mm lateral to the linea alba). The abdominal wall was closed in layers using interrupted 4–0 nylon sutures. Rats were given one dose of buprenorphine (0.05 mg/kg) and placed separately in clean cages to recover.

After 1, 2, 4, or 7 days, rats were given an overdose of isoflurane and exsanguinated by opening the thorax and lesioning the heart. This prevented bleeding during the necropsy, which we previously noted interfered with the grading. The abdominal skin was reflected distally. Using a surgical microscope at 3.75×, the abdominal wall was opened starting at the distal xiphoid, followed by a wide oval incision of the abdominal wall, which isolated but did not disturb the cecal–abdominal adhesion.

To record the data for rating, a digital camera (Nikon DS-Fi1, USA) attached to Nikon Elements software was started. The surgeon systematically examined the abdomen, identifying and then disrupting all adhesions. The necropsy proceeded by examining adhesions involving the greater omentum, then the left and right adnexal fat pads, and then any other adhesions between the cecum and other organs, or of the cecum to itself. The cecum–abdominal wall adhesion was not disrupted during this process but was removed in entirety after the video was finished (and not included in this report). Videos ranged from 2–3 min long. A total of 72 videos were reviewed and scored by the observers. The videos were of necropsies from 28 rats used as controls during a time course of PAs, 30 from rats used for interventions, and 14 “decoys” that were not used in the training and were not included in the analysis. Only results related to the characterization of the scoring system will be presented; results for the intervention experiment will be reported elsewhere. Videos were copied and renamed using a random number generator. They were then placed in ascending number, thus randomly mixing the experimental groups.

A ballot was created (Figure 1) to record data for each rater. The ballot had provisions for recording every type of adhesion that is seen following this and other similar surgical methods. For our purposes, the cecum to abdominal wall adhesion was being quantified using other methods, but this could be added to this ballot. We had the raters record the presence of each type of adhesion, and the severity and extent of each adhesion (1 or 2). We

asked the raters to record their impression of inflammation, and their impression of the overall adhesion severity on a scale of 0–10. The following parameters were analyzed: overall severity, overall impression of inflammation, number of adhesions, number of adhesions not involving the fatty structures, summed severity of adhesions, and the summed extent of adhesions.

Four raters volunteered to rate the videos. Two of the raters had some experience grading similar videos and otherwise viewing the surgeries and results. The other two raters were graduate students with little previous experience with these methods. The surgeon who performed all the surgeries and performed all the necropsies (GMB) trained the raters as a group during a single-narrated session using a number of videos not included in the analyzed data. First, a normal abdomen dissection was presented to familiarize the raters with anatomy and the normal positioning and mobility of the abdominopelvic contents. This was followed by narrated presentations of rats with various severities of adhesions. All raters had a ballot in their hands, and references were made throughout to the ballot and how the particular feature of the necropsy would be rated. The training session lasted less than an hour, including discussion.

The raters were shown the group-blinded videos in numerical order (thus randomly by group), sized full-screen on a 22'' monitor. Each rater scored the videos independently. The ballots were collected, stapled together. An overall score was reached by consensus, and written down on the back of the ballot package, allowing for later comparison of the overall mean score with the overall consensus score. The entire process took 3–4 min per video.

Ratings for 58 videos were analyzed (72 videos minus the “decoy” videos). Data were tabulated in excel and analyzed using Stata (USA) and Prism (USA). We constructed Bland-Altman plots for each parameter to visualize the discrepancy between the raters and also to examine the performance of the scale over the entire range, i.e., investigating for bias/instability at the ends plus any systematic difference between the measurements [20]. The X-axis was defined as the mean of the four raters' scores for each rat. The Y-axis was defined as the mean of all of the pairwise differences between the two raters (six sets of pairwise differences) for each rat [21]. The standard deviation calculated for the limits of agreement was based on the pairwise differences. The limits of agreement were calculated using Bland and Altman's formula [20]. We calculated intraclass correlation coefficients (ICCs) using Model 3, Form 1, chosen because the same raters judged the videos of the same rats independently and these raters were the only ones of interest [22].

To demonstrate the utility of the measure we graphed and analyzed the mean and individual results for the 28 control videos. These only varied by postoperative survival days [1, 2, 4, 7] following cecal abrasion surgery ( $n = 7$  in each group). As healing in the rat gut primarily occurs during this epoch, we hypothesized that our method could differentiate some parameters between these groups. These data were graphed and analyzed using 1-way and 2-way ANOVAs as appropriate, with Fisher's LSD post hoc tests.

## RESULTS

The process of rating the videos was found to be efficient, and all videos were rated during a single session. The majority of the points in the Bland-Altman plots were close to the mean difference line with few points lying beyond the limits of agreement (Figures 2A–F), indicating low measurement error. The limits of agreement were narrow (judgment is based on the size of the scale), indicating a small amount of variation and high agreement. These data support that the raters were subjectively scoring the videos in the same fashion. The plots for the impression of inflammation severity (Figure 2B) showed an increasing trend as the mean score increased. These data support that the raters had acceptable but lower agreement judging inflammation when it was more severe.

We used the results from all raters and calculated ICCs as a measure of inter-rater reliability for the primary parameters (Table 1). The ICCs indicated very high inter-rater reliability for each parameter [23], ranging from 0.77 to 0.92. The impression of the inflammation severity, while having the lowest ICC, still showed practical utility in detecting group differences (see next paragraph). Coupled with the Bland-Altman plot results, these data support that one rater should be sufficient to rate videos from experiments involving quantification of PAs using this method (however, see next paragraph and Discussion).

The analysis of 28 control videos was clearly able to distinguish between groups (Figures 3A–F and Table 2). The mean overall severity scores were compared to the consensus overall severity scores using a 2-way ANOVA and were found to be essentially identical (Figure 3A). Both methods revealed the same results by day, showing significant differences between Day 1 and other days. These same results were found for each individual rater (Figure 3G). The metric “inflammation impression” had the lowest ICC, but the mean results demonstrated the decrease in inflammation that is expected following this surgery (Figure 3B), with all days being significantly different from each other. These results were mostly but not entirely reflected for each individual rater (Figure 3H), where statistical analyses showed some differences for individual raters compared to the combined ratings (see Figure 3 Legend). The total number of adhesions (Figure 3C) and the number of nonfatty adhesions (Figure 3D) were significantly different, with the results for individual raters being identical (Figures 3I and J). The mean severity and mean extent per adhesion were not significant overall (Figures 3E and F), but the 2-way ANOVAs showed some differences for individual raters (Figures 3K and L).

## DISCUSSION

Our modified method of evaluating PAs in rats proved to be easy to implement. The high inter-rater reliability and the low measurement error (Figure 2) indicated that the scales used are appropriate for raters with similar levels of training and experience.

When applied to a typical experiment, using 7 rats per group, the method revealed some basic but expected information on the biology of postoperative adhesions. For instance, normal healing dictates that inflammation would subside following surgery, and this can be appreciated in Figure 3B. The number of adhesions is also expected to decrease over time

due to the fibrinolytic system, which is most prominent 48–72 hours postoperatively. Figure 3C shows this reduction of the number of adhesions and Figures 3D and F support that this difference is due to decreased adhesions between parts of the gut, rather than between the fatty structures (omentum and adnexal fat pads). These results support the utility and sensitivity of these evaluation methods.

Changes in the extent (size) and severity (difficulty pulling apart, or “tenacity”) per adhesion were not apparent, but this may be because the rating choices were limited to 1 and 2. We purposefully did not combine these features because in our experience they are not independent. If an adhesion’s strength is consistent over its area or attachment, then the size will necessarily contribute to the parameter “tenacity.” Adding these together would serve to magnify a result inappropriately. We also did not include “vascularity” as a metric because our attempts to subjectively evaluate this parameter seemed inconclusive. This parameter would be measured more accurately using histological methods.

Because the measurement error was low and the inter-rater reliability was high for all parameters, and since the results for the overall severity scores were essentially identical between raters, we can conclude based on the statistical analysis that one trained rater can reliably rate PAs using this method. However, we also show that discrepancies exist for other parameters, such as the impression of inflammation severity. To buffer these discrepancies, it might be advised to follow the adage “two heads are better than one” by having two raters view the videos and reach consensus for each rated parameter, yielding one score per parameter and one ballot for each video. This would have the additional benefit of offering camaraderie while performing an admittedly tedious task, which should help the process proceed more efficiently.

## Acknowledgments

### FUNDING

Research reported in this publication was supported by the National Institute General Medical Sciences of the National Institutes of Health under Award Number R01GM108041 to GMB. The content does not necessarily represent the views of the National Institutes of Health.

## REFERENCES

- [1]. Ten Broek RP, Bakkum EA, Laarhoven CJ, et al. Epidemiology and prevention of postsurgical adhesions revisited. *Ann Surg.* 2016;263:12–19. [PubMed: 26135678]
- [2]. Ten Broek RP, Stommel MW, Strik C, et al. Benefits and harms of adhesion barriers for abdominal surgery: a systematic review and meta-analysis. *Lancet.* 2014;383:48–59. [PubMed: 24075279]
- [3]. Metwally M, Cheong Y, Li TC. A review of techniques for adhesion prevention after gynaecological surgery. *Curr Opin Obstet Gynecol.* 2008;20:345–352. [PubMed: 18660685]
- [4]. Ozel H, Avsar FM, Topaloglu S, et al. Induction and assessment methods used in experimental adhesion studies. *Wound Repair Regen.* 2005;13:358–364. [PubMed: 16008724]
- [5]. Whang SH, Astudillo JA, Sporn E, et al. In search of the best peritoneal adhesion model: comparison of different techniques in a rat model. *J Surg Res.* 2011;167:245–250. [PubMed: 20304431]
- [6]. Poehnert D, Abbas M, Kreipe HH, et al. High reproducibility of adhesion formation in rat with meso-stitch approximation of injured cecum and abdominal wall. *Int J Med Sci.* 2015;12:1–6. [PubMed: 25552912]

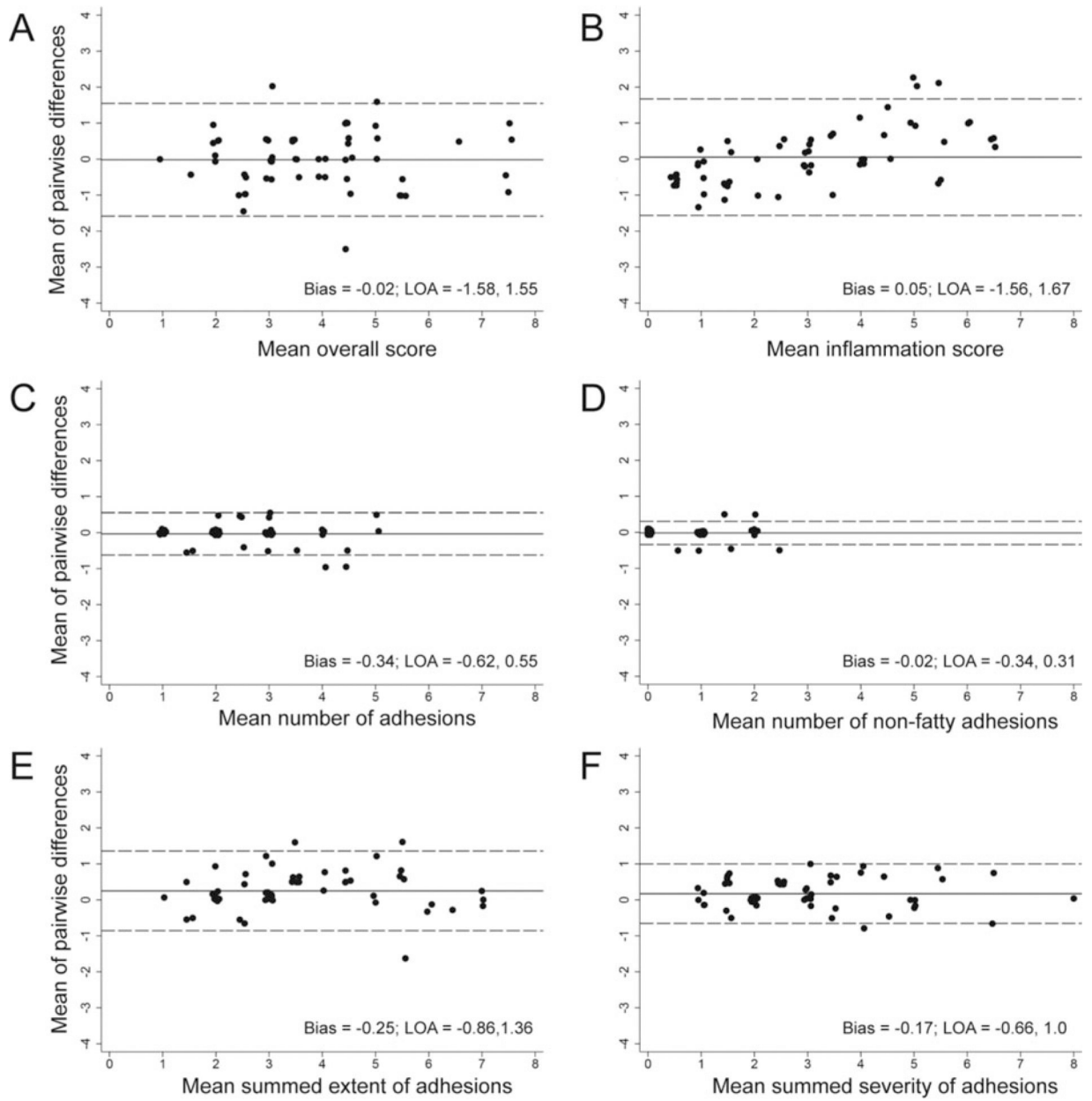
- [7]. Hansen KA, Lowman L, Fiedler EP, et al. Pelvic adhesion formation after intraperitoneal installation of gallstones in a rabbit model. *FertilSteril*. 1999;72:868–872.
- [8]. Cashman J, Burt HM, Springate C, et al. Camptothecin-loaded films for the prevention of postsurgical adhesions. *Inflamm Res*. 2004;53:355–362. [PubMed: 15316666]
- [9]. Reed KL, Fruin AB, Bishop-Bartolomei KK, et al. Neurokinin-1 receptor and substance P messenger RNA levels increase during intraabdominal adhesion formation. *J Surg Res*. 2002;108:165–172. [PubMed: 12443729]
- [10]. Cone DF. The effect of intestinal motility on the formation of adhesions. *Bull Johns Hopkins Hosp*. 1959;105:8–13. [PubMed: 13662717]
- [11]. Knightly JJ, Agostino D, Clifton EE. The effect of fibrinolysin and heparin on the formation of peritoneal adhesions. *Surgery*. 1962;52:250–258. [PubMed: 14457251]
- [12]. Mazuji MK, Fadhli HA. Peritoneal adhesions; prevention with povidone and dextran 75. *Archives of Surgery*. 1965;91:872–874. [PubMed: 5842944]
- [13]. Nair LS. Prolotherapy for tissue repair. *Transl Res*. 2011;158:129–131. [PubMed: 21867977]
- [14]. Diamond MP, Linsky CB, Cunningham T, et al. A model for sidewall adhesions in the rabbit: reduction by an absorbable barrier. *Microsurgery*. 1987;8:197–200. [PubMed: 3431416]
- [15]. Blauer KL, Collins RL. The effect of intraperitoneal progesterone on postoperative adhesion formation in rabbits. *Fertil Steril*. 1988;49:144–149. [PubMed: 3335262]
- [16]. Fielder EP, Guzick DS, Guido R, et al. Adhesion formation from release of dermoid contents in the peritoneal cavity and effect of copious lavage: a prospective, randomized, blinded, controlled study in a rabbit model. *Fertil Steril*. 1996;65:852–859. [PubMed: 8654650]
- [17]. Roman H, Canis M, Kamble M, et al. Efficacy of three adhesion-preventing agents in reducing severe peritoneal trauma induced by bipolar coagulation in a laparoscopic rat model. *FertilSteril*. 2005;83(Suppl 1):1113–1118.
- [18]. Kim TH, Park JS, An SS, et al. Inhibition of thrombin-activated fibrinolysis inhibitor decreases postoperative adhesion. *J Surg Res*. 2015;193:560–566. [PubMed: 25167782]
- [19]. Bove GM, Chapelle SL. Visceral mobilization can lyse and prevent peritoneal adhesions in a rat model. *J Bodywork Mov Ther*. 2012;16:76–82.
- [20]. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310. [PubMed: 2868172]
- [21]. Gwet KL, Kilem LI. In: *The Definitive Guide to Measuring the Extent of Agreement Among Raters*. 4th ed. Gaithersburg: Advanced Analytics, LLC; 2014.
- [22]. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428. [PubMed: 18839484]
- [23]. LeBreton JM, Senter JL. Answers to 20 questions about interrater reliability and interrater agreement. *Org Res Methods*. 2008;11:815–852.

**Adhesion Grading** Grader Initials 2 Video Number 1596 Decoded Number 118

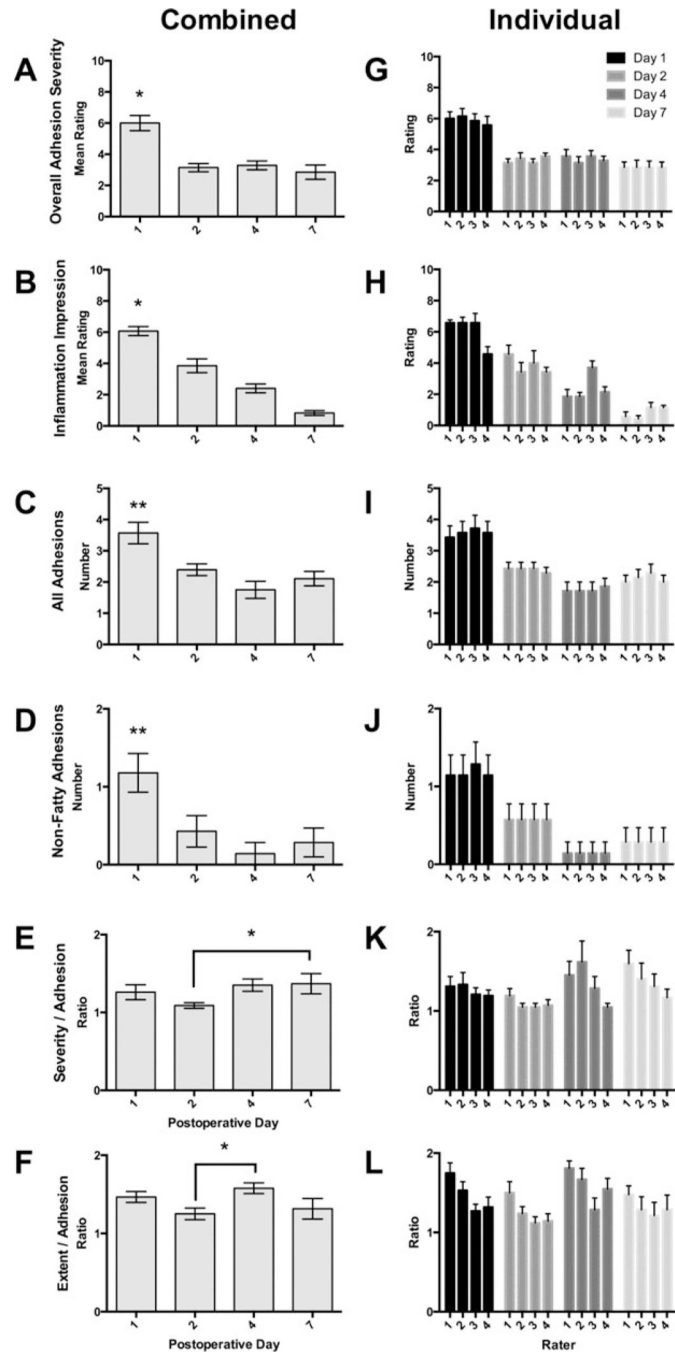
Anatomical Site of Adhesion	Adhesion Presence 0 = no 1 = yes	*If Yes Specify Below	
		Severity 1=Mild / flimsy 2=Severe / dense	Extent 1=localized / small 2=extensive / large
Omentum	1	1	1
Left Fatpad (top of screen)			
Right Fatpad (bottom of screen)			
Small Int. -- Cecum			
Small Int. -- Large Int.			
Large Int. -- Cecum	1	1	1
Other adhesion	1	1	1
How "inflamed" on a scale of 0-10?	4		
Total number of adhesions	3		
<b>Overall Impression on a Scale of 0-10</b>	<b>5</b>		

**FIGURE 1.**  
Rating ballot.





**FIGURE 2.** Bland-Altman plots showing levels of agreement for the four raters. For details see Results section.



**FIGURE 3.** Comparisons of means and individual ratings (+/- SEM) for control experiments by day (statistics in Table 2). (A) Comparison of overall mean ratings with consensus ratings showed no difference using 2-way ANOVA. The comparison by day showed significant differences between Day 1 and other three days, confirmed by post hoc tests (\*  $p < .001$ ). This same relationship was observed when individual ratings (G) were analyzed. (B) The rating of inflammation impression was statistically different for each group. All six post hoc comparisons were statistically different (\* =  $p < .005$ ). This same relationship was observed

when individual ratings (H) were analyzed. Post hoc tests confirmed all differences ( $p < .05$ ) with the exception of Rater 3 Days 2–4 and Rater 4 Days 1–2 and 4–7. (C) Total number of adhesions was different by day, the significance driven by Day 1 compared to the other groups ( $p < .005$ ). This same relationship was observed for all individual ratings, with all differences confirmed by post hoc tests ( $p < .005$ ). (D) Number of nonfatty adhesions was significantly different by day, the significance driven by Day 1 compared to the other groups ( $p < .005$ ). This same relationship was observed for all individual ratings, with all differences confirmed by post hoc tests ( $p < .05$ ). (E–F) Mean severity and extent per adhesion did not differ by day. (K) Individual severity per adhesion differed by day, but with only two differences with post hoc testing (Rater 1 Days 2–7, and Rater 2, Days 2–4). (L) Individual extent per adhesion differed by day, with three differences with post hoc testing (Rater 2 Days 2–4 and 4–7, and Rater 4 Days 2–4).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE 1**

Intraclass correlation coefficients (ICCs) and their 95% confidence intervals (CIs)

	<b>ICC</b>	<b>95% CI</b>	
Overall score	0.80	0.73	0.87
Inflammation score	0.77	0.68	0.85
Number of adhesions	0.91	0.87	0.94
Number of nonfatty adhesions	0.92	0.89	0.95
Severity of adhesions	0.86	0.79	0.90
Extent of adhesions	0.85	0.78	0.90

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE 2**

Statistical analysis of comparisons of Figure 3

<b>Comparison</b>	<b>F</b>	<b>p</b>
Mean vs. consensus severity ratings (A)	(1,12) = 0.02	= 0.88
Consensus severity ratings by day (A)	(3,24) = 14.3	< 0.0001
Mean severity ratings by day (A)	(3,24) = 16.8	< 0.0001
Individual ratings (G)	(3,96) = 48.8	< 0.0001
Inflammation (B)	(3,24) = 52.1	<.001
Individual ratings (H)	(3,96) = 105.4	< .0001
Total adhesions (C)	(3,24) = 8.9	<.001
Individual ratings (I)	(3,96) = 48.8	< .0001
Nonfatty adhesions (D)	(3,24) = 5.4	<.01
Individual ratings (J)	(3,96) = 20.1	< .0001
Mean severity (E)	(3,24) = 1.9	= 0.15
Individual ratings (K)	(3,96) = 3.52	= 0.018
Mean extent (F)	(3,24) = 2.7	= 0.07
Individual ratings (L)	(3,96) = 3.27	= 0.002

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript