

# Severity assessment using three common behavioral or locomotor tests after laparotomy in rats: a pilot study

Laboratory Animals  
0(0) 1–9  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0023677220911680  
journals.sagepub.com/home/lan



Leonie Zieglowski<sup>1</sup>, Anna Kümmecke<sup>1</sup>, Lisa Ernst<sup>1</sup>,  
Mareike Schulz<sup>1</sup>, Steven R Talbot<sup>2</sup> , Rupert Palme<sup>3</sup>,  
Michael Czaplak<sup>4</sup>  and René H Tolba<sup>1</sup> 

## Abstract

The aim of this pilot study was to evaluate whether behavioral or locomotor tests (Open Field (OF), rotarod (RR), and CatWalk (CW)) can help assess the severity of laparotomy in rats.

The new EU Directive (2010/63/EU) mandates severity assessment in experiments involving animals. However, validated and objective methods are needed to relate trial-specific monitoring results to the degree of distress caused to individual animals. Therefore, we focused on non-invasive or minimally invasive, simple, and convenient severity assessment methods in a surgical model.

To evaluate surgical severity in this model, we compared moving velocity among three commonly used behavioral test methods (OF, RR, and CW) after midline laparotomy within postoperative 7 days.

In this study, 30 adult male Wistar Han rats ( $n = 10$  per test) were trained in their assigned test method and subsequently subjected to surgery. Severity scoring was performed daily using a modified score sheet developed previously. In addition, blood and fecal samples were collected to analyze surgical and postoperative corticosterone metabolite levels. We found significant differences among the experimental groups in terms of the analyzed parameters. In this context, the OF test was found to be the most suitable method for severity assessment after laparotomy in rats.

## Keywords

Distress behavior, surgical techniques, severity assessment, rat, motor activity physiology

Date received: 7 August 2019; accepted: 13 February 2020

## Introduction

In 2010, the focus on animal welfare was renewed in Europe with the issuance of Directive 2010/63/EU.<sup>1</sup> Article 15 makes it mandatory to classify the extent of distress that an animal will suffer during an experiment as the degree of severity (DS).<sup>1–3</sup> DS is generally categorized as non-recovery, light, moderate, or severe. Severity rating is based on the assessment of the degree of pain, suffering, anxiety, or lasting harm experienced by the animal exposed to distress during an experiment. It is known that different forms of stress exist and are perceived differently by species.

The term “severity assessment” is still very new and is therefore often equated with pain assessment. Therefore, severity assessment in this study can be

<sup>1</sup>Institute for Laboratory Animal Science and Experimental Surgery, RWTH Aachen University, Faculty of Medicine, Germany

<sup>2</sup>Institute for Laboratory Animal Science, Hannover Medical School, Germany

<sup>3</sup>Department of Biomedical Sciences, University of Veterinary Medicine, Austria

<sup>4</sup>Department of Anesthesiology, University Hospital RWTH Aachen, Germany

### Corresponding author:

René H Tolba, Institute for Laboratory Animal Science and Experimental Surgery, RWTH Aachen University, Faculty of Medicine, Pauwelsstrasse 30, 52074 Aachen, Germany.  
Email: rtolba@ukaachen.de

described via the following six influencing factors: pain, suffering, anxiety, affective internal/emotional state, lasting harm, and distress (Figure 1). The aim of the study is to create a multimodal approach from the different methods and components of severity assessment in order to record, categorize, and evaluate severity entirely, already at the individual animal level. This pilot study identified and evaluated different approaches, such as behavioral or locomotor testing, in combination with biochemical analysis and clinical parameters. This enables an all-encompassing severity assessment down to the individual animal level.<sup>4</sup>

However, validated methods and objective measures are required to relate trial-specific monitoring results to the degree of distress caused to individual animals. Therefore, we focused on strategies for severity assessment that are non-invasive or minimally invasive as well as simple and convenient to apply.

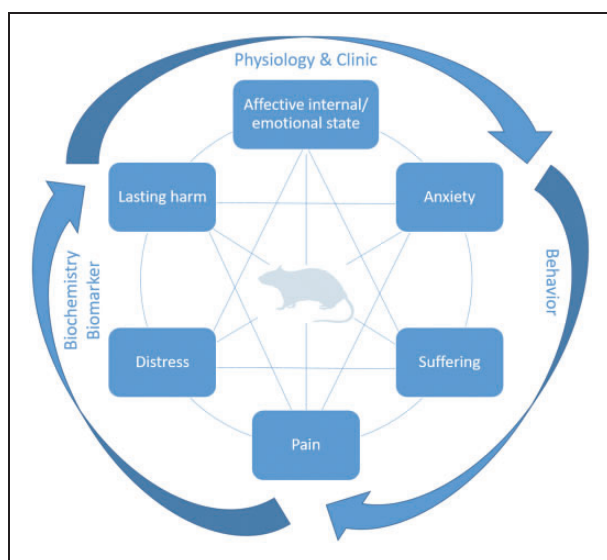
In the context of surgical animal models, severity assessment is mostly based on the subjective grading from a human perspective. This is due to the lack of objective, species-specific and intervention-specific assessment methods.<sup>5,6</sup> The commonly used score sheets are based on those developed by Morton et al. decades ago and have only been slightly adapted ever since.<sup>5</sup> Therefore, behavioral testing is a supplementary method to evaluate distress, pain, and severity in experimental research involving the use of living animals. Locomotor and real-time behavioral tests are the two primary approaches to investigate the behavior of an individual animal. The three commonly used test methods for this purpose are Open Field (OF), CatWalk (CW), and rotarod (RR) tests. The OF test investigates

voluntary space exploration and memory capacity in rodents. In addition, the duration of stay in segments or avoidance behavior can be analyzed, which falls within the domain of behavioral tests. The CW and RR tests characterize locomotion and coordination of a rodent. Herein, loading of individual limbs and step length are examples of parameters that can be measured using the CW test. On the other hand, latency to fall and fatigue resistance can be determined using the RR test.

The abovementioned test methods assess impairment in an animal during an experiment by evaluating spontaneous locomotor behavior and fitness. These methods are convenient for researchers and entail low personnel and financial costs; in addition, these involve a low risk of bias and are suitable for addressing surgical research questions.

The aim of this pilot study was to evaluate whether behavioral or locomotor tests (OF, RR, or CW) can help assess surgical severity after the implantation of a dummy telemetry transponder and laparotomy in rats. Then, the most suitable test method will be used in a future liver resection model to assess the severity of the surgical procedure and its treatment effect.

Therefore, we compared the abovementioned test methods in terms of their usability to assess the severity of laparotomy. The OF test is used to assess spontaneous locomotor and exploratory behaviors, whereas the CW test is used to assess locomotor behavior, which facilitates gait analysis. On the other hand, the RR test evaluates forced coordination and fatigue resistance in rats. In all three test methods, velocity can be analyzed as a common parameter. In this study, these locomotor and behavioral tests are hereafter referred to as “behavioral” tests.



**Figure 1.** Multimodal approach for severity assessment in animal-based research.

## Material and methods

### *Animals and ethical statement*

A total of 30 male Wistar Han rats (Janvier SAS, Saint-Berthevin, France) (mean body weight (BW):  $290 \pm 17$  g; age range: 6–8 weeks) were used in this study. Information on housing conditions and health monitoring are provided as supplementary materials.

The experiments were performed in accordance with the German animal welfare law (Tierschutzgesetz) and Directive 2010/63/EU pertaining to the protection of animals used for scientific purposes.<sup>1</sup> The official approval for this study was granted by the Governmental Animal Care and Use Committee (Protocol No.: 84-02.04.2017.A304; Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany). The study protocol complied with the Guide for the Care and Use of

Laboratory Animals.<sup>7</sup> Postoperative pain treatment is based on the recommendations of the German Society for Laboratory Animal Science as well as Initiative Veterinary Pain Therapy.<sup>8,9</sup>

### Experimental setup and performance of behavioral tests

After 1-week acclimatization, the rats were randomly allocated to three groups ( $n = 10$  per test method; RR, CW, or OF). Based on the group, the rats were trained for the respective test thrice every alternate training day (D1, 3, and 5). On D6, a dummy telemetry transponder (hereafter referred to as transponder implantation, TI) was surgically implanted (first surgery) followed by a 12-day recovery phase as recommended by the manufacturer without further behavior or locomotor testing. Retraining was conducted on D18 and 19, followed by the second and major surgery, that is sham laparotomy (hereafter referred to as SHAM) on D20. Behavioral tests were performed on postoperative days (PODs) 1, 3, 4, and 7. All tests were performed in the housing room and during the initial 3 h after the beginning of the light phase (Figure 2). Access to the housing room was restricted to female personnel only to avoid affection of behavioral or locomotor tests by male experimenter.<sup>10</sup>

BW was determined before each training, after each surgery, and once daily during scoring. Postoperative scoring was conducted three times a day on POD1–3 and once a day on POD4–7. On POD7, rats had a re-laparotomy under general anesthesia, as described below, and were euthanized during surgery by blood withdrawal from the inferior vena cava.

### RR test

The RR test involves an electronically and timer-controlled rod wherein rats are forced to walk on the RR.<sup>11,12</sup> This test (IITC Life Science, Los Angeles, CA, USA) was conducted for three cycles. Each cycle comprised two alternating runs at 10 and 10–20 rpm for 60 s

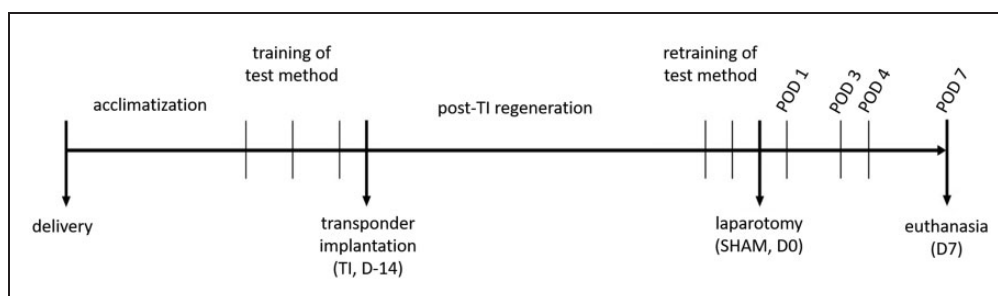
each followed by a recovery time of 60 s after each cycle, according to von Wrangel et al. (with minor modification of speed).<sup>13</sup> The recovery time began as soon as a rat tumbled or after the time span for the run was completed. The run time and distance were noted for each cycle, and velocity was calculated using the increasing speed (10–20 rpm) as a function of the covered distance in the corresponding time (velocity = distance/time). If a rat failed the RR test on the first training day, it was reassigned randomly to the CW or OF group. Therefore, the rats in the RR group were first tested and allocated to this group to retain 10 animals per group. No negative enforcement was used. After the completion of the respective test method, rats were returned to their housing cages.

### CW test

During the CW test, animals have to travel a specified distance (length: 1 m; CW XT gait pattern analysis system; Noldus, version 10.6, Wageningen, The Netherlands) by traversing a running glass surface through an unlit tunnel, and this act is simultaneously filmed from below. Thereafter, locomotor behavior (such as foot faults) and gait are analyzed based on the video.<sup>14,15</sup> The housing cage was placed on the opposite side at the same height to encourage rats to move through the tunnel. Velocity was analyzed using the Noldus software.

### OF test

The OF test was conducted according to referred studies, with minor dimensional modifications.<sup>13,16</sup> Rats were placed in the middle of the test field (L 72 × W 72 × H 40 cm; water-resistant plastic with dark underground) and then video recorded for 10 min (Media Recorder 4, Noldus, Wageningen, The Netherlands; camera: Camera GigE monochrome, 1/1"; lens: Lens Std CS mount, 4.5–12.5 mm 1/2", Basler AG, Ahrensburg, Germany) without further adaptation time. Analyses were performed using the Noldus



**Figure 2.** A schematic of the study timeline.

POD: postoperative day; SHAM: sham laparotomy; TI: transponder implantation.

EthoVision XT 14 software (Noldus, Wageningen, The Netherlands) with a focus on velocity. After each test run, the OF surface was cleaned and disinfected with wet wipes (Incidin Perfekt 5%, #104206E; Ecolab Deutschland GmbH, Germany) to remove not only feces and urine but also the odor of individual rats.

### Surgical procedures and treatment

Operations were randomized and performed always in the same timeframe under aseptic conditions in a separate operating room and under general anesthesia (induction: 5 vol% isoflurane + 5 L O<sub>2</sub>/min; maintenance: 2 vol% isoflurane + 2 L O<sub>2</sub>/min) with additional analgesia (metamizole/dipyrone; Novaminsulfon-ratiopharm<sup>®</sup> 1 g/2 mL; Ratiopharm GmbH, Germany; 100 mg/kg, subcutaneously (s.c.), single dose)<sup>9</sup> and antibiotics (cefuroxime, 16 mg/kg, s.c.). Blood samples were obtained from the vena sublingualis (TI) or vena cava inferior (SHAM and re-laparotomy/euthanasia).

Telemetric transponders (HD-S11; Data Sciences International, Minnesota, USA) were implanted into a subcutaneous pocket on the left flank, and a blood pressure catheter was inserted into the femoral artery.

Electrocardiogram electrodes were placed subcutaneously at the regio pectoralis (Figure 3).

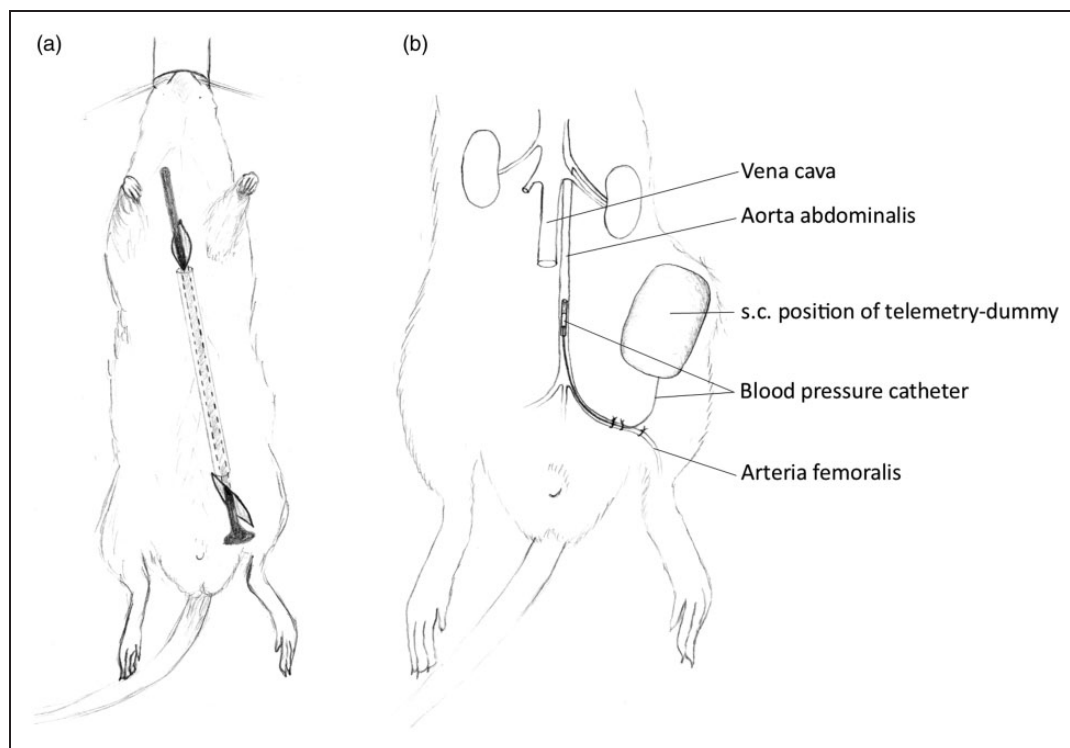
To mimic abdominal surgery, SHAM was performed using a midline incision (from the xiphoid to cranial pelvic brim, approximately 5 cm). The wound was left open for 15 min, followed by closure using two-layer continuous suture for the muscle layer and interrupted suture for the skin. Postoperative analgesia (metamizole, Novaminsulfon-ratiopharm<sup>®</sup> 1 g/2 mL; 400 mg/kg/day in sweetened drinking water) and antibiotics (cefuroxime, 16 mg/kg, s.c., once daily) were administered daily after both surgeries and on POD3. Pain medication was administered orally to reduce postoperative injections to once daily (only for antibiotics).

### Determination of BW

Determination of BW is provided as supplementary material.

### Scores for DS

DS was scored on a daily basis according to the scoring system described in 1985 by Morton et al.<sup>5</sup> to assess general conditions (DS=0-4 points (no distress), DS=5-9 points (mild distress), DS=10-19 points



**Figure 3.** (a) Ventral view of a rat showing subcutaneous tunneling for the placement of electrocardiogram leads at the regio pectoralis. (b) Ventral view of the regio inguinalis of the rat showing the implantation of a telemetric transponder in a subcutaneous (s.c.) pocket on the left flank; blood pressure catheter inserted via the arteria femoralis up to the aorta abdominalis.

(moderate distress), and  $DS \geq 20$  points (severe distress)).  $DS$  was determined on the respective postoperative examination day using the following four criteria: (a) BW; (b) overall state; (c) spontaneous locomotor behavior and readiness to walk; and (d) surgical procedure and wound healing. Spontaneous locomotor behavior was usually evaluated remotely prior to the evaluation of the other three criteria. Scores between 0 and 20 points could be obtained for each criterion, with increasing values indicating increasing  $DS$ . Rats with a sum of  $\geq 15$  points were not examined in the test setup on that day and were monitored more frequently. A time limit of 24 h was set for a score of 15 points. Scores  $\geq 20$  points led to a sudden removal of the animals from the experiment and euthanasia, due to the predefined humane endpoint. However, this did not occur at any time.

### Measurement of serum corticosterone and its fecal metabolites

Blood and fecal samples were used to detect the level of corticosterone or its metabolites. Fecal samples were collected (always in the same timeframe in the morning) after being weaned during behavioral tests and during surgical procedures. The samples were analyzed using an enzyme immunoassay as described previously.<sup>17,18</sup> The levels of fecal corticosterone metabolites (FCMs) are expressed as micrograms/gram feces. Blood samples were obtained under general anesthesia during surgical procedures, and serum was obtained and analyzed using the MagPix<sup>®</sup> multiplex analyzer (Assay-Kit: Rat Stress Hormone Magnetic Bead Panel, #RSHMAG-69K, Merck, Germany) according

to the manufacturer's instructions. Serum corticosterone levels are expressed as nanograms per milliliter.

### Statistical analysis

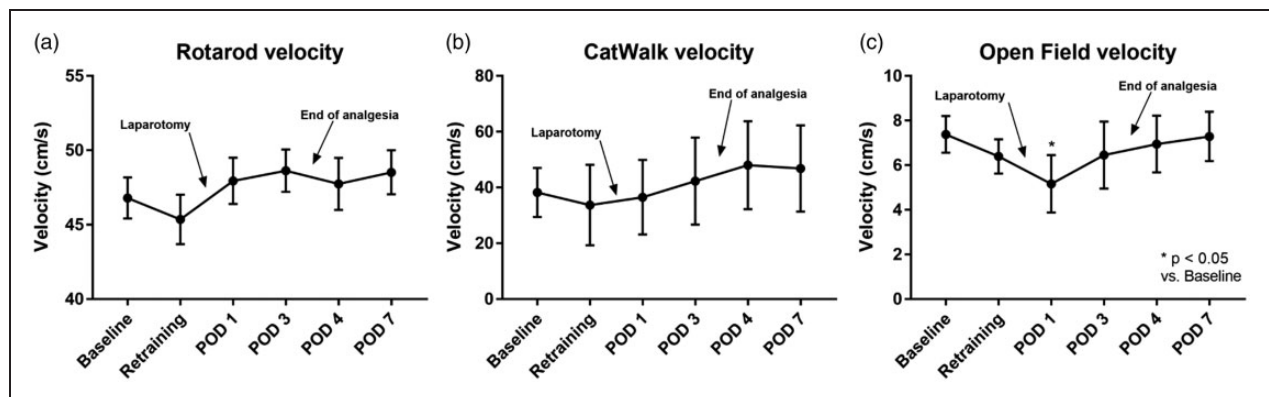
Detailed information on statistical analysis is provided as supplementary material.

## Results

### Performance in behavioral tests

**RR test.** In the RR group, several rats ( $n = 14$ ) had to be reassigned to other groups due to the lack of learning ability of movement on the RR. The analysis of velocity showed a slight decrease in velocity during retraining with an increase on POD1 compared with baseline velocity. Between POD1 and 7, almost no changes were detected in velocity, with no significant differences compared with baseline velocity (Figure 4 (a)). Analysis using area under the curve (AUC) of the receiver operating characteristic (ROC) was neither sensitive nor specific in detecting changes due to surgical intervention (AUC = 0.65;  $CI_{95\%} = 0.3868-0.9243$ ,  $p = 0.2530$ ).

**CW test.** In the CW group, the analysis of velocity showed a slight decrease in velocity from baseline to retraining. From POD3 onward, the mean velocity was consistently greater than baseline velocity; however, there was no significant difference in velocity at any time point compared with baseline velocity (Figure 4 (b)). AUC of ROC was neither sensitive nor specific in detecting changes due to surgical intervention (AUC = 0.51;  $CI_{95\%} = 0.188-0.8324$ ,  $p = 0.9491$ ).



**Figure 4.** Moving velocity in test performances (mean  $\pm$  standard deviation). Results of retraining are averaged and shown as "retraining": (a) Rotarod: 3 runs for 60 s with 10–20 rpm in 30 s, repeated measures one-way analysis of variance ( $F[2,152,62.42] = 3.416$ ;  $p = 0.0359$ ), Dunnett's post hoc test; (b) CatWalk: repeated measures one-way analysis of variance ( $F[2,932,17.59] = 3.641$ ;  $p = 0.034$ ), Dunnett's post hoc test; (c) Open Field: repeated measures one-way analysis of variance ( $F[3,177,28.59] = 5.372$ ;  $p = 0.0041$ ), Dunnett's post hoc test. POD: postoperative day.

**OF test.** In the OF test, velocity decreased from baseline to retraining until it reached its lowest value on POD1 ( $p < 0.05$ ). However, it increased and reached baseline velocity on POD7 (Figure 4(c)). AUC of ROC could detect changes due to surgical intervention (AUC = 0.8; CI<sub>95%</sub> = 0.5799–1.02,  $p = 0.0404$ ).

### Determination of BW

The rats in the OF and CW groups experienced a significant BW loss on POD1 ( $p < 0.05$ ) compared with their baseline BW. This phenomenon was not observed in the RR group. After this time point, all groups showed a continuous increase in BW and exceeded their postoperative BW on POD4 (Figure 5(a)).

### Scores for DS

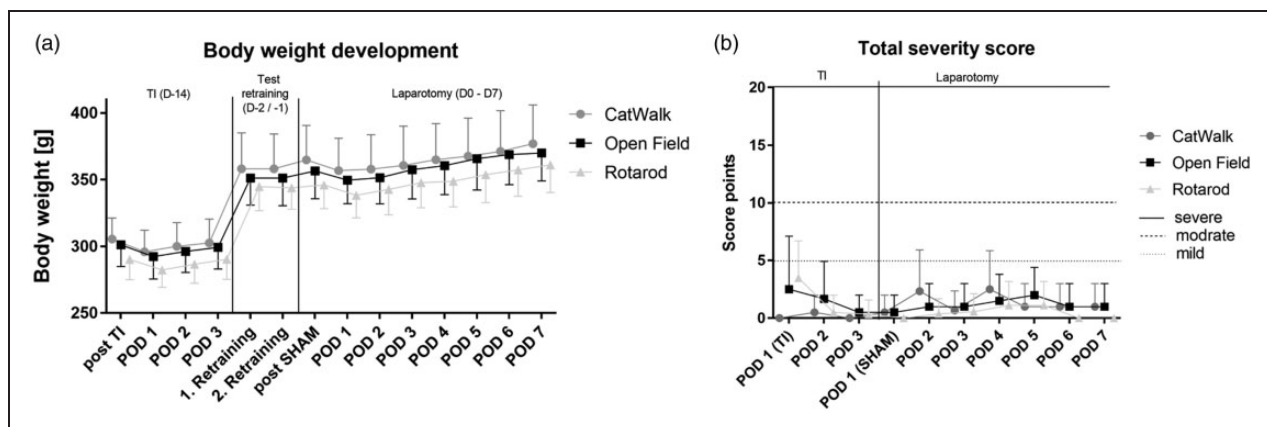
Except for one rat, all rats reached the planned end of the study, that is POD7. Due to a technical failure during laparotomy, one rat in the RR group had to be euthanized prematurely (on the same day) because the humane endpoint (opening of wound sutures and intestinal prolapse) was achieved after 2 h postoperative. Therefore, the rat was excluded from the analysis. Evaluation using DS score sheets showed no significant differences with respect to the analyzed parameters. Median analysis showed no significant difference between any time points or groups (Figure 5(b)).

### Analysis of FCM levels

No significant differences were observed with respect to serum corticosterone levels among the groups at all time points or among different time points within the same group (Figure 6(a)). All groups showed significantly higher FCM levels than the baseline levels (Figure 6(b)). After SHAM, the levels increased and reached the highest levels on POD1. However, the levels decreased from POD3 to 4. Thereafter, the levels decreased and approached the baseline levels again on POD7, and no significant differences were detected. FCM levels after POD2 approximated the preoperative FCM levels. The end of analgesia did not increase FCM levels, and the levels continued to decrease toward the end of the study.

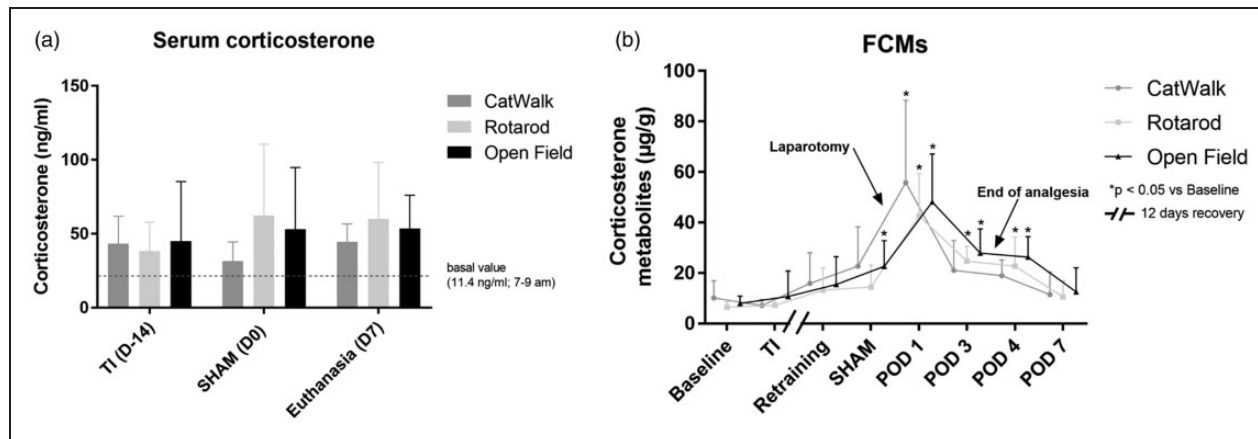
### Discussion

With the implementation of Directive 2010/63/EU in Europe, severity assessment of animal experiments has become an essential part of the approval process. However, severity assessment is typically based on subjective parameters that are not backed by robust evidence.<sup>2,19</sup> Therefore, the implementation of more objective parameters for severity assessment is imperative with the evaluation of the currently used severity parameters for animals in experimental research. The scientific hypothesis of this study was that surgical intervention involving midline laparotomy would cause a measurable effect on the applied behavioral tests (OF, RR, or CW) for severity assessment of



**Figure 5.** (a) Body weight (mean  $\pm$  standard deviation) changes after transponder implantation (D–14), retraining (D–2) and laparotomy [from D0 to POD7]; time span of 21 days; using body weights right after surgery as baseline values (100%): two-way analysis of variance ( $F(7,208) = 40.68$ ;  $p < 0.05$ ). Multiple comparisons of BW on different postoperative days with baseline body weight were performed using Dunnett's post hoc test. (b) Total score (body weight, general condition, wound healing, and spontaneous locomotor behavior) of groups shown as median and upper limit on postoperative days after surgeries (TI till POD3; laparotomy till POD7) with gradual allocation of severity [mild: 5 points, moderate: 10–15 points, and severe:  $\geq 20$  points]. [TS: There is a small typo in Figure 5. Please change "modrate" to "moderate" in panel (b).]

POD: postoperative day; SHAM: sham laparotomy; TI: transponder implantation.



**Figure 6.** (a) Serum corticosterone levels (mean  $\pm$  standard deviation): two-way analysis of variance, comparison of columns within each row ( $F(2,65) = 0.7723$ ;  $p = 0.4661$ ), Tukey's post hoc test; samples obtained during surgery (TI/D-14, SHAM/D0 and re-laparotomy/D7) under general anesthesia. (b) Levels (mean  $\pm$  standard deviation) of fecal corticosterone metabolites: two-way analysis of variance, comparison of rows within each column ( $F(7,201) = 34.32$ ;  $p = 0.0001$ ), Dunnett's post hoc test.

FCM: fecal corticosterone metabolite; POD: postoperative day; SHAM: sham laparotomy; TI: transponder implantation.

laparotomy in rats. This hypothesis was based on the clinical observation that humans and animals experience significant locomotor impairment and reduced quality of life after an abdominal surgical intervention.<sup>20</sup> In this study, no significant difference was observed in postoperative velocity in the RR and CW groups compared with the respective baseline velocity; in addition, AUC of ROC showed poor sensitivity and specificity. In contrast, we observed a significant decrease in the postoperative velocity curve in the OF group compared with baseline velocity. These findings are consistent with the measurements of FCM levels and BW during the postoperative phase. The exact reason for the observed differences in terms of BW loss in the OF and RR groups compared with that in the CW group remains unclear. We hypothesize that the OF and RR tests are more energy-consuming procedures due to the approximate 10-min test duration than the CW test ( $< 6$  s). Several studies have demonstrated the ability of the CW,<sup>21,22</sup> OF,<sup>23,24</sup> and RR<sup>12,25</sup> tests to assess distress or behavioral changes in rats. Thus, we subjected animals to the RR test after midline laparotomy. Regarding animal welfare, we decided not to use negative enforcement for the RR test (e.g. electric shocks) if an animal fell or jumped on the platform bottom. However, this is a potential cause for selection bias within the RR group because only the animals that learned to move on the RR were included in the RR group. However, the use of only the rats that performed in the RR and its depending selection bias was chosen on purpose, because the assessment of severity with negative enforcement could lead to severity itself. Remarkably, in contrast to the findings

reported by Whishaw et al.,<sup>26</sup> we did not observe post-laparotomy locomotor or behavioral impairment in the RR group under analgesia with metamizole. Even after midline laparotomy, moving velocity increased, although the results were not statistically significant. In line with the results of the RR test, moving velocity in the CW system also increased after laparotomy; however, this velocity was significantly different from baseline velocity. This finding is in contrast to that of a previous study that employed the CW system for the assessment of pain behavior in a rat model of intervertebral disk injury.<sup>21</sup> The OF test clearly showed a significant decrease in moving velocity after laparotomy, consistent with the findings reported by Cittolin-Santos et al.;<sup>23</sup> they also demonstrated the ability of the OF test to detect behavioral changes after surgical intervention. To summarize our results, the OF test was the only test that indicated a change in behavior. This is supported by the fact that the OF test allows a higher degree of freedom in movement. This is further supported by the ROC curve analysis, wherein only the OF test was significantly different from random testing. Additionally, the OF test can be used for further behavioral testing such as grooming and exploration.<sup>27,28</sup>

In 1993, Flecknell et al.<sup>29</sup> reported that "a relatively simple surgical procedure (laparotomy) results in a major reduction in food and water consumption in rats," thereby leading to BW loss. This observation is consistent with the observed postoperative decrease in BW in all groups in the present study. However, the baseline values were achieved after POD4 in all groups.

To assess whether the results of BW analysis were affected by the test method used, we compared these

with the postoperative BW curves reported by Kanzler et al.<sup>2</sup> We observed a similar progression in both studies, which suggests that the postoperative BW loss was primarily influenced by the surgical procedure and not by the test method. In addition, in the study by Kanzler et al., none of the experimental groups showed scores exceeding 10 points, which corresponds to the upper limit of moderate distress. They also showed only a mild-to-moderate postoperative load after liver resection in rats.

In the present study, we could not determine any measurable influence of the behavioral test methods on the assessed severity of impairment in animals, regardless of the test method or surgical procedure. There is a species-specific time delay between increased plasma cortisol levels and respective FCM levels, which peaks approximately 12 h after acute stress in rats.<sup>17,30,31</sup> In our model, FCM levels were the highest ( $p < 0.0001$ ) on POD1. In comparison with TI and SHAM values (here feces were taken during surgery, representing the impact of the test on the day before the surgery), POD1 represents the impact of the prolonged corticosterone secretion during SHAM (due to the species-specific time delay of the metabolism). In contrast to FCM levels, serum corticosterone levels showed no correlation with any surgical procedure or time point. Further, we observed no correlation between serum corticosterone and FCM levels (data not shown). This may be attributed to the fact that blood collection is stressful in general; serum corticosterone levels are estimated at a point-in-time, are only shortly elevated after intervention (minutes to hours),<sup>32</sup> and do not depend on the duration and degree of stress. This is supported by the results reported by Palme,<sup>30</sup> who showed that FCMs constitute a better and non-invasive marker for acute stress compared with single blood samples. Therefore, this parameter should be recommended on the basis of the 3Rs.

In conclusion, the OF test can help detect the severity of abdominal surgical intervention in rats, and the results are consistent with the measurements of FCM level and BW. Furthermore, the OF test showed the highest sensitivity and specificity for assessing the severity of laparotomy. It should be noted that the current study was designed as a pilot study to investigate which behavior/locomotor tests are most suitable to assess severity within a subsequent study of partial hepatectomy. According to the 3R principles, the animals of this pilot study will serve as a control group to evaluate the additional impact of the hepatectomy in the main study. Therefore, no actual telemetric devices but dummies were used to mimic the influence of the operation with regard to changed running performance with the device. However, we are aware of the limits of the study, for example the impact of the first surgery

(TI) which is not addressed here. Moreover, the influence of sex has not yet been taken into account. It should be pointed out that there is an existing sex gap in the current preclinical research culture.<sup>33</sup>

Therefore, further research is required to study the impact of sex and more specific surgeries on the well-being of laboratory animals. The extent to which comparable surgical procedures with organ-specific interventions, such as organ resection or transplantation, influence results and how gradations of severity assessment can be made by using the OF must be investigated in future studies. Further, we assume that the use of the OF test could enable the evaluation of the severity of a new surgical procedure within pilot studies in the future. Then the prospective classification of severity would be based on objective data and not based only on subjective parameters. The OF test may allow a more objective severity assessment in animal experiments and therefore can be recommended in the future.

#### Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


#### Funding


The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Deutsche Forschungsgemeinschaft, TO 542/5-1.

#### Acknowledgements

We are thankful to Pascal Paschenda and Pramod Kadaba Srinivasan for skillful technical assistance and to Edith Klobetz-Rassam for FCM analysis.

#### ORCID iDs

Steven R Talbot  <https://orcid.org/0000-0002-9062-4065>

Michael Czaplak  <https://orcid.org/0000-0002-7118-7728>

René H Tolba  <https://orcid.org/0000-0002-0383-3994>

#### References

1. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. 2010.
2. Kanzler S, Rix A, Czigany Z, et al. Recommendation for severity assessment following liver resection and liver transplantation in rats: part I. *Lab Anim* 2016; 50: 459–467.
3. Pinkernell S, Becker K and Lindauer U. Severity assessment and scoring for neurosurgical models in rodents. *Lab Anim* 2016; 50: 442–452.



4. Jirkof P, Abdelrahman A, Bleich A, et al. A safe bet? Inter-laboratory variability in behaviour-based severity assessment. *Lab Anim* 2020; 54: 73–82.
5. Morton DB and Griffiths PH. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Vet Rec* 1985; 116: 431–436.
6. Oliveira M, Nascimento-Goncalves E, Silva J, et al. Implementation of humane endpoints in a urinary bladder carcinogenesis study in rats. *In Vivo* 2017; 31: 1073–1080.
7. National Research Council Committee. *Guide for the care and use of laboratory animals*. 8th ed. Washington, DC: National Academies Press, 2011.
8. Tacke S, Henke J and Erhardt W. Metamizol (dipyrone) for pain therapy. A literature review and own clinical experiences. *Tierarztl Prax* 2008; 36: 19–25.
9. Committee on Anaesthesia of GV-SOLAS Supported by The working group 4 of TVT (German veterinary association for animal welfare). Pain management for laboratory animals. 2015. [http://www.gv-solas.de/fileadmin/user\\_upload/pdf\\_publication/Anaest.\\_Analgesie/Schmerztherapie\\_Mai2015\\_e.pdf](http://www.gv-solas.de/fileadmin/user_upload/pdf_publication/Anaest._Analgesie/Schmerztherapie_Mai2015_e.pdf).
10. Sorge RE, Martin LJ, Isbester KA, et al. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat Methods* 2014; 11: 629–632.
11. TSE Systems. RotaRod: motor function & coordination. 2019.
12. Hamm RJ, Pike BR, O'Dell DM, et al. The rotarod test: an evaluation of its effectiveness in assessing motor deficits following traumatic brain injury. *J Neurotrauma* 1994; 11: 187–196.
13. von Wrangel C, Schwabe K, John N, et al. The rotenone-induced rat model of Parkinson's disease: behavioral and electrophysiological findings. *Behav Brain Res* 2015; 279: 52–61.
14. Gabriel AF, Marcus MA, Honig WM, et al. The CatWalk method: a detailed analysis of behavioral changes after acute inflammatory pain in the rat. *J Neurosci Methods* 2007; 163: 9–16.
15. Deumens R, Van Gorp SF, Bozkurt A, et al. Motor outcome and allodynia are largely unaffected by novel olfactory ensheathing cell grafts to repair low-thoracic lesion gaps in the adult rat spinal cord. *Behav Brain Res* 2013; 237: 185–189.
16. Schulz D. Acute food deprivation separates motor-activating from anxiolytic effects of caffeine in a rat open field test model. *Behav Pharmacol* 2018; 29: 543–546.
17. Lepschy M, Touma C, Hruby R, et al. Non-invasive measurement of adrenocortical activity in male and female rats. *Lab Anim* 2007; 41: 372–387.
18. Touma C and Palme R. Measuring fecal glucocorticoid metabolites in mammals and birds: the importance of validation. *Ann N Y Acad Sci* 2005; 1046: 54–74.
19. Moller C, Wolf F, van Dijk RM, et al. Toward evidence-based severity assessment in rat models with repeated seizures: I. Electrical kindling. *Epilepsia* 2018; 59: 765–777.
20. Strik C, Stommel MWJ, Hol JC, et al. Quality of life, functional status and adhesiolysis during elective abdominal surgery. *Am J Surg* 2018; 215: 104–112.
21. Miyagi M, Ishikawa T, Kamoda H, et al. Assessment of pain behavior in a rat model of intervertebral disc injury using the CatWalk gait analysis system. *Spine* 2013; 38: 1459–1465.
22. Sakuma T, Kamoda H, Miyagi M, et al. Comparison of CatWalk analysis and von Frey testing for pain assessment in a rat model of nerve crush plus inflammation. *Spine* 2013; 38: E919–E924.
23. Cittolin-Santos GF, Guazzelli PA, Nonose Y, et al. Behavioral, neurochemical and brain oscillation abnormalities in an experimental model of acute liver failure. *Neuroscience* 2019; 401: 117–129.
24. Boada MD, Martin TJ and Ririe DG. Nerve injury induced activation of fast-conducting high threshold mechanoreceptors predicts non-reflexive pain related behavior. *Neurosci Lett* 2016; 632: 44–49.
25. Rodrigues TA, Sampaio AJBJ, Nunes IDP, et al. Effect of strontium ranelate on pain behavior in an experimental model of osteoarthritis. *Braz J Med Biol Res* 2017; 50: e6314.
26. Wishaw IQ, Li K, Wishaw PA, et al. Use of rotorod as a method for the qualitative analysis of walking in rat. *J Vis Exp* 2008; 22: 1–4.
27. Asano Y. Characteristics of open field behavior of Wistar and Sprague-Dawley rats. *Jikken Dobutsu* 1986; 35: 505–508.
28. Whimbey AE and Denenberg VH. Two independent behavioral dimensions in open-field performance. *J Comp Physiol Psychol* 1967; 63: 500–504.
29. Flecknell PA, Orr HE, Roughan JV, et al. Comparison of the effects of oral or subcutaneous carprofen or ketoprofen in rats undergoing laparotomy. *Vet Rec* 1999; 144: 65–67.
30. Palme R. Non-invasive measurement of glucocorticoids: advances and problems. *Physiol Behav* 2019; 199: 229–243.
31. Royo F, Bjork N, Carlsson HE, et al. Impact of chronic catheterization and automated blood sampling (Accusampler) on serum corticosterone and fecal immunoreactive corticosterone metabolites and immunoglobulin A in male rats. *J Endocrinol* 2004; 180: 145–153.
32. Yarigholi F, Zare Mehrjardi A, Azizi Z, et al. An analysis of the hypothalamic-pituitary-adrenal axis functions in cirrhotic rats in response to surgical stress. *Surg Res Pract* 2018; 2018: 7606304.
33. Beery AK and Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev* 2011; 35: 565–572.

## Résumé

Cette étude pilote vise à déterminer si les tests comportementaux ou locomoteurs [Open Field (OF), rotarod (RR) et CatWalk (CW)] peuvent permettre d'évaluer la sévérité de la laparotomie pratiquée chez les rats.

La nouvelle directive européenne 2010/63/UE impose une évaluation de la gravité dans les expériences réalisées sur les animaux. Toutefois, les méthodes validées et objectives doivent établir un rapport entre les résultats de la surveillance spécifique à l'essai et le degré de souffrance causée aux animaux concernés. Par conséquent, nous sommes concentrés sur des méthodes d'évaluation de la gravité qui sont simples, commodes et soit non invasives, soit mini-invasives dans le cadre d'un modèle chirurgical.

Afin d'évaluer la gravité chirurgicale de ce modèle, nous avons comparé la vitesse de déplacement dans trois méthodes de tests comportementaux couramment utilisées (OF, RR et CW) après une laparotomie, dans la période opératoire de 7 jours.

Dans le cadre de cette étude, 30 rats Wistar Han adultes mâles ( $n = 10/\text{test}$ ) ont été entraînés à la méthode de test qui leur avait été affectée, puis ils ont ensuite subi l'intervention chirurgicale. Le degré de gravité a été calculé quotidiennement au moyen d'une feuille de pointage, élaborée antérieurement et modifiée. En outre, des échantillons de matières fécales et de sang ont été prélevés afin d'analyser les niveaux de métabolites de corticostérone chirurgicaux et postopératoires. Nous avons constaté des différences importantes au sein du groupe expérimental en termes de paramètres analysés. Dans ce contexte, il a été conclu que le test OF était la méthode la mieux adaptée à l'évaluation de la gravité chez les rats après une laparotomie.

## Abstract

Ziel dieser Pilotstudie war es, zu evaluieren, ob Verhaltens- oder Lokomotionstests [Open Field (OF), Rotarod (RR) und CatWalk (CW)] zur Beurteilung des Schweregrades einer Laparotomie bei Ratten dienen können.

Die neue EU-Richtlinie (2010/63/EU) schreibt die Bewertung des Schweregrades bei Tierversuchen vor. Es bedarf jedoch validierter und objektiver Methoden, um die Ergebnisse der studienspezifischen Überwachung mit dem Grad der Belastung für einzelne Tiere in Beziehung zu setzen. Daher haben wir uns auf nicht-invasive oder minimal-invasive, einfache und praktische Methoden zur Bewertung des Schweregrades in einem chirurgischen Modell konzentriert.

Um den chirurgischen Schweregrad in diesem Modell zu beurteilen, verglichen wir die Bewegungsgeschwindigkeit zwischen drei häufig angewandten Verhaltenstestmethoden (OF, RR und CW) nach einer Mittellinienlaparotomie innerhalb von 7 Tagen nach der Operation.

In der vorliegenden Studie wurden 30 adulte männliche Wistar-Han-Ratten ( $n = 10/\text{Test}$ ) in der ihnen zugewiesenen Testmethode trainiert und anschließend einer Operation unterzogen. Die Schweregradbewertung wurde täglich unter Verwendung eines zuvor entwickelten modifizierten Score Sheets durchgeführt. Zusätzlich wurden Blut- und Kotproben entnommen, um die chirurgischen und postoperativen Kortikosteron-Metabolitenwerte zu analysieren. Wir ermittelten signifikante Unterschiede zwischen den Versuchsgruppen hinsichtlich der analysierten Parameter. Im vorliegenden Kontext erwies sich der OF-Test als die am besten geeignete Methode zur Schweregradbeurteilung nach Laparotomie bei Ratten.

## Resumen

El objetivo de este estudio piloto es evaluar si las pruebas locomotoras y de comportamiento [Open Field (OF), rotarod (RR) y CatWalk (CW)] pueden ayudar a evaluar la gravedad tras una laparotomía en ratas.

La nueva Directiva de la UE (2010/63/UE) exige realizar una evaluación de gravedad en los experimentos con animales. No obstante, se requieren métodos validados y objetivos para vincular los resultados de un control específico de las pruebas con el grado de molestia causada a animales individuales. Por tanto, nos hemos centrado en métodos de evaluación de gravedad convenientes, simples y no invasivos/mínimamente invasivos con un modelo quirúrgico.

Para evaluar la gravedad quirúrgica en este modelo, hemos comparado la velocidad de movimiento entre tres métodos de pruebas de comportamiento normalmente utilizados (OF; RR y CW) tras una laparotomía de línea mediana en un plazo de 7 días tras la operación.

En este estudio, 30 ratas adultas Wistar Han macho (N = 10/prueba) fueron formadas en su método de prueba asignado y posteriormente se sometieron a una intervención quirúrgica. La calificación de la gravedad fue realizada a diario utilizando una hoja de calificación modificada que se creó anteriormente. Asimismo, se recogieron muestras fecales y sanguíneas para analizar los niveles de metabolitos de corticosterona post-operativos y quirúrgicos. Encontramos diferencias significativas entre los grupos de experimentación en cuanto a los parámetros analizados. En este contexto, la prueba OF se consideró el método más adecuado para la evaluación de gravedad tras una laparotomía en ratas.