

Effects of Multimodal Therapy, Blinding, and Multi-laboratory Protocol Conduct on the Robustness of the Rat Model of Adjuvant Induced Arthritis

DASA SEVELJEVIC-JARAN^{1,2}, MAJA ANTOLIC¹, ANJA OGNJENOVIC¹,
OTTO KALLIOKOSKI², KLAS S.P. ABELSON² and JANN HAU²

¹*In Vivo Pharmacology and Toxicology Department and Toxicology and Laboratory Animal Science Department, Selvita Ltd., Zagreb, Croatia;*
²*Department of Experimental Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark*

Abstract. *Background/Aim: The aim of this study was to investigate the effects of multimodal therapy comprising buprenorphine (BUP) and indomethacin (IND) on key translational parameters in the rat adjuvant induced arthritis (AIA) model. Furthermore, we investigated the difference between visual assessment scores and histology scores generated by blinded and non-blinded assessors and the robustness and generalizability of results by conducting a multi-laboratory study. Materials and Methods: The experiment was terminated on day 26 after 11 days (days 15-25) of voluntarily ingested buprenorphine and 7 days of gavage delivered indomethacin treatment (days 19-25). The treatment effects were assessed on the last day of the study, relying on body weight assessment, serum concentrations of $\alpha 1$ -acid glycoprotein, and assessment of affected hind paws swelling, in-life and post mortem. Results: Across two laboratories, the combined analgesic treatments had minimal effects on the measured model parameters indicating that multimodal treatment did not affect the translatability of the model. We found an improvement in clinical scores (a negative change in scores) in nearly all medicated animals when scored informed, whereas it was essentially 50:50 for*

the blinded scorings and no difference between the blinded and informed histological scoring. Conclusion: The present results support the use of more effective analgesic treatment regimens and the good practice recommendations advocating blinding as a mandatory practice in conduct of preclinical in vivo efficacy studies. In spite of minor differences between results obtained at the two sites, there was good agreement between them indicating robustness of the AIA model.

In humans, chronic autoimmune diseases, such as rheumatoid arthritis (RA), affect women more often than men, the sex ratio typically being around 3:1 (1). A typical model for studying the pathology of RA, and potential treatments in rats, is the adjuvant induced arthritis (AIA) model (2, 3). The AIA model is categorized as moderate to severe in the severity assessment framework of Directive 2010/63/EU. Animals modelling painful human conditions should not be deprived of the analgesic care that the modelled human could expect. Consequently, the pain relief therapy in RA animal models should mimic the clinical management of humans with a similar condition. Untreated pain and stress may also interfere with the experimental outcomes and translational properties of the model (4, 5). Nonsteroidal anti-inflammatory drugs (NSAIDs) are the therapy of choice for symptomatic treatment of RA. They are known to modulate the inflammatory response and to decrease the disease severity in both animals and man. In humans, opioids are often the therapy of choice in severe pain management and may be used in animal models as an alternative or addition to NSAID therapy. Buprenorphine has been shown to be less immunosuppressive than other opioids, which makes it a suitable candidate for pain management in rodent models of inflammation and pain (6, 7). Pain relief in rodents is, however, often restricted to a single agent rather than the multimodal approaches that are the current standard for human patients (8, 9).

Correspondence to: Dasa Seveljevic-Jaran, DVM, MSc LAS, Toxicology and Laboratory Animal Science Department, Selvita Ltd., 10000 Zagreb, Croatia. Tel: +385 8886436, e-mail: dasa.seveljevicjaran@selvita.com

Key Words: Rat, adjuvant induced arthritis, buprenorphine, multimodal therapy, blinding, multi-laboratory study.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) 4.0 international license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

The value, reliability and reproducibility of animal models has come under scrutiny over the past decade because of multiple reports documenting their poor translatability to human studies (10-16). Reproducibility depends on experimental design and conduct, which together account for the internal and external validity of experimental results (17-19). Controlling all major sources of bias is a central issue in the design of an experiment. These are context-dependent, and some are specific to the animal model under investigation, and to the laboratory where the study is conducted (20, 21). Biases generally stem from inadequacies in the design, conduct, and analysis of an experiment. They significantly confound *in vivo* studies, reducing their internal validity (22-24).

The reproducibility of results from animal experiments is poor across many research areas. Moreover, seemingly successful therapies in animal models frequently fail in translating to efficacious treatments in clinical trials. Systematic reviews of clinical and preclinical research have convincingly shown that inadequate design, conduct and reporting of study quality attributes (*e.g.*, randomization, blinding *etc.*) correlate with substantially larger treatment effects (25-31).

The aim of the present study was to investigate the effects of two combined analgesics – buprenorphine and indomethacin – on key translational parameters in the rat AIA model. We also investigated the difference between scores generated by blinded and non-blinded assessors assessing the effects of the different therapies. Specifically, we focused on the swelling of the arthritis-affected hind paws and on their histopathological findings, *post mortem*. Since women are more susceptible to RA, we chose to focus this study on female rats. Furthermore, the study protocol was carried out in two facilities/laboratories, one located in Croatia and another in Israel, to ensure that our findings were robust and generalizable.

Materials and Methods

Ethical statement. The present studies were conducted at two different sites. One was conducted in 2018 in the AAALAC-accredited animal facility at Fidelta Ltd. (currently Selvita Ltd., legal successor of Fidelta Ltd.) in Zagreb, Croatia. A confirmatory study using the same study protocol in 2019 was carried out in an AAALAC-accredited Sharett facility at Hebrew University Medical School, Ein Kerem campus, Jerusalem, Israel. The project authorization (KLASA UP/I-322-01/17-01/162, URBROJ: 525-10/0255-18-4) was issued in Croatia by the National Competent Authority, the Ministry of Agriculture, Animal Protection Department, Veterinary and Food Safety Directorate in consultation with the National Ethics Committee and in compliance with the Directive 2010/63/EU. Project authorization (MD-17-14881-4) for conducting the study in Israel was issued in advance by the Hebrew University Ethics Committee.

Animals and housing. Fidelta facility (Croatia): Seven-weeks-old female (n=32) SPF Lewis rats (LEW/Crl; Charles River, Sulzfeld, Germany) were pair-housed in open-top cages (“1291H”; 425×266×185 mm; Tecniplast, Buguggiate, Italy) on arrival from

the breeder. The animal holding rooms were maintained at 22±2°C, and 55±10% relative humidity, with ≥15 h⁻¹ air changes. A diurnal rhythm was kept through a 12:12 h dark/light cycle (lights on at 7:00 AM), with a daytime intensity of 150 Lux at 1 m over the floor. The conventional cages (Tecniplast Type III H) were structured with raised lids; bedding consisted of corncob grit (“Scobis Due”; Mucedola, Milan, Italy) and enrichment items were provided in the form of cardboard shelters and wooden gnawing blocks (LBS Biotechnology, Horley, UK). Feed pellets (“VRF1”; SDS Diets, Horley, UK) and tap water were provided *ad libitum*. Pellets were also placed on the bedding for easier access for arthritic animals. Cages (including water) were changed twice a week, transferring a handful of used bedding to maintain olfactory cues.

Sharett facility (Israel): Seven and eight-weeks-old female (n=32) SPF Lewis rats (LEW/SsNHsd, Envigo, Jerusalem, Israel) were group-housed (4 per cage) in Tecniplast Greenline GR900 racks for rats. The animal holding rooms were maintained at 22±2°C, and 55±10% relative humidity, with ≥15 h⁻¹ air changes. A diurnal rhythm was kept through a 12:12 h dark/light cycle (lights on at 7:00 AM). The cages (Seal Safe Plus, cage floor area 904 cm²) were provided with wood chip bedding (autoclaved Teklad sani-chips, Envigo) and rodent maintenance diet formula (gamma irradiated, Teklad 2918SC4 Envigo). Enrichment was provided in the form of an autoclaved grey PVC tunnel (cross section diameter 12 cm, replaced once a month). The municipal water was locally processed (acidified to pH 2.8-3.2) and provided *ad libitum*. Pellets were also placed on the bedding for easier access for arthritic animals. Cages (including water) were changed once a week.

Model induction. Following 5-7 days of acclimatization, arthritis was induced through adjuvant injections (marking day 1 of the experiment). Complete Freund’s adjuvant (CFA) was prepared by suspending 10 mg/ml heat-inactivated *M. tuberculosis* (Myc t, “Difco H37 Ra”, Fisher Scientific, Hampton, VA, USA) in an oil vehicle (“Freund’s adjuvant, incomplete”, prod. no. F5506; Sigma-Aldrich, St. Louis, MO, USA). Rats were anesthetized with 4% isoflurane (“Forane”, Abbot, Chicago, IL, USA was used in Croatia and Isoflurane, USP, Terell, Piramal critical care Inc., Bethlehem, PA, USA was used in Israel) delivered in pure oxygen, in an induction chamber. Anesthesia was maintained on a facemask (with 3% isoflurane) while a small area of the lower back was shaved with electric clippers. A 100 µl injection of the adjuvant was administered intradermally, approximately 1 cm cranially to the tail base, before the subject was returned to its cage for recovery.

Stratification (inclusion/exclusion criteria, humane end points). On day 18 post-induction, rats not developing early stages of arthritis were excluded from the study and sacrificed. These non-responders were identified based on body weight (gain), absence of bilateral midfoot edema and lack of a visible limp when allowed to explore the flat surface of a bench top.

Moreover, animals were excluded from further research and euthanised if they reached one of the following endpoints at any point during the study: extreme paw swelling encompassing the entire foot and ankle; ulceration of the immunisation site; weight loss up to 25% of baseline that did not reverse within 5 days.

Treatments. Animals were randomly assigned to receive treatment, twice daily, with unmedicated (the negative control group) and buprenorphine-medicated nut paste for voluntary ingestion (*v.i.*) on days 15-25. The medicated paste was prepared in several larger

batches and, following the recommendations in Abelson *et al.* (32), kept refrigerated at 4–8°C for up to a week. Rats were conditioned to eat the nut paste for 7 days, once a day, prior to the therapy start. Rats demonstrated high motivation to eat the nut paste, even the medicated nut paste, regardless of their health/disease status. In Croatia, Buprenorphine Sandoz 8 mg sublingual tablets (Salutas Pharma GmbH, Barleben, Germany, Batch no. HK0071) were used. In Israel, Subutex 8 mg sublingual tablets (Indivior UK Ltd, Slough, Berkshire, UK, Batch no. 818613) were used. The tablets were mortared to a fine powder and mixed into the nut paste in a kitchen blender (0.2 mg buprenorphine per 1 g of paste). A dose of 0.4 mg/kg b.w. was delivered twice a day, on a piece of paper taped to the wall of the rat's home cage, approximately at rats' nose level while standing. Cage mates were temporarily separated for 10 min during the feeding, with the dosed rat's cage mate being placed temporarily in an empty Type I cage.

The NSAID indomethacin was used as a pharmacologic standard in the positive control group: it was administered by intragastric gavage to arthritic rats once a day on days 19–25. The formulation was prepared fresh every day. Indomethacin (0.1 mg/ml; prod. no. I7378; Sigma-Aldrich) was dissolved in an aqueous solution of 0.5% (w/v) methyl cellulose (prod. no. M0262; Sigma-Aldrich). Animals were dosed using a 5-cm curved stainless steel gavage needle with a ball tip, delivering intragastric bolus of 10 ml/kg rat body weight, containing 1 mg/kg indomethacin.

Each experimental group consisted of 8 females. Animals were randomized in blocks based on caging, such that two animals from the same group would never share the same cage – thus attempting to mitigate ambient effects on the model. All experimental procedures were carried out in a randomized order.

Clinical assessments (blinded and informed approach). The animals' disease progression was monitored throughout the study. The animals were weighed on days 1 (pre challenge), 7, 10, 15, 18, 23, and 26. On days 0 and 19 (just before the start of indomethacin therapy) and on day 26 (pre sacrifice) the animals' hind paws were photographed. Two standardized digital photographs showing the bottom/plantar and lateral views of both hind paws were used for clinical assessments by four independent assessors who did not participate in the hands-on part of the study. Assessors viewed anonymised images in a randomised order. After blinded assessments, assessors were informed about the therapy regimens and photos from the same animals were shown as clearly numbered sets that could easily be linked to the treatment received. For both assessments, the arthritic joints were scored using the methods of Hawkins *et al.* (7).

Biomarkers. Blood samples were collected on days 1 (pre challenge), 11, 19, and 26 in Croatia and on days 1 (pre challenge), 8, 15, 19, and 26 in Israel. The blood sampling was facilitated either by use of a warming cabinet in Croatia (cage mates were warmed for 10 min in 39°C inside VetTech warming chambers) or by use of a red electrical bulb in Israel (cage mates were warmed for 10 min under a bench top red lamp). A blood sample of approximately 250 µl was obtained from the tail vein. Serum was separated and stored at –80°C until analysis. Samples were analyzed for α_1 -acid glycoprotein content (α_1 AGP, µg/ml) using commercially available ELISA kits (Rat α_1 -AGP; MyBioSource, San Diego, CA, USA).

Histopathology. On day 26, the rats were euthanized by exsanguination under anesthesia delivered *i.p.* comprising 100 mg/kg ketamine (5%

solution for injection, Narkamon, Bioveta, Ivanovice na Hané, Czech Republic) and 20 mg/kg xylazine (2% solution for injection, Alfasan, JA Woerden, the Netherlands). Both hind paws, severed at the level of the ankle, were submerged in 10% neutral buffered formalin (Thermo Shandon Ltd., subsidiary of Thermo Fisher Scientific, Cheshire, UK). After 10 days at room temperature, the paws were decalcified in 10% formic acid (Honeywell Fluka, Charlotte, NC, USA in Israel and Kemika, Zagreb, Croatia in Croatia) in 10% formalin solution (Thermo Shandon Ltd.) for another 10 days. Subsequently, the joints were cut, mid-longitudinally, through the center, and embedded in paraffin. Microtome sections of 1 µm were prepared and stained using hematoxylin-eosin. The slides were assessed by two pathologists, independently, first in blinded and then in informed fashion. A combined score (Table I), grading inflammation, bone erosion, and osteoclast activity, was employed (33, 34).

Statistics. We confined statistical analysis of the model parameters to the final day of the experiments, since this was the only time point where all parameters could be expected to have been affected by the analgesic treatments. All of the measurements were analyzed using two-way analysis of variance, using indomethacin and buprenorphine treatments as independent variables. Analyses were conducted separately for the two facilities since the levels of many of the parameters were greatly different at baseline. We contrasted blinded and informed scoring using paired *t*-tests. Here, we combined data from the two facilities, given that analysis was carried out at the level of observation. Finally, for correlations, we calculated Pearson's correlation coefficient and evaluated model fit by calculating R^2 from a linear regression. We considered $p < 0.05$ to be statistically significant, and the variability of model parameters is presented as the standard deviation.

Results

Across both facilities, the analgesic treatments had minimal effects on the measured model parameters (Figure 1). We found no effect on any of the parameters from the buprenorphine treatment. Treatment with indomethacin showed some changes suggestive of improved welfare. In Croatia, the indomethacin-treated animals regained more body weight. All animals lost weight following model induction, but, on average, the indomethacin-treated animals regained 6.4 g [± 3.2 g (SEM)] in the treatment period ($F_{1,28}=5.5$; $p=0.026$). In Israel, the clinical score – the number of swollen joints – was reduced slightly by the indomethacin treatment. On average, the treatment reduced the change in the clinical scores from Day 19 to Day 26 by 4.5 (± 1.9) points ($F_{1,26}=6.3$; $p=0.019$). Importantly, these changes did not appear to manifest as differences in the histological analyses (Figure 1).

The histological scores were consistent between the blinded and informed assessments (Figure 2). Minor differences at the level of the individual slides could be observed, but once added up (inflammation, erosion, osteoclasts) to a score, combining these scores for both legs, and then averaging the results resulted in small effects being averaged out. Consequently, we observed no difference between the blinded and informed scoring for the histological assessments.

Table I. Histopathological scoring of arthritic severity. The scores have been developed by Bolon et al. (33) and Coxon et al. (34).

Score	Inflammation
0	None. Normal tissue.
1	Few inflammatory cells.
2	Mild inflammation.
3	Moderate inflammation (often, but not always, diffuse)
4	Marked inflammation (diffuse and dense, with large periarticular abscesses).
Score	Bone erosion
0	None. Normal tissue.
1	Minimal loss of cortical or trabecular bone at a few sites.
2	Mild loss of cortical or trabecular bone at modest number of sites (generally tarsals).
3	Moderate loss of bone at many sites. Usually the trabeculae of the tarsals, but sometimes the cortex of the distal tibia.
4	Marked loss of bone at many sites. Usually as extensive destruction of trabeculae in the tarsals, but sometimes with partial loss of cortical bone in the distal tibia.
5	Marked loss of bone at many sites. With fragmenting of tarsal trabeculae and full-thickness penetration of cortical bone in the distal tibia.
Score	Osteoclasts
0	No osteoclasts. Normal tissue.
1	Few osteoclasts (lining less than 5% of the most affected bone surfaces).
2	Some osteoclasts (lining 5-25% of the most affected bone surfaces).
3	Many osteoclasts (lining 25-50% of the most affected bone surfaces).
4	Myriad osteoclasts (lining more than 50% of the most affected bone surfaces).

For the clinical assessments, a difference could be seen between when the assessors were informed of what they were scoring and when not (Figure 2). We found an improvement (a negative change in scores) in nearly all animals when scored informed, whereas it was essentially 50:50 for the blinded scorings. As a result, there was a significant lowering of the difference in the clinical scores over the treatment period across all animals (paired-samples *t*-test: $t_{61}=4.1, p<0.001$). When the assessors knew they were looking at photos from before a treatment was administered, they were more inclined to score the full 24 inflamed joints, particularly for the no treatment group (Figure 3).

For data generated in Israel, the blinded assessments had already found an effect of the indomethacin treatment. The effect was pronounced in the informed scorings, as the scores before the treatments were, on average, higher. Informed scoring did not affect the overall conclusions; it did however, also lower the scores for the buprenorphine-treated animals. We have no reason to believe that buprenorphine should affect the number of inflamed joints. Yet, informed scoring significantly lowered the clinical scores when compared to blinded scoring ($t_{15}=2.4, p=0.03$) by 3.8 (± 1.6) points. With a larger sample size, this difference might have manifested as a difference also in the overall hypothesis test.

The blinded and informed clinical scores on day 26 (combining data from both facilities, Figure 3) were not well-correlated ($r=0.51, p<0.001$) with the exception of some clustering in the lower left corner. The linear relation

explained only a fraction of the variation in the data ($R^2=0.28$) suggesting that the task of scoring the number of inflamed joints was difficult and highly subjective.

Discussion

Both unrelieved pain and the use of analgesia in animal experimentation may act as significant confounders, potentially skewing results as well as increasing inter- and intra-animal variation (8, 35, 36). Buprenorphine, is the most commonly used opioid analgesic for laboratory animals. Buprenorphine did not affect the measured model parameters, which agrees well with its weak immunosuppressive properties. Indomethacin, by contrast, reduces inflammation and may potentially alter the model. In the present investigation, we did not find it affecting model pathology. Instead, we observed less body weight decrease in the animals in Croatia and improvement in clinical scores in the rats in Israel treated with indomethacin.

In rats, $\alpha 1$ AGP is the dominant acute phase protein whereas in humans, C-reactive protein is more prominent (37). In most disease states, including inflammation, $\alpha 1$ AGP is detectable for several days after the stimulus (38). Under chronic inflammation conditions (e.g., arthritis) increased serum concentrations of acute phase proteins are generally observed but the increase is lower than during acute episodes of inflammation or infection. Under experimental arthritis conditions in rats, AGP levels have been reported to increase

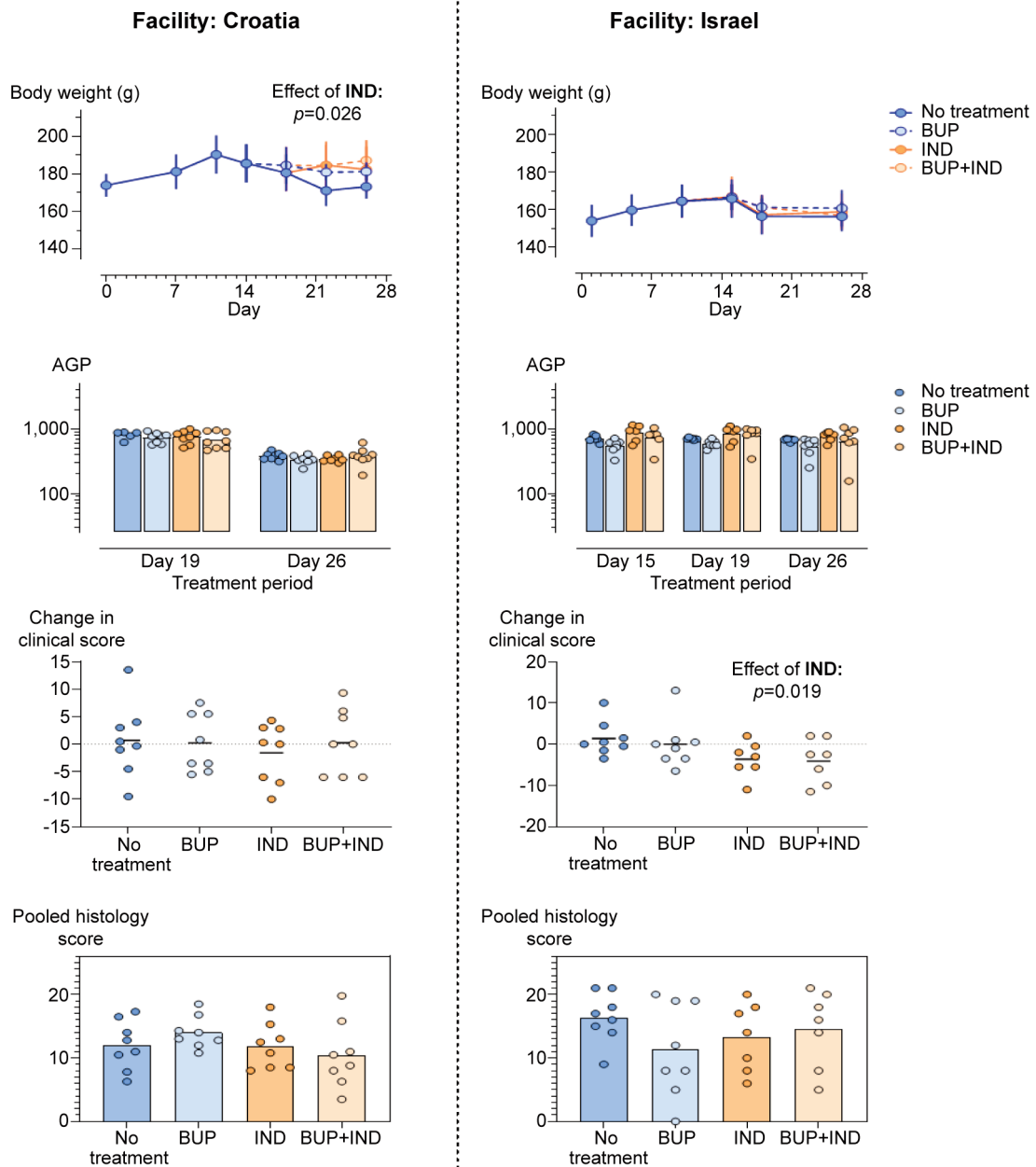


Figure 1. Model parameters measured in female rats across two facilities (in Croatia on the left and in Israel on the right). Error bars for the body weights represent standard deviations. Clinical scores are shown as the difference in the scores (number of inflamed joints) before/after treatment. For both the clinical scores and histology scores, the results of the blinded assessments are shown. Bars represent the median.

10-fold immediately before the onset of arthritis after which it steadily declines towards the chronic phase of the disease (39). Quantitative analysis of AGP in this investigation suggests that this response was not affected by the analgesic treatments under investigation.

Animal research often relies on visual observations, methods, and scales that quantify a physical or a clinical condition, all of

which are to some extent subjective. An improper skew in these observations, referred to as observer bias, is particularly likely when the investigator has preconceptions and expectations and when the underlying data are ambiguous (29). In the present study, there was a clear difference between blinded and informed clinical scores. The strongest effect seemed to stem from knowing whether the photo was taken before or after

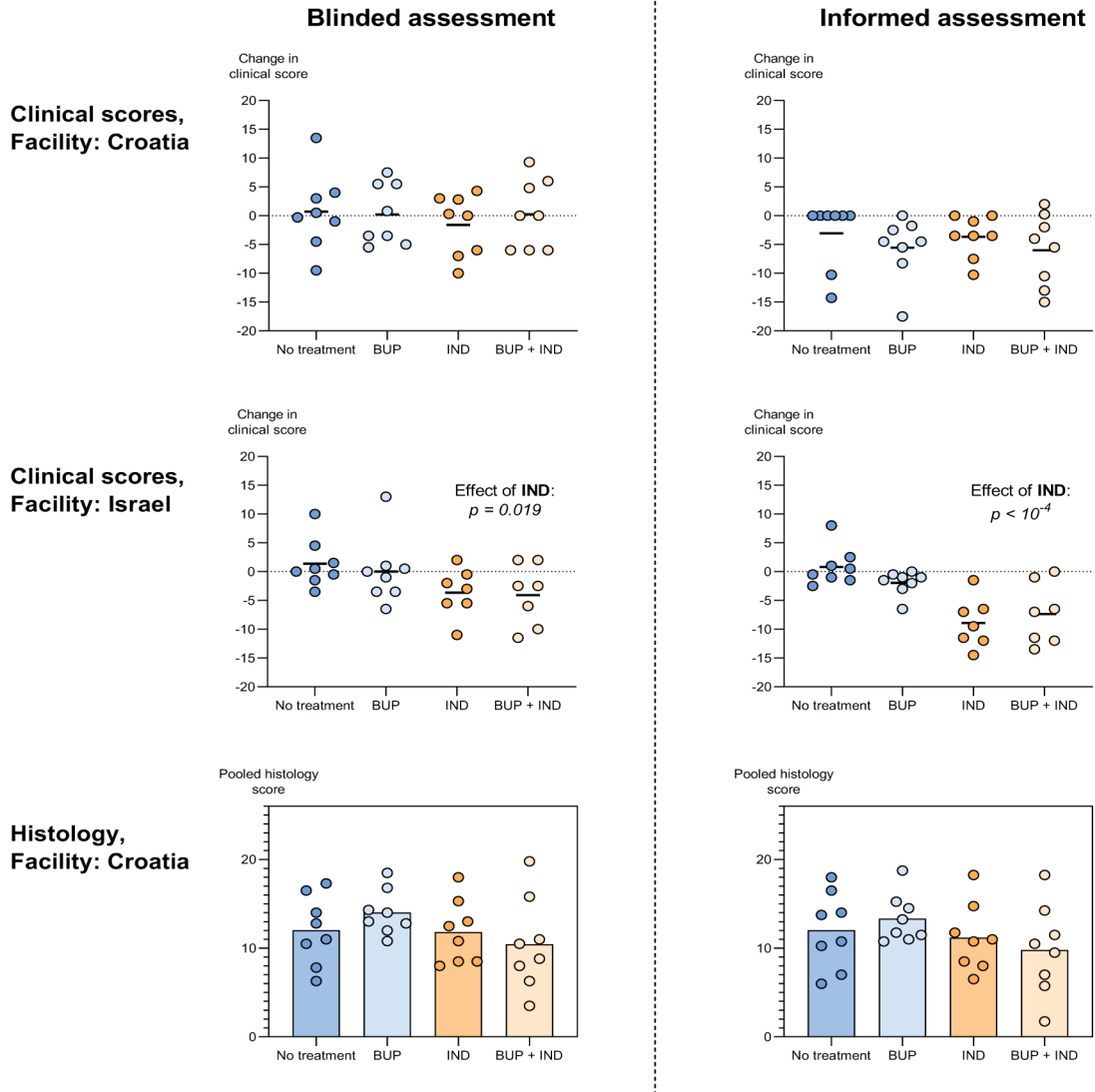


Figure 2. Blinded and informed scorings of paw swelling and histology. Clinical scores (change in number of swollen joints comparing before and after the treatment period) show much clearer effects when scored informed (right hand side). A similar effect was not found for the histological scores. Bars represent the median.

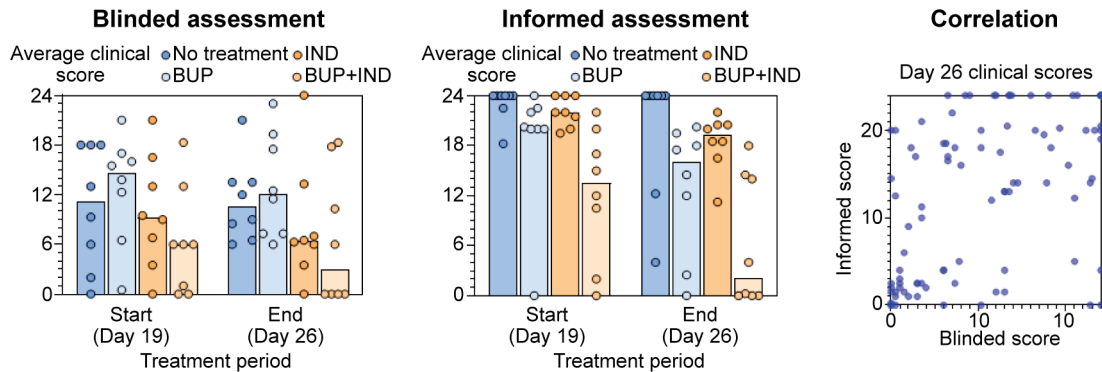


Figure 3. Blinded and informed clinical assessments performed at the facility in Croatia, and the correlation between blinded and informed scores (for day 26) across both facilities. Bars represent the median.

model induction/treatments. In the scores from Croatia, the assessors observed an anti-inflammatory effect of BUP that did not exist, which demonstrates the influence of assessors' expectation on their scoring decisions.

By contrast, we found no effect of blinding on histology scores. This may be due to the routine histology-scoring framework being more rigid, leaving less room for subjectivity or perhaps owing to the histologists being sensitive to being monitored.

Despite rigorous standardization of experimental conditions across laboratories, several multi-laboratory studies have revealed large proportions of results that were laboratory specific (40-43). Many environmental factors (*e.g.*, personnel, equipment used, husbandry procedures, environmental conditions, research procedures) can never be perfectly standardised between laboratories (22, 41). In the present study, we carried out similar protocols in Croatia and Israel. We noted certain differences in experimental results. However, the AIA model was reasonably robust, which is positive given that it is a widely used model.

Conclusion

The combined indomethacin and buprenorphine therapy did not significantly impact the validity of the AIA model. The present results support the use of more effective analgesic treatment regimens – treatments more in line with that of human patients – ensuring better animal welfare all the while improving the quality of the associated studies. Comparing blinded assessments and informed assessments confirms the best *in vivo* preclinical practice recommendations advocating blinding as a mandatory practice in conduct of preclinical *ex vivo* efficacy studies.

Conflicts of Interest

The Authors declare that they have no competing interests in relation to this study.

Authors' Contributions

Dasa Seveljevic-Jaran confirms contribution to the study conception, design, data collection, and analysis. All Authors confirm contribution to the interpretation of results and draft article preparation. Finally, all Authors reviewed the results and approved the final version of the article.

Acknowledgements

The Authors would like to thank Rony Kalman (Jerusalem, Israel) for offering the Sharett animal facility at Hebrew University (HU) as an alternative facility for conduct of PhD research and for financial support. The Authors are grateful to HU managers, research and animal care taking personnel for providing the necessary technical assistance during research conduct at HU. The

Authors also extend their thanks to Selvita's managers, the peers and the technicians for providing much appreciated financial and technical assistance for all *in vivo* and *ex vivo* research activities.

References

- van Vollenhoven RF: Sex differences in rheumatoid arthritis: more than meets the eye. *BMC Med* 7: 12, 2009. PMID: 19331649. DOI: 10.1186/1741-7015-7-12
- Sardar S and Andersson Å: Old and new therapeutics for Rheumatoid Arthritis: *in vivo* models and drug development. *Immunopharmacol Immunotoxicol* 38(1): 2-13, 2016. PMID: 26769136. DOI: 10.3109/08923973.2015.1125917
- Fischer BD, Adeyemo A, O'Leary ME and Bottaro A: Animal models of rheumatoid pain: experimental systems and insights. *Arthritis Res Ther* 19(1): 146, 2017. PMID: 28666464. DOI: 10.1186/s13075-017-1361-6
- Carbone L and Austin J: Pain and laboratory animals: publication practices for better data reproducibility and better animal welfare. *PLoS One* 11(5): e0155001, 2016. PMID: 27171143. DOI: 10.1371/journal.pone.0155001
- Peterson NC, Nunamaker EA and Turner PV: To treat or not to treat: the effects of pain on experimental parameters. *Comp Med* 67(6): 469-482, 2017. PMID: 29212578.
- Walker JS, Chandler AK, Wilson JL, Binder W and Day RO: Effect of mu-opioids morphine and buprenorphine on the development of adjuvant arthritis in rats. *Inflamm Res* 45(11): 557-563, 1996. PMID: 8951507. DOI: 10.1007/BF02342227
- Hawkins P, Armstrong R, Boden T, Garside P, Knight K, Lilley E, Seed M, Wilkinson M and Williams RO: Applying refinement to the use of mice and rats in rheumatoid arthritis research. *Inflammopharmacology* 23(4): 131-150, 2015. PMID: 26168847. DOI: 10.1007/s10787-015-0241-4
- Flecknell P: Rodent analgesia: Assessment and therapeutics. *Vet J* 232: 70-77, 2018. PMID: 29428096. DOI: 10.1016/j.tvjl.2017.12.017
- Foley PL, Kendall LV and Turner PV: Clinical management of pain in rodents. *Comp Med* 69(6): 468-489, 2019. PMID: 31822323. DOI: 10.30802/AALAS-CM-19-000048
- Prinz F, Schlange T and Asadullah K: Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10(9): 712, 2011. PMID: 21892149. DOI: 10.1038/nrd3439-c1
- Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubitza AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetz W, Krainc D, Lazic SE, Levine MS, Macleod MR, McCall JM, Moxley RT 3rd, Narasimhan K, Noble LJ, Perrin S, Porter JD, Steward O, Unger E, Utz U and Silberberg SD: A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490(7419): 187-191, 2012. PMID: 23060188. DOI: 10.1038/nature11556
- Begley CG and Ellis LM: Drug development: Raise standards for preclinical cancer research. *Nature* 483(7391): 531-533, 2012. PMID: 22460880. DOI: 10.1038/483531a
- Begley CG and Ioannidis JP: Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* 116(1): 116-126, 2015. PMID: 25552691. DOI: 10.1161/CIRCRESAHA.114.303819

- 14 Freedman LP and Gibson MC: The impact of preclinical irreproducibility on drug development. *Clin Pharmacol Ther* 97(1): 16-18, 2015. PMID: 25670378. DOI: 10.1002/cpt.9
- 15 Freedman LP, Cockburn IM and Simcoe TS: The economics of reproducibility in preclinical research. *PLoS Biol* 13(6): e1002165, 2015. PMID: 26057340. DOI: 10.1371/journal.pbio.1002165
- 16 Macleod M and Mohan S: Reproducibility and rigor in animal-based research. *ILAR J* 60(1): 17-23, 2019. PMID: 31687758. DOI: 10.1093/ilar/ilz015
- 17 Henderson VC, Kimmelman J, Fergusson D, Grimshaw JM and Hackam DG: Threats to validity in the design and conduct of preclinical efficacy studies: a systematic review of guidelines for *in vivo* animal experiments. *PLoS Med* 10(7): e1001489, 2013. PMID: 23935460. DOI: 10.1371/journal.pmed.1001489
- 18 Reichlin TS, Vogt L and Würbel H: The researchers' view of scientific rigor-survey on the conduct and reporting of *in vivo* research. *PLoS One* 11(12): e0165999, 2016. PMID: 27911901. DOI: 10.1371/journal.pone.0165999
- 19 Vogt L, Reichlin TS, Nathues C and Würbel H: Authorization of animal experiments is based on confidence rather than evidence of scientific rigor. *PLoS Biol* 14(12): e2000598, 2016. PMID: 27911892. DOI: 10.1371/journal.pbio.2000598
- 20 Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M and Langendam MW: SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol* 14: 43, 2014. PMID: 24667063. DOI: 10.1186/1471-2288-14-43
- 21 Cheleuitte-Nieves C and Lipman NS: Improving replicability, reproducibility, and reliability in preclinical research: a shared responsibility. *ILAR J* 60(2): 113-119, 2019. DOI: 10.1093/ilar/ilaa009
- 22 Bailoo JD, Reichlin TS and Würbel H: Refinement of experimental design and conduct in laboratory animal research. *ILAR J* 55(3): 383-391, 2014. PMID: 25541540. DOI: 10.1093/ilar/ilu037
- 23 Osborne N, Avey MT, Anestidou L, Ritskes-Hoitinga M and Griffin G: Improving animal research reporting standards: HARRP, the first step of a unified approach by ICLAS to improve animal research reporting standards worldwide. *EMBO Rep* 19(5): e46069, 2018. PMID: 29669797. DOI: 10.15252/embr.201846069
- 24 Bepalov A, Wicke K and Castagné V: Blinding and Randomization. In: *Good Research Practice in Non-Clinical Pharmacology and Biomedicine*. Bepalov A, Michel MC and Steckler T (eds.). Cham, Switzerland, Springer, pp. 81, 2020. DOI: 10.1007/978-3-030-33656-1
- 25 Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaud P and Brorson S: Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 344: e1119, 2012. PMID: 22371859. DOI: 10.1136/bmj.e1119
- 26 Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, Ravaud P and Brorson S: Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ* 185(4): E201-E211, 2013. PMID: 23359047. DOI: 10.1503/cmaj.120744
- 27 Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, Boutron I, Ravaud P and Brorson S: Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *Int J Epidemiol* 43(3): 937-948, 2014. PMID: 24448109. DOI: 10.1093/ije/dyt270
- 28 Bello S, Krogsbøll LT, Gruber J, Zhao ZJ, Fischer D and Hróbjartsson A: Lack of blinding of outcome assessors in animal model experiments implies risk of observer bias. *J Clin Epidemiol* 67(9): 973-983, 2014. PMID: 24972762. DOI: 10.1016/j.jclinepi.2014.04.008
- 29 Tuytens FAM, de Graaf S, Heerkens JLT, Jacobs L, Nalon E, Ott S, Stadig L, Van Laer E and Ampe B: Observer bias in animal behaviour research: can we believe what we score, if we score what we believe? *Anim Behav* 90: 273-280, 2014. DOI: 10.1016/j.anbehav.2014.02.007
- 30 de Vries RB, Wever KE, Avey MT, Stephens ML, Sena ES and Leenaars M: The usefulness of systematic reviews of animal experiments for the design of preclinical and clinical studies. *ILAR J* 55(3): 427-437, 2014. PMID: 25541545. DOI: 10.1093/ilar/ilu043
- 31 Kahan BC, Rehal S and Cro S: Blinded outcome assessment was infrequently used and poorly reported in open trials. *PLoS One* 10(6): e0131926, 2015. PMID: 26120839. DOI: 10.1371/journal.pone.0131926
- 32 Abelson KS, Jacobsen KR, Sundbom R, Kalliokoski O and Hau J: Voluntary ingestion of nut paste for administration of buprenorphine in rats and mice. *Lab Anim* 46(4): 349-351, 2012. PMID: 22969145. DOI: 10.1258/la.2012.012028
- 33 Bolon B, Campagnuolo G and Feige U: Duration of bone protection by a single osteoprotegerin injection in rats with adjuvant-induced arthritis. *Cell Mol Life Sci* 59(9): 1569-1576, 2002. PMID: 12440777. DOI: 10.1007/s00018-002-8530-7
- 34 Coxon A, Bolon B, Estrada J, Kaufman S, Scully S, Rattan A, Duryea D, Hu YL, Rex K, Pacheco E, Van G, Zack D and Feige U: Inhibition of interleukin-1 but not tumor necrosis factor suppresses neovascularization in rat models of corneal angiogenesis and adjuvant arthritis. *Arthritis Rheum* 46(10): 2604-2612, 2002. PMID: 12384918. DOI: 10.1002/art.10546
- 35 Jirkof P: Side effects of pain and analgesia in animal experimentation. *Lab Anim (NY)* 46(4): 123-128, 2017. PMID: 28328895. DOI: 10.1038/labana.1216
- 36 Herrmann K and Flecknell P: Retrospective review of anesthetic and analgesic regimens used in animal research proposals. *ALTEX* 36(1): 65-80, 2019. PMID: 30222179. DOI: 10.14573/altex.1804011
- 37 Petersen HH, Nielsen JP and Heegaard PM: Application of acute phase protein measurements in veterinary clinical chemistry. *Vet Res* 35(2): 163-187, 2004. PMID: 15099494. DOI: 10.1051/vetres:2004002
- 38 Stolina M, Bolon B, Middleton S, Dwyer D, Brown H, Duryea D, Zhu L, Rohner A, Pretorius J, Kostenuik P, Feige U and Zack D: The evolving systemic and local biomarker milieu at different stages of disease progression in rat adjuvant-induced arthritis. *J Clin Immunol* 29(2): 158-174, 2009. PMID: 18726678. DOI: 10.1007/s10875-008-9238-8
- 39 Smith SA and Waters NJ: Pharmacokinetic and pharmacodynamic considerations for drugs binding to Alpha-1-Acid glycoprotein. *Pharm Res* 36(2): 30, 2018. PMID: 30593605. DOI: 10.1007/s11095-018-2551-x
- 40 Vierron E and Giraudeau B: Design effect in multicenter studies: gain or loss of power? *BMC Med Res Methodol* 9: 39, 2009. PMID: 19538744. DOI: 10.1186/1471-2288-9-39

- 41 Richter SH, Garner JP and Würbel H: Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nat Methods* 6(4): 257-261, 2009. PMID: 19333241. DOI: 10.1038/nmeth.1312
- 42 Richter SH, Garner JP, Zipser B, Lewejohann L, Sachser N, Touma C, Schindler B, Chourbaji S, Brandwein C, Gass P, van Stipdonk N, van der Harst J, Spruijt B, Vöikar V, Wolfer DP and Würbel H: Effect of population heterogenization on the reproducibility of mouse behavior: a multi-laboratory study. *PLoS One* 6(1): e16461, 2011. PMID: 21305027. DOI: 10.1371/journal.pone.0016461
- 43 Voelkl B, Vogt L, Sena ES and Würbel H: Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biol* 16(2): e2003693, 2018. PMID: 29470495. DOI: 10.1371/journal.pbio.2003693

Received November 8, 2022
Revised November 23, 2022
Accepted November 24, 2022