Cross-centre replication of suppressed burrowing behaviour as an ethologically relevant pain outcome measure in the rat: a prospective multicentre study

Rachel Wodarski	extsuperscript{a,b}, Ada Delaney	extsuperscript{c}, Camilla Ultenius	extsuperscript{d}, Rosie Morland	extsuperscript{a}, Nick Andrews	extsuperscript{b}, Catherine Baastrup	extsuperscript{1}, Luke A. Bryden	extsuperscript{b}, Ombretta Caspani	extsuperscript{h}, Thomas Christoph	extsuperscript{i}, Natalie J. Gardiner	extsuperscript{i}, Wenlong Huang	extsuperscript{a}, Jeffrey D. Kennedy	extsuperscript{b}, Suguru Koyama	extsuperscript{l}, Dominic Li	extsuperscript{k}, Marcin Ligocki	extsuperscript{l}, Annika Lindsten	extsuperscript{m}, Ian Machin	extsuperscript{n}, Anton Pekcec	extsuperscript{o}, Angela Robens	extsuperscript{i}, Sanziana M. Rotariu	extsuperscript{i}, Sabrina Voß	extsuperscript{b}, Marta Segerdah	extsuperscript{d,m}, Carina Stenfors	extsuperscript{d}, Camilla I. Svensson	extsuperscript{e}, Rolf-Detlef Treede	extsuperscript{k}, Katsuhiro Uto	extsuperscript{i}, Kazumi Yamamoto	extsuperscript{i}, Kris Rutten	extsuperscript{a,*}, Andrew S.C. Rice	extsuperscript{a}

Abstract

Burrowing, an ethologically relevant rodent behaviour, has been proposed as a novel outcome measure to assess the global impact of pain in rats. In a prospective multicentre study using male rats (Wistar, Sprague-Dawley), replication of suppressed burrowing behaviour in the complete Freund adjuvant (CFA)-induced model of inflammatory pain (unilateral, 1 mg/mL in 100 µL) was evaluated in 11 studies across 8 centres. Following a standard protocol, data from participating centres were collected centrally and analysed with a restricted maximum likelihood-based mixed model for repeated measures. The total population (TP—all animals allocated to treatment; n = 249) and a selected population (SP—TP animals burrowing over 500 g at baseline; n = 200) were analysed separately, assessing the effect of excluding “poor” burrowers. Mean baseline burrowing across studies was 1113 g (95% confidence interval: 1041-1185 g) for TP and 1329 g (1271-1387 g) for SP. Burrowing was significantly suppressed in the majority of studies 24 hours (7 studies/population) and 48 hours (7 TP, 6 SP) after CFA injections. Across all centres, significantly suppressed burrowing peaked 24 hours after CFA injections, with a burrowing deficit of −374 g (−479 to −269 g) for TP and −498 g (−609 to −386 g) for SP. This unique multicentre approach first provided high-quality evidence evaluating suppressed burrowing as robust and reproducible, supporting its use as a tool to infer the global effect of pain on rodents. Second, our approach provided important informative value for the use of multicentre studies in the future.

Keywords: Non-evoked pain, Validation, Reproducibility, Preclinical controlled trials

1. Introduction

Prospective multicentre controlled trials are an important clinical research tool, providing high-quality evidence to inform health care systems about the validity of new treatments and outcome measures. In contrast, preclinical studies are predominantly single-centre studies conducted using experimental protocols varying significantly across laboratories. This likely contributes to the well-recognized poor level of experimental reproducibility. Implementing the concept, ethos, and design of multicentre clinical trials into preclinical studies could significantly increase reproducibility and translatability by standardising experimental design, monitoring data, and improving reporting standards.

The use of multicentre designs in preclinical studies is rare; however, a few pioneering studies in mice have demonstrated the effects of environmental conditions, strain, and study design on behavioural outcomes across centres. A multicentre approach
for preclinical pain studies could equally provide robust validation and evidence for reproducibility of models, outcome measures, and pharmacological interventions. This could be achieved by systematically assessing replication and interlaboratory variability and identifying factors that may be associated with such variability.

A major shortcoming in chronic pain research has been the lack of newly developed analytics, because translation from promising preclinical studies to successful clinic trials has been poor. Animal studies mostly rely on stimulus-evoked measures of sensory gain, whereas clinical trials use patient-reported outcomes focusing on spontaneous pain intensity, functioning, and quality of life. Although spontaneous pain and decreased quality of life are reported across chronic pain conditions, sensory gain is common only in a subset of clinical pain conditions caused by peripheral nerve trauma and inflammation but is rare in others, such as polyneuropathies caused by diabetes and HIV. This has brought into question the translational strength of relying on sensory gain as the sole pain outcome measure preclinically. We have previously proposed ethologically relevant rodent behaviours, such as predator avoidance behaviours, as novel nonevoked pain outcome measures to assess the global impact of pain on animals.

Among these ethologically relevant behaviours is burrowing. Burrowing, a social behaviour of rats that is important for building underground habitats and nests, is highly conserved in laboratory rats. Suppressed burrowing, quantified by a reduced amount of displaced substrate from an artificial burrow, is indicative of behavioural dysfunction. Pain-depressed behaviours, such as reduction in feeding, locomotion, or operant behaviours, have been proposed as pain-relevant measures that assess functioning instead of sensory changes. Although burrowing is a relatively novel pain outcome measure, a few studies have reported decreased burrowing in rat models of inflammatory and neuropathic pain; importantly, burrowing deficits were reversed by known analgesics, suggesting a degree of predictive validity for this outcome measure.

We hypothesise that by using a prospective multicentre design, the reliability of novel outcome measures can be efficiently and rapidly evaluated. As part of the IMI Europain collaboration (http://www.imieuppain.org), we investigated rat burrowing behaviour after induction of complete Freund adjuvant (CFA)-associated inflammation across 8 centres to assess the reliability of suppressed burrowing as a pain-related outcome measure. We hope this will facilitate future routine use of multicentre studies to evaluate novel outcome measures and models.

2. Materials and methods

The purpose of this study was to identify whether the impact of CFA-induced inflammation on burrowing behaviour can be reproduced across multiple centres when following a basic protocol (Supplemental digital content 1, http://links.lww.com/PAIN/A302). For this, a draft protocol was written and then reviewed by the participating centres before it was finalised. Minor changes to the protocol were permitted to accommodate for local variations in laboratory practice and procedures (eg, different animal suppliers, variations in environmental conditions, substrate and equipment availability, and precise randomisation procedures).

2.1. Ethical statement

All animal experiments conformed to local Government and Institutional guidelines on the care and use of animals in research and the IASP guidelines for in vivo research. Guidance was given on Good Laboratory Practice standards, but exact methods were not specified, and local variations were allowed and recorded (Table 1). Experiments are reported in accordance with the ARRIVE Guidelines (Supplemental digital content 6, http://links.lww.com/PAIN/A307).

2.2. Experimental animals and environmental conditions

Adult male rats were used for all experiments. Animals were group-housed in standard cages in temperature and humidity-controlled conditions (Table 1). Bedding and environmental enrichment was provided according to local convention, including nesting material, jolly balls, or cardboard/plastic tubes. Animals were provided with standard rat chow pellets and tap water ad libitum and were allowed to acclimatise to their housing environment for a minimum of 4 to 7 days before experiments started. Animals were monitored during regular husbandry activity to ensure their well-being.

2.3. Complete Freund's adjuvant—induced inflammation

Inflammation was induced under isoflurane/O₂ anaesthesia by unilateral intraplantar injection of CFA. Sham animals received a saline or incomplete CFA injection of the same volume. The CFA model was chosen because of the high likelihood of ongoing spontaneous pain, as reported in humans after an accidental CFA injection and in animals.

2.4. Burrowing

Burrowing experiments were performed as previously described. For this, a burrowing tube (320 × 100 mm; open end elevated by 60 mm), made of either steel (Boehringer Ingelheim) or plastic (all other groups), was filled with 2500 g substrate and placed in an empty cage. No floor bedding was provided during the burrowing task, as this could hinder cleaning of the substrate. In some studies, however, the cage floor was covered with paper towels to create a more comfortable environment for the animals. If multiple animal cohorts were tested per day, test cages were emptied of displaced substrate, faecal bolus and tissue paper, and were wiped clean or replaced with a clean testing cage before starting the next session. As free access to food and water could distract animals from burrowing, most studies opted not to provide food and water during testing. At studies performed at Eli Lilly (United Kingdom), water was accessible during testing, as local regulations only allowed a maximal period of 2 hours without water. Details of the experimental set-ups can be found in Table 2.

2.4.1. Training—social facilitation

For animals to learn the task, 2 to 3 training sessions were conducted over consecutive days (apart from weekends). After acclimatisation to the testing room and/or experimental set-up without a burrowing tube, a filled burrowing tube was placed in the test cage, and animals, in pairs, were allowed to burrow for 60 to 120 minutes. If a pair showed poor burrowing behaviour during the first training session, one of the animals was swapped with an animal from a known burrowing pair to facilitate burrowing in subsequent sessions. As no criterion was set a priori to define a poor burrowing pair, experimenters determined this on a case-by-case basis. At Boehringer Ingelheim, pairs burrowing less than 1500 g were classified as poor burrowers; because strong burrowers would displace nearly all of the 2500 g substrate, this high limit was set to ensure the behaviour was strongly expressed.
### Table 1

**Major domains of Good Laboratory Practice.**

| Description of the procedure | Boehringer Ingelheim Pharma GmbH | Grünenthal GmbH | Heidelberg University | Asahi Kasei Pharma | University of Manchester | Eli Lilly and Company (United Kingdom) | Imperial College London | Eli Lilly and Company (USA) |
|------------------------------|---------------------------------|-----------------|----------------------|---------------------|-------------------------|----------------------------------------|-------------------------|---------------------------|
| Housing environment/ experimental lighting conditions | 20-22°C; 45%-60% humidity; 12/12 h light/dark cycle/ lights 15 lux | 19-23°C; 15%-50% humidity; 12/12 h light/dark cycle/ dimmed lights 0.8-7.6 lux | 24-26°C; 49%-66% humidity; 12/12 h light/dark cycle/ dimmed lights (<50 lux) | 21-21°C; 50% humidity; 12/12 h light/dark cycle/ bright lights (study 1 and 2); dimmed lights (study 3) | 30-60 lux | 30-60 lux | 30-60 lux |
| Sample size estimation (SigmaPlot 11.0 [analysis of variance]) was calculated with pilot burrowing data at 3 d after induction of inflammation (n = 6 and n = 5 for control and complete Freund adjuvant groups, respectively). Parameters were difference in means = 951.79 g, SD = 549.29 g, α = 0.05, and power = 0.8. An adequately powered study required an n = 7 per group and n = 8 when testing 2 and 3 groups, respectively. N numbers for each study are reported in the results and in Figure 2.

**Inclusion/exclusion criteria**

In the basic protocol, animals that burrowed more or less than 2 SD of the total mean at baseline sessions should have been excluded before model induction. However, because of high variability, this never occurred. Alternative criteria were discussed including the exclusion of animals that burrowed less than 500 g at baseline. Thus, at Eli Lilly (United Kingdom), 5 of 32 animals in study 1 and at Grünenthal 1 of 22 animals were excluded before model induction as they burrowed less than 500 g during baseline measurements. No data from these animals was included. At Boehringer Ingelheim, rats were excluded when burrowing on average <1000 g or had an SD >450 g after 3 individual training days; no animals had to be excluded according to this criterion. For all other studies investigators opted to not exclude animals before model induction but instead analyse burrowing behaviour with a total population and selected population approach as described below. Exclusions are reported in Figure 2.

**Randomisation to groups: naive, sham, CFA or sham, CFA**

Animals were randomly allocated to groups by using a random number generator in Excel. Baseline values (mean 3 baseline sessions) were then used to shift individual values, in a blinded fashion, from group to group until each group had a comparable baseline in terms of mean value and variation.

Baseline values were used to rank animals from low to high performers and pseudorandomised by allocating them to groups following a repeated sequence. Individual values were shifted from group to group until each had a comparable baseline in terms of mean value and variation.

Last baseline session values were used to randomly allocate animals to groups using StatLight#11 (Yukms Co, Ltd).

Baseline values were used to rank animals from low to high performers before randomly allocating animals to groups by using a random sequence generator in Excel. Individual values were shifted blinded from group to group until each group had a comparable baseline in terms of mean value and variation.

Baseline values were used to rank animals from low to high performers before randomly allocating animals to groups by using a random sequence generator in Excel. Individual values were shifted blinded from group to group until each group had a comparable baseline in terms of mean value and variation.

Baseline values were used to randomly allocate animals to groups using a random sequence generator in Excel.

Allocation concealment

No allocation concealment procedure was followed.

No allocation concealment procedure was followed because only CFA animals were injected because of the control group being naive animals.

Syringes for model induction were prepared by an independent investigator. The investigator performing model induction was only aware of animal ID.

Syringes for model induction were prepared by the investigator only aware of the letter coded group name. During model induction, the investigator was not aware of the treatment as animals and syringes were picked by an independent researcher outside the experimental room.

Syringes for model induction were prepared by the investigator only aware of the letter coded group name. During model induction, the investigator was not aware of the treatment as animals and syringes were picked by an independent researcher outside the experimental room.

(continued on next page)
For baseline performance and testing, animals were acclimatised to the testing room/experimental set-up as described above. Importantly, animals were placed into the burrowing cage alone rather than in pairs, and allowed to burrow for 60 minutes. At the end of the session, the weight of the displaced substrate was measured. For this, each burrowing tube was weighed before and after burrowing to allow for accurate measurement. In each study, 1 to 3 baseline sessions were performed. Animals were randomised into experimental groups according to their baseline performance by evenly distributing “poor” and “good” burrowers between groups; it was ensured that each experimental group burrowed on average a similar mean amount to reduce the impact of this source of variability on following experiments (Table 1). Burrowing performance was measured on days 1, 2, 3, 7, and 10 after CFA/sham injections.

2.5. Study design

This prospective multicentre validation study was divided into 2 parts. The first part comprised the experimental work at participating laboratories following the basic protocol (Supplementary digital content 1, http://links.lww.com/PAIN/A302). All data were centrally collected at Grünenthal and then processed for statistical analysis at H. Lundbeck A/S, who did not participate in the experimental work or data collection.

All burrowing experiments were performed during the light phase. The primary outcome was the change from baseline in the amount burrowed, with a negative value representing a decrease in burrowing behaviour. For this, the mean of the last 2 baseline sessions was used as the baseline value. In cases in which only one session was performed, this was used as the baseline value. The secondary outcome was the effect of protocol variations on burrowing behaviour. Because the study was not designed to investigate the effect of protocol variations, no statistical hypothesis testing has been performed for these elements.

2.6. Statistical analysis

Details for sample size calculation are given in Table 1. The study was powered for individual self-contained experiments at each centre.
For statistical analyses, we used an approach adapted from an intention-to-treat and per-protocol analysis. For this, we separately analysed 2 populations: the total population (TP), including all animals that were allocated to treatment groups, and the selected population (SP) population, including all animals that were allocated to treatment groups and burrowed over 500 g at baseline, a cutoff that has been previously described for burrowing.

Change from baseline was analysed using a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM). Experimental group (naive, sham, CFA), time of assessment (days 1, 2, 3, 7, and 10), group-by-time interaction, and baseline burrowing-by-time interaction were used as fixed effects for analysis of individual studies. Laboratory ID (8 participating centres) was added as the fixed effect for combined analysis across all studies. An unstructured covariance design was used to model the within-animal errors (type 3 tests). A Kenward–Roger approximation was used to estimate denominator degrees of freedom. The analysis was based on the missing-at-random assumption and performed using all available observations. The mean differences between naive and CFA, and sham and CFA, were estimated based on the least squares means for the treatment-by-time interaction in the MMRM model. The estimates are presented with 95% confidence intervals (CIs). A P value <0.05 was considered statistically significant. Because of the exploratory nature of the study, no adjustment for multiplicity was made. Within the text, burrowing data are presented as mean with 95% CI.

Strain (Sprague-Dawley, Wistar [Wistar and Wistar Hannover]), animal weight at the start of study (200-225 g, 225-250 g, >250 g), substrate size (2-5 mm, 4-8 mm, <1 mm [sand]), number of training sessions (2, 3), weight of substrate provided (2000 g, 2500 g), and sex of the experimenter (female, male, mixed) were investigated as protocol variations. As many of these factors were given over a range, with equal values for all animals within a study or laboratory, ranges have been redefined to fit a set of studies. This could result in overlapping values but always with clear cutoff values given by the original ranges. When applicable, the mean values of the minimum and maximum were used as numeric values. No statistical analyses have been performed on explanatory factors, and data were presented as observational findings with descriptive statistics only.

For sample size recommendations, sample sizes for a range of mean differences and SD were calculated. Sample size calculations were based on a 2-sample t test informed by the descriptive statistics and MMRM analysis for both TP and SP populations 24 hours after CFA injections. Analyses were performed following the exact method under a fixed scenario adopting normal distribution, 2 sides, a nominal power of 0.8, α = 0.05, and a null hypothesis assuming no difference between groups. For data manipulation and analysis, SAS software Version 9.4 (SAS Institute Inc, Cary, NC, USA) was used.

### 3. Results

#### 3.1. Study profile, animal characteristics, and experimental design

Eleven studies, completed between June 2013 and September 2014, were performed at 8 different laboratories across 4 countries (Fig. 2). Of 255 recorded animals, 249 were randomly assigned to 3 treatment groups: naive (49), sham (96), and CFA-treated (104). No adverse events were reported across the studies, with animals tolerating sham injections, CFA injections, and behavioural testing well. Six animals that displayed no burrowing behaviour during baseline sessions were excluded before group assignment. The study protocol regarding animal exclusion was adjusted over the course of the project, and no further animals were excluded at this stage, as a TP and SP analysis approach was favoured instead. All 249 animals allocated to treatment groups were part of the TP analysis, whereas 200 animals were included in the SP analysis; the latter was defined as all rats in the TP population that burrowed at least 500 g on average at baseline (Fig. 2). Although all eleven studies included a CFA group, control groups varied. One study chose naive as the control group, 4 studies chose sham as the control group, and 6 studies chose naive and sham as control groups.
All 11 studies reported randomisation procedures for group allocation. In 4 studies, no allocation concealment was performed. In the remaining 7 studies, allocation concealment procedures were followed during model induction; however, because of the occasional presence of a slight yellowish colour and the higher viscosity of CFA suspensions, allocation concealment could be maintained only in 2 studies. Ten studies reported that outcome assessment was performed blinded. Oedema of the paw in some CFA animals potentially revealed the identity of the experimental group to the investigator in 7 of these studies.

In all studies, male rats from outbred albino strains with a starting weight range of 200 to 335 g were used (Table 3). All CFA-injected animals received 100 μg CFA (1 mg/mL in 100 μL, intraplantar), whereas sham animals received a saline or incomplete Freund adjuvant injection of the same volume (Table 3). Animals were tested at days 1, 2, 3, 7, and 10 after CFA/sham injections with the exception of the study performed at Eli Lilly (USA), in which animals were only tested up to 7 days. Because of a weighing error, in study 3 at Eli Lilly (United Kingdom), only 2000 g substrate was provided.

Some variations in the features of experimental design were permitted across studies (Fig. 3). Time for animals to acclimatise to the test room ranged from 0 minutes to 60 minutes. Time for acclimatisation to the experimental set-up (empty test cage) again ranged from 0 minutes to 60 minutes. The study performed at the University of Heidelberg introduced a 3-hour habituation session for animals in their home cages before training started. On the first training day, another habituation session was performed in the morning, during which animals were allowed to acclimatise to the experimental room for 30 minutes and for a further 30 minutes in the empty test cage. An empty test tube was then added to study the...
the cage, and animals spent another 60 minutes to acclimatise. The training session was then conducted in the afternoon. As animals were still in the behavioural room, they were allowed to acclimatise for only 30 minutes to the experimental set-up before an empty burrowing tube was placed in the test cage for 30 minutes. Subsequently, a filled burrowing tube was added, and animals were allowed to burrow for 60 minutes. All subsequent training sessions followed the standard protocol (Fig. 3). For all training, baseline, and testing sessions, animals were allowed to burrow for 60 minutes. The only exception was for training sessions performed at Boehringer Ingelheim, in which animals were allowed to burrow for 120 minutes. Across studies, 2 to 3 training sessions and 1 to 3 baseline sessions were performed (Fig. 3). Testing sessions followed the same protocol as outlined for baseline sessions.

3.2. Burrowing behaviour over time

Baseline burrowing behaviour was highly variable between individual animals and ranged from 0 to 2286 g at baseline. Mean baseline burrowing across studies was 1113 g (95% confidence interval: 1041-1185 g) and 1329 g (1271-1387 g) in the TP and SP population, respectively (Supplemental digital content 2, http://links.lww.com/PAIN/A303). To reduce the impact of baseline variability between studies, mean change from baseline in the amount burrowed was chosen as the primary outcome, with a negative value representing a decrease in burrowing behaviour. A summary of the nonnormalised data can be found in the supplemental material (Supplemental digital content 2, http://links.lww.com/PAIN/A303). Statistical analysis of individual studies showed, for both TP and SP populations, that in 7 of 11 studies, burrowing behaviour was significantly suppressed 24 hours after CFA injections as compared with a control group (naive or sham) (Tables 4 and 5). Forty-eight hours after CFA injections, significantly suppressed burrowing behaviour was observed in 7 studies for the TP and in 6 studies for the SP population. Burrowing performance over time in naive and sham groups from individual studies was largely within the 95% CI of the overall mean (all animals combined of relevant group) (Fig. 4A–D). In contrast, burrowing behaviour in CFA groups was more variable (Fig. 4E, F). Burrowing behaviour was comparable between the TP and SP population, with a burrowing deficit of

### Table 3
Animal and model characteristics.

| Animal characteristics | Animal characteristics | Animal characteristics | Animal characteristics |
|------------------------|------------------------|------------------------|------------------------|
| **Strain** | **Supplier** | **Country of origin** | **Supplier** | **Country of origin** |
| Wistar Hannover | Charles River | Germany | Sprague-Dawley | Charles River | United Kingdom |
| Sprague-Dawley | Charles River | United Kingdom | Sprague-Dawley | Charles River | Japan |
| Sprague-Dawley | Charles River | United Kingdom | Sprague-Dawley | Charles River | United Kingdom |
| Sprague-Dawley | Harlan | USA |

| Model characteristics | Model characteristics | Model characteristics | Model characteristics |
|------------------------|------------------------|------------------------|------------------------|
| **CFA supplier** | **Form of CFA** | **Concentration, mg/mL** | **Volume, µL** | **Vehicle used for CFA solution** |
| Sigma (Germany) | Liquid | 1 | 100 | NA |
| DIFCO Laboratories (United Kingdom) | Powder | 1 | 100 | Incomplete Freund adjuvant |
| Sigma (Germany) | Liquid | 1 | 100 | NA |
| Sigma (Japan) | Liquid | 1 | 100 | NA |
| Sigma (United Kingdom) | Liquid | 1 | 100 | NA |
| Sigma (United Kingdom) | Liquid | 1 | 100 | NA |
| Sigma (United Kingdom) | Liquid | 1 | 100 | NA |
| Sigma (United Kingdom) | Liquid | 1 | 100 | NA |
| Sigma (St. Louis, USA) | Liquid | 1 | 100 | NA |

CFA, complete Freund adjuvant; NA, not applicable because CFA solution was provided in liquid form by the supplier and was injected neat; when CFA was purchased from Sigma, saline was used for sham injections.
−374 g (−479 to −269 g) and −498 g (−609 to −386 g), at 24 hours after CFA injection, respectively.

Statistical analysis of the combined data showed suppressed burrowing peaked 24 hours after CFA injections and, although less pronounced, was present for up to 10 days, gradually regressing to baseline values (Fig. 5). Each of the fixed effects, namely, group, time of assessment, laboratory ID, group-by-time interaction, and baseline burrowing-by-time interactions, significantly contributed to heterogeneity (Table 6).

### 3.3. Effect of protocol variations on burrowing behaviour—observational data

The secondary outcome was the effect of local protocol variations on burrowing behaviour. Because of the nature of the study design, no statistical hypothesis testing was performed, and data are presented as observational findings only with descriptive statistics. Observations were made including all animals allocated to a treatment group (TP population), focusing on the 24-hour time point at which burrowing deficits were most pronounced. Data summarising burrowing behaviour at all time points by protocol variation can be found in the supplemental material (Supplemental digital content 3, http://links.lww.com/PAIN/A304). There seemed to be no strain difference between Wistar and Sprague-Dawley rats, but there was a tendency towards a more prominent CFA-associated burrowing deficit in animals of lower body weight. Burrowing deficits after CFA injection were also more noticeable when a substrate of smaller size was used (Fig. 6). An increased number of training sessions negatively affected burrowing behaviour in naive animals. In addition, when a larger amount of substrate was provided, the burrowing deficit was more pronounced. The sex of the experimenter also might have affected burrowing behaviour; increased burrowing behaviour was observed with a male experimenter in naive animals, whereas in contrast, a more pronounced burrowing deficit was present in CFA animals with a male experimenter (Supplemental digital content 4, http://links.lww.com/PAIN/A305). It should be noted that because the study was not designed to identify variables that affect burrowing outcome, confounding factors were identified, and observational data should be interpreted cautiously. In particular, animals were unevenly distributed across groups. An n = 4 was recorded for the naive group that underwent 3 training sessions and for the naive group that was tested by a male experimenter. Furthermore, sand as the burrowing substrate and provision of 2000 g of the substrate was only reported by one centre in one study each.

### 4. Discussion

A prospective preclinical multicentre study across 8 laboratories assessed the reliability of CFA-associated suppressed burrowing in rats. Overall, reduced burrowing was partially replicated at 6 of the 8 participating centres with an element of variability between and within centres. The prospective multicentre approach was important in that it enabled the evaluation of variability of suppressed burrowing, and it could prove important for future studies aiming to identify the factors underlying such variability.

We showed prominent CFA-associated suppressed burrowing in 7 of 11 studies. Consistent with our results, CFA-associated suppression of behaviour previously has been shown in feeding behaviours, locomotion, and operant behaviours. Analysis of the combined data demonstrated that burrowing deficits peaked 24 hours after CFA injection, although with high variability between individual studies, but some studies show no suppression of burrowing. Notably, the original sample size calculation was based on a pilot study with a large effect size; however, across centres, burrowing deficits ranged from 1570 to 273 g. Therefore, some studies were underpowered using the originally estimated sample size. This could result in a reduced
## Table 4

| Group                                      | N number | 24 h | 48 h | 72 h | 7 d   | 10 d  |
|--------------------------------------------|----------|------|------|------|-------|-------|
| **Boehringer Ingelheim Pharma**            |          |      |      |      |       |       |
| Naive                                      | 9        |      |      |      |       |       |
| Sham                                       | 10       |      |      |      |       |       |
| CFA                                        | 10       |      |      |      |       |       |
| **Grüenthal GmbH**                         |          |      |      |      |       |       |
| Naive                                      | 7        |      |      |      |       |       |
| Sham                                       | 7        |      |      |      |       |       |
| CFA                                        | 7        |      |      |      |       |       |
| **Heidelberg University**                  |          |      |      |      |       |       |
| Naive                                      | 8        |      |      |      |       |       |
| Sham                                       | 8        |      |      |      |       |       |
| CFA                                        | 8        |      |      |      |       |       |
| **Asahi Kasei Pharma**                     |          |      |      |      |       |       |
| Sham                                       | 12       |      |      |      |       |       |
| CFA                                        | 12       |      |      |      |       |       |
| **University of Manchester**               |          |      |      |      |       |       |
| Naive                                      | 8        |      |      |      |       |       |
| Sham                                       | 8        |      |      |      |       |       |
| CFA                                        | 8        |      |      |      |       |       |
| **Eli Lilly and Company (United Kingdom)** |          |      |      |      |       |       |
| Study 1                                    |          |      |      |      |       |       |
| Naive                                      | 9        |      |      |      |       |       |
| Sham                                       | 9        |      |      |      |       |       |
| CFA                                        | 9        |      |      |      |       |       |
| Study 2                                    |          |      |      |      |       |       |
| Sham                                       | 16       |      |      |      |       |       |
| CFA                                        | 16       |      |      |      |       |       |
| Study 3                                    |          |      |      |      |       |       |
| Sham                                       | 16       |      |      |      |       |       |
| CFA                                        | 16       |      |      |      |       |       |
| **Imperial College London**                |          |      |      |      |       |       |
| Study 1                                    |          |      |      |      |       |       |
| Naive                                      | 4        |      |      |      |       |       |
| Sham                                       | 4        |      |      |      |       |       |
| CFA                                        | 4        |      |      |      |       |       |
| Study 2                                    |          |      |      |      |       |       |
| Naive                                      | 8        |      |      |      |       |       |
| Sham                                       | 8        |      |      |      |       |       |
| CFA                                        | 8        |      |      |      |       |       |
| **Eli Lilly and Company (USA)**            |          |      |      |      |       |       |
| Sham                                       | 10       |      |      |      |       |       |
| CFA                                        | 10       |      |      |      |       |       |

| % of naive vs CFA comparisons reporting significant suppression of burrowing | 71 | 29 | 14 | 14 | 0 |
| % of sham vs CFA comparisons reporting significant suppression of burrowing | 60 | 80 | 30 | 10 | 10 |

Data shown as mean (95% confidence interval) and analysed using a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM) approach with fixed effects (group [naive, sham, CFA]), time of assessment (days 1, 2, 3, 7, 10), group-by-time, and baseline burrowing-by-time interactions. P values given for naive and sham groups refer to comparison with the complete Freund adjuvant (CFA) group. Significance level has been set at $P < 0.05$, bold values indicate significant differences to the CFA treated groups.

CFA, complete Freund adjuvant.
| Group                                      | N number | 24 h              | 48 h              | 72 h              | 7 d              | 10 d             |
|--------------------------------------------|----------|-------------------|-------------------|-------------------|------------------|------------------|
| **Boehringer Ingelheim Pharma GmbH**       |          |                   |                   |                   |                  |                  |
| Naive                                      | 9        | $-38$ to $-289$    | $-289$ to $-193$, | $-91$ to $-320$, | $-16$ to $-179$, |                  |
| Sham                                       | 10       | $-60$ to $-280$    | $-280$ to $160$,  | $-39$ to $-245$, | $-20$ to $-198$, |                  |
| CFA                                        | 10       | $-740$ to $-960$   | $-960$ to $-520$, | $-369$ to $-575$,| $-308$ to $-526$,|                  |
| **Grunenthal GmbH**                        |          |                   |                   |                   |                  |                  |
| Naive                                      | 7        | $-49$ to $-163$    | $-163$ to $-261$, | $-227$ to $-18$,  | $-70$ to $-308$, |                  |
| Sham                                       | 7        | $-84$ to $-126$    | $-126$ to $294$,  | $-56$ to $-187$,  | $-27$ to $-403$, |                  |
| CFA                                        | 7        | $-896$ to $-1110$  | $-1110$ to $-689$,| $-307$ to $-552$,| $-653$ to $-1009$,|                  |
| **Heidelberg University**                  |          |                   |                   |                   |                  |                  |
| Naive                                      | 6        | $303$ to $-142$    | $-142$ to $748$,  | $265$ to $-162$,  |                  |                  |
| Sham                                       | 6        |                   |                   |                   |                  |                  |
| CFA                                        | 7        | $-106$ to $-518$   | $-518$ to $306$,  | $345$ to $-51$,   |                  |                  |
| **Asahi Kasei Pharma**                     |          |                   |                   |                   |                  |                  |
| Sham                                       | 7        | $-87$ to $-524$    | $-524$ to $350$,  | $-68$ to $-493$,  |                  |                  |
| CFA                                        | 6        |                   |                   |                   |                  |                  |
| **University of Manchester**               |          |                   |                   |                   |                  |                  |
| Naive                                      | 6        | $119$ to $-179$    | $-179$ to $417$,  | $399$ to $11$,    |                  |                  |
| Sham                                       | 6        |                   |                   |                   |                  |                  |
| CFA                                        | 6        | $507$ to $282$     | $282$ to $878$,   | $838$ to $450$,   |                  |                  |
| **Eli Lilly and Company1 (United Kingdom)**|          |                   |                   |                   |                  |                  |
| Study 1                                    |          |                   |                   |                   |                  |                  |
| Naive                                      | 8        | $-302$ to $-661$   | $-661$ to $57$,   | $-361$ to $-161$, | $-7$ to $-521$,  |                  |
| Sham                                       | 7        | $-493$ to $-862$   | $-862$ to $-136$, | $-232$ to $-518$, | $-357$ to $-902$,|                  |
| CFA                                        | 7        | $-958$ to $-1321$  | $-1321$ to $-595$,| $-798$ to $-1068$,| $-401$ to $-944$,|                  |
| **Study 2**                                |          |                   |                   |                   |                  |                  |
| Sham                                       | 8        | $15$ to $-248$     | $-248$ to $278$,  | $49$ to $-243$,   |                  |                  |
| CFA                                        | 10       | $-259$ to $-494$   | $-494$ to $-24$,  | $-322$ to $-583$  |                  |                  |
| **Study 3**                                |          |                   |                   |                   |                  |                  |
| Sham                                       | 12       | $266$ to $-108$    | $-108$ to $640$,  | $404$ to $102$,   |                  |                  |
| CFA                                        | 13       | $-351$ to $-698$   | $-698$ to $-4$,   | $-135$ to $-425$, |                  |                  |
| **Imperial College London**                |          |                   |                   |                   |                  |                  |
| Study 1                                    |          |                   |                   |                   |                  |                  |
| Naive                                      | 4        | $660$ to $107$     | $107$ to $1213$,  | $229$ to $-283$,  |                  |                  |
| Sham                                       | 3        | $-162$ to $-815$   | $-815$ to $491$,  | $-114$ to $-720$, |                  |                  |
| CFA                                        | 3        | $-985$ to $-1614$  | $-1614$ to $-356$,| $190$ to $-392$,  |                  |                  |
| **Study 2**                                |          |                   |                   |                   |                  |                  |
| Naive                                      | 4        | $-974$ to $-1186$  | $-1186$ to $-162$,| $-193$ to $-640$, |                  |                  |
| Sham                                       | 3        | $63$ to $-531$     | $-531$ to $657$,  | $317$ to $-200$,  |                  |                  |
| CFA                                        | 4        | $-752$ to $-1266$  | $-1266$ to $-238$,| $-952$ to $-1401$,|                  |                  |
| **Eli Lilly and Company2 (USA)**           |          |                   |                   |                   |                  |                  |
| Sham                                       | 10       | $48$ to $-164$     | $-164$ to $260$,  | $46$ to $-109$,   |                  |                  |
| CFA                                        | 10       | $-742$ to $-954$   | $-954$ to $-530$, | $-335$ to $-490$, |                  |                  |

| % of naive vs CFA comparisons reporting significant suppression of burrowing | 57 | 42 | 14 | 14 | 0 |
| % of sham vs CFA comparisons reporting significant suppression of burrowing | 50 | 50 | 30 | 30 | 10 | 10 |
chance to detect a true effect. We have calculated sample size recommendations for a range of mean differences and SD (Supplemental digital content 5, http://links.lww.com/PAIN/A306) to provide guidance for future studies. Increasing the sample size of underpowered studies would result in a more accurate estimated effect size and could reduce variability within studies. Group allocation, time of assessment, laboratory ID, group-by-time interaction, and baseline burrowing-by-time interaction all significantly contributed to the heterogeneity across studies. Adjustment for laboratory ID showed a change from baseline similar to data without adjustment, demonstrating that suppression of burrowing is robust across laboratories, despite the variability in the effect size.

We also observed effects of local protocol variations on burrowing behaviour. No statistical analyses were performed on these observations, as the study was not designed to formally detect such effects. Confounding factors such as uneven distribution of animals across groups and variables reported only by one centre should be considered when interpreting the data. Observations made were reported both for transparency and to identify variability factors in burrowing behaviour meriting future study. Although strain differences have been reported for other

Figure 4. Burrowing behaviour in individual studies—mean change from the baseline amount burrowed. (A, C, E) Total population (n = 249): burrowing behaviour in individual studies in naive (A), sham (C), and complete Freund adjuvant (CFA) (E) groups. Gray area represents the mean with 95% confidence interval (CI). (B, D, F) Selected population (n = 200): burrowing behaviour in individual studies in naive (B), sham (D), and CFA (F) groups. Gray area represents the mean with 95% CI.
outcome measures,\textsuperscript{8,42,63} in this study, no strain differences between Sprague-Dawley and Wistar rats were observed for suppressed burrowing. Animals with a lower body weight developed an increased burrowing deficit, whereas burrowing in sham or naive groups was unaffected. As all animals received the same volume and concentration of CFA, it may be that smaller animals received relatively more CFA per paw mass, resulting in a more severe inflammatory response leading to a larger burrowing deficit. The suppression of burrowing observed with a smaller amount of provided substrate is most likely due to the reduced amount of substrate available. In naive animals, we observed an increased burrowing deficit with male experimenters as compared with female experimenters or mixed experimenter teams. Because a male experimenter was only reported by one centre and only 4 rats were in this group, this result may be due to chance or a centre-specific effect. A study in mice showed that the experimenter’s sex affects pain outcome measures in mice.\textsuperscript{67} Further studies are required to verify whether burrowing behaviour in rats is also affected by the experimenter’s sex. Additional studies would be required to assess the impact of the substrate size and number of training sessions on burrowing, specifically whether a small sized substrate is superior to larger sized substrate and whether an increased number of training sessions reduces burrowing behaviour.

To assess the effect of excluding “poor” burrowers, we analysed the TP (all animals allocated to treatment groups) and an SP (all allocated animals that burrowed above 500 g at baseline), an approach adapted from the intention-to-treat and per-protocol analyses used in clinical trials.\textsuperscript{22} In both populations, the pattern of suppressed burrowing was comparable, which suggests that suppressed burrowing is a robust measure. Therefore, we recommend not to exclude “poor” burrowers. Excluding animals would increase variability, resulting in a less accurate effect size estimates. Exclusions could also result in attrition bias, an issue particularly important for studies using relatively small sample sizes.\textsuperscript{1,13,24}

Adding nonevoked ethologically relevant outcome measures to assess the global impact of pain previously has been suggested as a potential means to improve translation.\textsuperscript{49,56,68,75} Development and validation of these measures is of key relevance, particularly as spontaneous pain, functioning, and quality of life are primary outcome measures in clinical trials. Suppression of burrowing reflects the global impact of purportedly pain-induced reduction of general “well-being”.\textsuperscript{27–29,74} Although suppressed burrowing is not a pain-specific test, treatment with known analgesics has been shown to attenuate decreased burrowing behaviour in various pain models, suggesting a pain-specific component.\textsuperscript{2,9,35,61,62} A limitation of our study was the lack of an independent validation of suppressed burrowing as indicative of a pain-specific outcome measure; it will be crucial to address the interdependence of this connection in future studies. Correlation with other nonevoked

| Fixed effect | Type 3 tests of fixed effects. | Total population | Selected population |
|--------------|-------------------------------|------------------|---------------------|
|              | Nominator, \(df\) | Denominator, \(df\) | \(F\) | \(Pr > F\) | Nominator, \(df\) | Denominator, \(df\) | \(F\) | \(Pr > F\) |
| Experimental group (naive, sham, CFA) | 2 | 239 | 13.44 | <0.0001* | 190 | 16.85 | <0.0001* |
| Time of assessment (days 1, 2, 3, 7, 10) | 4 | 238 | 5.75 | 0.0002* | 189 | 2.49 | 0.0444* |
| Laboratory ID (8 participating centres) | 7 | 239 | 4.74 | <0.0001* | 189 | 7.68 | <0.0001* |
| Group-by-time interaction | 8 | 340 | 3.59 | 0.0005* | 270 | 4.24 | <0.0001* |
| Baseline burrowing-by-time interaction | 5 | 249 | 14.50 | <0.0001* | 193 | 10.34 | <0.0001* |

Data were analysed using a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM).
* Significant difference between variables within fixed effect.
CFA, complete Freund adjuvant; \(df\), degrees of freedom; \(Pr\), probability.
pain-related outcomes would be essential, particularly as a lack of correlation has been shown between burrowing performance and evoked mechanosensory thresholds in a neuropathic pain rat model, which suggests that suppressed burrowing may reflect a pain component not directly linked to sensory gain. Furthermore, pharmacological validation studies should be conducted, preferably informed by robust meta-analyses of clinical trials to guide both drug and dose selection. First, it should be demonstrated to what extent clinically efficacious drugs reverse suppression of burrowing. In the CFA model, ibuprofen has been shown to reverse suppressed burrowing, whereas gabapentin, which has a large body of evidence supporting efficacy in neuropathic, but not inflammatory, pain, is appropriately inefficacious, suggesting good pharmacologic sensitivity of suppressed burrowing.

Figure 6. Burrowing performance 24 hours after complete Freund adjuvant (CFA) injections in subgroups factoring in protocol variations (total population). (A) Burrowing dependent on strain. (B) Burrowing dependent on the weight of animals at the start of study. (C) Burrowing dependent on the substrate size. Data shown as single values (diamonds) and mean (square) with 95% confidence interval (whiskers).
Second, compounds such as neurokinin 1 antagonists and cannabinoid 2 receptor agonists that have been shown to be efficacious in animal pain models measuring evoked endpoints but have failed in clinical trials should be tested to assess the degree of pain specificity of burrowing.

We chose the CFA model, as we expected a high likelihood of ongoing spontaneous pain. Importantly, sex differences in behavioural responses have been shown in rodents. As the primary interest of our study was to evaluate burrowing across centres and not establish the model’s predictive validity as a pain outcome measure, only male rats were used. However, future validation studies should include female rats.

A multicentre approach for preclinical studies is very novel. Similar to clinical multicentre trials, the study design should be of a high standard, and results should be reported transparently. In this study, all participating centres followed a basic protocol that was previously reviewed and agreed upon by all parties; however, minor changes were permitted to pragmatically accommodate for local variations in laboratory practice and procedures. No detailed specifications were given regarding the scope of these changes, inevitably resulting in some degree of uncontrolled heterogeneity between studies. An external review of the protocol could have identified this issue before study start. Variations were also reported concerning bias reduction procedures. Although guidance on Good Laboratory Practice was given, there was variability between centres as to the extent to which such practice was followed, most notably as a result of constraints imposed by established local procedures. Future preclinical multicentre studies should not only provide Good Laboratory Practice training and validation but also establish an independent central monitor, similar to phase III clinical trials, to ensure protocol compliance and bias reduction. It should be noted that despite following bias reduction procedures, because of CFA-induced paw oedema, allocation concealment and blinding could not be maintained in all studies, potentially resulting in an overestimation of the effect size. As it was not possible to control for all model-specific factors, it is crucial to report data as transparently as possible to clearly highlight study limitations related to internal validity issues. In clinical trials, the near-universal implementation of the CONSORT reporting guidelines has noticeably improved reporting rigour and transparency. Although ARRIVE guidelines provide a similar framework for preclinical studies, they are not yet as well established as CONSORT; however, a similar positive impact on preclinical studies is expected as these guidelines achieve broader acceptance and implementation. In this study, we reported according to ARRIVE guidelines and presented the data as transparently as possible. Recommendations for future studies, based on our practical experience, are summarised in Table 7. An audio abstract of this study is available in the supplemental material (Supplemental digital content 7, Audio, http://links.lww.com/PAIN/A310). In conclusion, our approach demonstrates how implementation of a multicentre study design to evaluate novel preclinical outcome measures can yield robust data and can help accelerate the validation of outcome measures, pain models, and pharmacological interventions. This hopefully may help inform the design and conduct of similar future multicentre studies.

Conflict of interest statement

The following authors are/ were employees of the following companies at the time this study was undertaken: R. Wodarski, M. Ligocki, D. Li, and J. D. Kennedy; employees of Eli Lilly and Company; C. Uttenius and C. Stenfors: employees of Astra Zeneca; L. A. Bryden and A. Pekcec: employees of Boehringer Ingelheim; T. Christoph, A. Robens, and K. Rutten: employees of Grünenthal; K. Uto, S. Koyama, and K. Yamamoto: employees of Asahi Kasei; A. Lindsten and M. Segerdahl: employees of Lundbeck, Imperial College London: A. S.C. Rice also received grant funding from Pfizer and Astellas. Heidelberg University: R.-D. Treede also received research funding from AbbVie, Astellas, and Boehringer Ingelheim. The other authors have no conflicts of interest to declare.

Joint first authors: R. Wodarski, A. Delaney, C. Uttenius, and R. Morland; Joint senior authors: K. Rutten and A. Rice.

Acknowledgements

This European project has received support from the Innovative Medicines Initiative (IMI) Joint Undertaking (under grant agreement number 115007), resources of which are composed of financial contributions from the European Union’s Seventh Framework Programme (FP7/2007-2013) and European Federation of Pharmaceutical Industries and Associations (EFPIA) companies in-kind contribution (see http://www.imieurope.org/ for details). Eli Lilly and Company United Kingdom: We would like to thank Dr Gary Gilmour for his assistance in bringing forward this publication. Karolinska Institute: A. Delaney also received grant funding from Ulla & Gustaf af Ugglas Foundation; EU Project FP7-Health-2013-Innovation-1602919-2. Boehringer Ingelheim: We would like to thank Stacey Gould for technical assistance.

Table 7

| Table 7 | Recommendations for future multicentre animal studies. |
| --- | --- |
| **Recommendation** | **Study design** |
|  | Appoint study manager |
|  | Protocol development and refinement: each participating centre has to agree on protocol |
|  | External review might be beneficial to identify problems |
|  | Required training (inclusive Good Laboratory Practice) should be identified and provided at this stage |
|  | Plan for statistical analysis should be agreed upon in the protocol—changes to the analysis should be reported and explained |
|  | Independent protocol registration might be beneficial to ensure compliance during experimental phase |

**Experimental phase**

- Centralised administration: centralised coordination and monitoring, Web-based data entry portal, randomisation process
- Enables easier monitoring and quality assurance
- Proactive site monitoring and audit through central administration could be beneficial to identify problems during the experimental phase

**Reporting**

- Transparent reporting such as ARRIVE guidelines—it may be that some adaptation of these guidelines is required to assist reporting of preclinical multicentre studies
Appendix A. Supplemental Digital Content

Supplemental Digital Content associated with this article can be found online at http://links.lww.com/PAIN/A302, http://links.lww.com/PAIN/A303, http://links.lww.com/PAIN/A305, http://links.lww.com/PAIN/A306, http://links.lww.com/PAIN/A307, http://links.lww.com/PAIN/A310.

Supplemental media

A supplemental video accompanying this article can be found online at http://links.lww.com/PAIN/A310.

Article history:

Received 11 January 2016
Received in revised form 19 May 2016
Accepted 20 May 2016

References

[1] Abdel-Sattar M, Krauth D, Anglemyer A, Bero L. The relationship between risk of bias criteria and research outcomes, and study sponsorship in a cohort of preclinical thiazolidinedione animal studies: a meta-analysis. Evidence Based Preclin Med 2014;1:11–20.

[2] Andrews N, Legg E, Lisak D, Issop Y, Richardson D, Harper S, Pheby T, Huang W, Burgess G, Machin I, Rice AS. Spontaneous burrowing behaviour in the rat is reduced by peripheral nerve injury or inflammation. Sci Rep 2016;6:30909–14. Eur J Pain 2012;16:845–95.

[3] Andrews NA, Latremoliere A, Barsabai AM, Mogli JS, Porreca F, Rice ASC, Woolf CJ, Currie GL, Dworkin RH, Eisenach JC, Evans S, Gewandter JS, Gover TD, Handwerker H, Huang W, Iyengar S, Jensen MP, Kennedy JD, Lee N, Levine J, Lidster K, Machin I, McDermott MP, McMahon SB, Price TJ, Ross SE, Scherrer G, Saul RB, Sena ES, Silva E, Stone L, Svensson CI, Turk DC, Whiteside G. Ensuring transparency and minimalization of methodologic bias in preclinical pain research: PPREECE considerations. PAIN 2016;157:901–9.

[4] Baker D, Lidster K, Sottomayor A, Amor S. Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for preclinical animal studies. PLoS Biol 2014;12:e1001756.

[5] Barnett SA. The rat: A study in behavior. Chicago: University of Chicago Press, 1975.

[6] Barnett SA, Strain-dependent modification of neuropathic pain behaviour in the rat hindpaw by a priming painful trigeminal nerve injury. PAIN 2002;97:203–12.

[7] Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. Brain Res 2014;158:158–65.

[8] Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. Brain Res 2014;158:158–65.

[9] Bryden LA, Nicholson JR, Doods H, Pekcec A. Deficits in spontaneous burrowing behavior in the rat bilateral monosodium iodoacetate model of osteoarthritis: an objective measure of pain-related behavior and analgesic efficacy. Osteoarthr Cartil 2015;23:1605–12.

[10] Calvo M, Karu K, Olausen HR, Bathgate G, Okuse K, Bennett DL, Huang W, Calvo M, Karu K, Olausen HR, Bathgate G, Okuse K, Bennett DL, Rice AS. A clinically relevant rodent model of the HIV antiretroviral drug stavudine induced painful peripheral neuropathy. PAIN 2013;154:560–75.

[11] Canville DJ, Arcement J, Des Marais P, LeBlanc L, Nyberg T, Baribeau C. Reversal of mechanical hypersensitivity and anxiety-related behavior, and the influence of analgesic drugs. Neuroscience 2007;141:496–508.

[12] Carmona J, Molina L, Martinez-Sanchez D, Gaya M, Barrios B, Barrios J, Absence of signs of pain in a human perception. PAIN 2000;85:301–3.

[13] Carmona J, Molina L, Martinez-Sanchez D, Gaya M, Barrios B, Barrios J, Absence of signs of pain in a human perception. PAIN 2000;85:301–3.

[14] Carmona J, Molina L, Martinez-Sanchez D, Gaya M, Barrios B, Barrios J, Absence of signs of pain in a human perception. PAIN 2000;85:301–3.

[15] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[16] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[17] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[18] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[19] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[20] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[21] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[22] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[23] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[24] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[25] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[26] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[27] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[28] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[29] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[30] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[31] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[32] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[33] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[34] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[35] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[36] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[37] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[38] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[39] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[40] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[41] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.

[42] Cesarovic N, Jirkof P, Seifert B, Arras M, Burrowing behavior as an indicator of post-laparotomy pain in mice. Front Behav Neurosci 2010;4.
[38] Macleod MR, Fisher M, O’Collins V, Sena ES, Dimagli U, Bath PM, Buchanan A, van der Worp HB, Traystman R, Minematsu K, Donnan GA, Howells DW. Good laboratory practice: preventing introduction of bias at the bench. Stroke 2009;40:e50–2.

[39] Maier C, Baron R, Tolle TR, Binder A, Birbaumer N, Birkenfeld F, Gierthmuellen J, Flor H, Geber C, Hulse V, Krumova EK, Landwehrmeyer GB, Magier W, Maholcher H, Richter H, Rokke R, Schieren A, Schwarz A, Sommer C, Tronnier V, Uceyler N, Valet M, Wasner G, Treede RD. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. PAIN 2010;150:439–50.

[40] Mandillo S, Tucci V, Holter SM, Meziane H, Al Banchaabouchi M, Kalinkin M, Lad H, Nolan PM, Quagazzam AL, Cogheli E, Gale K, Golini E, Jacquot S, Krezel W, Parker A, Ret F, Schneider I, Maraziti D, Auer J, Brown SDM, Chambon P, Rosenthal N, Tocciini-Valentini G, Wurth W, Reliability, robustness, and reproducibility in mouse behavioral phenotyping: a cross-laboratory study. Physiol Genomics 2008;34:243–55.

[41] Miletic T, Kovacevic–Jovanovic V, Stanjovic S, Vujic V, Kosec D, Mitic K, Dimitrijevic M. Strain differences and the role for HSP74 and HSP70 in adjuvant arthritis in rats. Scand J Immunol 2006;64:623–32.

[42] Naj-Esfahani H, Vaseghi G, Safaeian L, Pilehvarian AA, Abed A, Rafieian-Kopaei M. Gender differences in a mouse model of chemotherapy-induced neuropathic pain. Lab Anim 2016;50:15–20.

[43] National Centre for the Replacement, Refinement and reduction of animals in research. ARRIVE guidelines. Available at: www.nc3rs.org.uk/arrive-guidelines.

[44] National Institute of Neurological Disorders and Stroke. Improving the quality of NINDS-supported preclinical and clinical research through rigorous study design and transparent reporting. 2015 Available at: http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf. Accessed October 22, 2015.

[45] Negus SS, Billsy EJ, Do Carma GP, Stevenson GW. Rationale and methods for assessment of pain-depressed behavior in preclinical assays of pain and analgesia. Methods Mol Biol 2010;617:79–91.

[46] Okun A, DeFelice M, Eyde N, Ren J, Mercado R, King T, Porreca F. Painful diabetic neuropathy: a systematic review and meta-analysis. PAIN 2014;155:2461–70.

[47] Parent AJ, Beaudet N, Beaudry H, Bergeron J, Berube P, Drolet G, Sarret LM, Machin I, Preclinical Pain C, Mogil JS, Stohr T. Animal models and the prediction of efficacy in clinical trials of analgesic drugs: a critical appraisal and call for uniform reporting standards. PAIN 2008;139:243–7.

[48] Rice ASC, Morland R, Huang W, Currie GL, Sena ES, Macleod MR. Transparency in the reporting of in vivo pre-clinical pain research: the relevance and implications of the ARRIVE (Animal Research: reporting in Vivo Experiments) guidelines. Scand J Pain 2013;4:58–62.

[49] Rischi SH, Garner JP, Rue C, Kunert J, Wurbel H. Systematic variation improves reproducibility of animal experiments. Nat Methods 2010;7:167–8.

[50] Roeska K, Doods H, Arndt K, Treede RD, Ceci A. Anxiety-like behaviour in rats with mononeuropathy is reduced by the analgesic drugs morphine and gabapentin. PAIN 2008;139:349–57.

[51] Rutten K, Arndt S, Read SJ, Christoph T. Pharmacological validation of a refined burrowing paradigm for prediction of analgesic efficacy in a rat model of sub-chronic knee joint inflammation. Eur J Pain 2014;18:213–22.

[52] Rutten K, Schiene K, Robens A, Leipelt A, Pasqualon T, Read SJ, Christoph T. Burrowing as a non-reflex behavioural readout for analgesic action in a rat model of sub-chronic knee joint inflammation. Eur J Pain 2014;18:204–12.

[53] Schaap MW, van Oostrom H, Doornenalb A, van ’t Klooster J, Baars AM, Arndt SS, Hellebrekers LJ. Nicotocision and conditioned fear in rats: strains matter. PLoS One 2013;8:e83339.

[54] Seretny M, Curie GL, Sena ES, Rammarine S, Grant R, Macleod MR, Colina LA, Fallon M. Incidence, prevalence, and predictors of chemotherapy-induced peripheral neuropathy: a systematic review and meta-analysis. PAIN 2014;155:2461–70.

[55] Sindrup SH, Graf A, Sifkas N. The N1K1-receptor antagonist TKA731 in painful diabetic neuropathy: a randomised, controlled trial. Eur J Pain 2006;10:567–71.

[56] Sorge RE, Maplebeck JC, Rosen S, Beggs S, Taves S, Alexander JK, Martin LJ, Austin JS, Sotocinal SG, Chen D, Yang M, Shi XQ, Huang H, Pillon NJ, Balian PJ, Tu Y, Klip A, Ji RR, Zhang J, Salter MW, Mogil JS. Different immune cells mediate mechanical pain hypersensitivity in male and female mice. Nat Neurosci 2015;18:1081–3.

[57] Sorge RE, Martin LJ, Labeser KA, Sotocinal SG, Rosen S, Tulle AH, Wileskopf JS, Acland EL, Dokova A, Kadoura B, Leger P, Maplebeck JC, McPhail M, Delaney A, Wigerblad G, Schumann AP, Quinn T, Frasnelli J, Svensson CI, Sterberg WF, Mogil JS. Oblatory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods 2014;11:629–32.

[58] Spruijt BM, Peters SM, de Heer RC, Pothenuis HJJ, van der Harst JE. Reproducibility and relevance of future behavioral sciences should benefit from a cross fertilization of past recommendations and today’s technology: “Back to the future”. J Neurosci Methods 2014;234:2–12.

[59] Thut PD, Hermanstye TO, Flake NM, Gold MS. An operant conditioning model to assess changes in feeding behavior associated with tempomandibular joint inflammation in the rat. J Orofac Pain 2010;24:166–70.

[60] Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Thut PD, Hermanstye TO, Flake NM, Gold MS. An operant conditioning model to assess changes in feeding behavior associated with tempomandibular joint inflammation in the rat. J Orofac Pain 2010;24:166–70.

[61] Turner L, Shamsseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz FK, Platt AC, Moher D. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev 2012;11:MR00003.

[62] Wallace VC, Blackhead J, Pheby T, Segerdahl AR, Davies M, Hasnie F, Hall S, McMahon SB, Rice AS. Pharmacological, behavioural and mechanistic analysis of HIV-1 gp120-induced painful neuropathy. PAIN 2007;133:47–63.

[63] Wallace VC, Blackhead J, Segerdahl AR, Hasnie F, Pheby T, McMahon SB, Rice AS. Characterization of rodent models of HIV-1 gp120 and anti-retroviral-associated neuropathic pain. Brain 2007;130:9688–702.

[64] Wallace VC, Segerdahl AR, Blackhead J, Pheby T, Rice AS. Anxiety-like behaviour is attenuated by gabapentin, morphine and diazepam in a rodent model of HIV anti-retroviral-associated neuropathic pain. Neurosci Lett 2008;448:153–6.

[65] Whittaker AL, Lynn KA, Nicholson A, Howarth GS. The assessment of good general well-being using spontaneous burrowing behaviour in a short-term model of chemotherapy-induced mucositis in the rat. Lab Anim 2015;49:30–9.

[66] Woolf CJ. Overcoming obstacles to developing new analgesics. Nat Med 2010;16:1241–7.

[67] Zimmermann M. Ethical guidelines for investigations of experimental pain in conscious animals. PAIN 1983;16:109–10.