Aus der Medizinischen Klinik mit Schwerpunkt Rheumatologie und klinische Immunologie der Charité – Universitätsmedizin Berlin

Eingereicht über das Institut für Tierschutz,
Tierverhalten und Versuchstierkunde
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Pain assessment and management in mouse femoral fracture models

Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin

vorgelegt von
Angélique Wolter
Tierärztin aus Stahnsdorf

Berlin 2024
Journal-Nr.: 4437

Aus der Medizinischen Klinik mit Schwerpunkt Rheumatologie und klinische Immunologie der Charité – Universitätsmedizin Berlin

Eingereicht über das Institut für Tierschutz, Tierverhalten und Versuchstierkunde des Fachbereichs Veterinärmedizin der Freien Universität Berlin

Pain assessment and management in mouse femoral fracture models

Inaugural-Dissertation

zur Erlangung des Grades eines

Doktors der Veterinärmedizin

an der

Freien Universität Berlin

vorgelegt von

Angélique Wolter

Tierärztin

aus Stahnsdorf

Berlin 2024

Journal-Nr.: 4437

Gedruckt mit Genehmigung

des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Uwe Rösler
Erste Gutachterin:	Univ.-Prof. Dr. Christa Thöne-Reineke
Zweiter Gutachter:	Univ.-Prof. Dr. Robert Klopfleisch
Dritter Gutachter:	Univ.-Prof. Dr. Lars Lewejohann

Deskriptoren (nach CAB-Thesaurus):
mice, animal models, bone fractures, pain, animal welfare, animal behaviour

Tag der Promotion**:** 13.02.2024

Table of content

List of figures

Abbreviations

3D	Three dimensional
ARRIVE	Animal Research: Reporting of In Vivo Experiments
BfR	Bundesinstitut für Risikobewertung
BtMG	Betäubungsmittelgesetz
BUP	Buprenorphine
DW	Drinking water
EU	European Union
FDA	U.S. Food and Drug Administration
GV-SOLAS	German Society of Laboratory Animal Science
Ex-vivo µCT	Ex-vivo micro computed tomography
i.p.	Intraperitoneal
MGS	Mouse Grimace Scale
NSAID	Nonsteroidal anti-inflammatory drug
PLGA	Poly-lactic-co-glycolic acid
PREPARE	Planning Research and Experimental Procedures on Animals: Recommendations for Excellence
PRISMA	Preferred Reported Items for Systematic Review and Meta-analyses
s.c.	Subcutaneous
SPF	Specific-pathogen-free
SYRCLE	SYstematic Review Center for Laboratory animal Experimentation
TierSchG	Tierschutzgesetz
TierSchVersV	Tierschutz-Versuchstierverordnung
USA	United States of America

1 Introduction

In 1959, Russell and Burch first described the 3R principle (*Replace*, *Reduce* and *Refine*) as a framework of experimental scientific work in *The Principles of Humane Experimental Technique* (Russell and Burch 1959). According to current legal standards, alternative methods should replace animal experiments whenever possible (*Replacement)*, the number of animals used should be decreased (*Reduction*) and the harm, suffering, or pain for the used animals should be reduced, and their well-being increased (*Refinement*).

In *The Principles of Humane Experimental Technique*, Russell and Burch introduce Chapter 7, which is focusing on *Refinement* as following: *"…Suppose, for a particular purpose, we cannot use replacing techniques. Suppose it is agreed that we shall be using every device of theory and practice to reduce to a minimum the number of animals we have to employ. It is at this point that refinement starts, and its object is simply to reduce to an absolute minimum the amount of distress imposed on those animals that are still used…" (Russell and Burch 1959).* Therefore, they emphasize that if *Replacement* and *Reduction* options are exhausted, *Refinement* measures must be implemented to ensure the avoidance of any inhumanity e.g., suffering of the animal. Although the initial definition of *Refinement* by Russell and Burch seems to be mainly focused on the reduction of distress, today's interpretation of the term is rather broader, also including the general improvement of animal welfare and well-being and the minimization or elimination of stress, pain and suffering without compromising scientific objectives. Exemplary *Refinement* measures comprise, e.g., non-invasive handling measures, enrichment to improve housing and husbandry, minimally invasive procedures and evidence-based analgesic regimens for sufficient pain coverage and avoidance of suffering.

Even though the use of analgesia in animal experiments is not a *Refinement* measure but is required by European and German law, the analgesia itself can be adequate or inadequate. The implementation of model-adapted, stress-free, pre-, intra and post-operative analgesic protocols can, therefore, be considered as *Refinement*. Adequate pain management in laboratory animals is essential for ethical but also for scientific reasons, as insufficiently treated pain can alter mechanical, physiological, and immunological responses and thereby increase the variability of the gained data (Page 2003; Stasiak et al. 2003; Jirkof 2017). To assess whether pain management is adequate or not, painful states must be reliably detected. However, this remains a great challenge in (laboratory) animals, and pain assessment is largely based on the absence or presence of clinical signs that are indicative of pain or reduced well-being. Pain assessment in rodents is further complicated by the fact that these animals can hide signs of disease and weakness for long periods of time. Clear changes often only become noticeable when the animal is already experiencing greater pain (Stasiak et al. 2003;

Arras et al. 2007; Miller and Leach 2015). Thus, some indicators may not be sensitive enough to detect mild or even moderate pain in laboratory rodents and it is recommended to combine multiple general and model-specific parameters (Turner et al. 2019). Pain assessment is further complicated by the fact that anesthesia and analgesia can also affect animal behavior and indicators suggestive of pain. Therefore, distinguishing between anesthesia and analgesia associated effects and actual (surgery-related) pain can be difficult (Hohlbaum et al. 2018; Oliver et al. 2018). The ongoing critical assessment and continued development of parameters for the assessment of pain and reduced well-being in laboratory animals is, therefore, of great importance to further improve animal welfare in animal-based research (Jirkof 2017).

Over the past decade, different guidelines have been proposed to enhance the planning and reporting in animal experimentation. Pain assessment and management in laboratory rodents has also been the subject of many research efforts. However, challenges in animal-based research remain regarding adequate reporting and appropriate pain assessment and management (Leung et al. 2018; Jirkof et al. 2019b; Tihanyi et al. 2019). This is also true for bone fracture-related research involving animals. Despite their small size and differences in general bone structure compared to humans, mice are often used in fracture healing studies due to multiple other advantages. In fracture healing studies, pain management is mostly conducted through opioid analgesia (Lang et al. 2016; Jirkof et al. 2019a). However, as the half-life of the used opioids is relatively short, reapplications must take place frequently. Other possibilities are oral application routes (e.g., via the drinking water) or the use of sustained-release formulations, that ensure analgesic efficacy for a longer period post-injection. A sustained-release opioid formulation could, therefore, help to expand the possibilities for pain management in mice in Europe greatly. However, no such product is currently available on the European market.

In this work, the status quo of the reporting accuracy on general and model-specific parameters as well as on commonly used analgesia protocols in studies using mouse femoral fracture models was assessed. Subsequently, the analgesic efficacy of an experimental sustained-release buprenorphine formulation for prolonged post-surgical analgesia was evaluated in two mouse osteotomy models. Furthermore, it was investigated whether the analyses of rear-up behavior could add value as an additional model-specific indicator for pain in mouse osteotomy models.

2 Literature review

2.1 The 3R principle

In 1959, Russell and Burch described the 3R principle (*Replacement, Reduction* and *Refinement*) as a framework for animal-based experimental scientific work in *The Principles of Humane Experimental Technique* (Russell and Burch 1959). While the 3R principle initially received little recognition, it has gradually become widely established (Flecknell 2002) and is now an integral component in German animal experiments with manifestation in the European (Directive 2010/63 EU) and national legislation (Tierschg ; Tierschversv). Russell and Burch originally defined the meaning of *Replacement* as *"…the substitution for conscious living higher animals of insentient material…"*, *Reduction* as *"…the reduction in the numbers of animals used to obtain information of a given amount and precision…"* and *Refinement* as *"…any decrease in the incidence or severity of inhumane procedures applied to those animals which still have to be used…"* (Russell and Burch 1959). Over time, the definition of the 3Rs has been modified, based on different interpretations. Today's understanding of the 3Rs, therefore, deviates in part from the original definitions from 1959 (Schuppli et al. 2004; Tannenbaum and Bennett 2015). Nowadays, the term *Replacement* often relates to the use of less sentient species or the use of alternative, non-animal methods (NAMs) excluding the use of cells, tissue or other components generated from animals (Tannenbaum and Bennett 2015). Even though today's use of *in-vitro* and computer-based *in-silico* methods are promising to replace specific animal experiments, *Replacement* is currently proving difficult or not feasible, including research areas like fundamental research, regulatory toxicology, research on animals for animals and others (Flecknell 2002; Benfenati et al. 2010; Madden et al. 2020). In terms of *Reduction*, the number of animals used in an experiment should be reduced to the necessary minimum by optimizing the study design and performing precise sample size calculations to ensure the generation of valid data and avoid the need to repeat animal experiments (Flecknell 2002; Smith et al. 2018). In the guide for the care and use of laboratory animals, *Reduction* is defined as *"…strategies for obtaining comparable levels of information from the use of fewer animals or for maximizing the information obtained from a given number of animals (without increasing pain or distress) so that in the long run fewer animals are needed to acquire the same scientific information…"* (National Research Council 2011). Additionally, combining alternative methods and animal experiments can help to reduce the animal use while maximizing translation (Chapman et al. 2013). Nowadays, the term *Refinement* often refers to the enhancement of the animal's well-being and the minimization or elimination of distress and pain experienced by the animals that are still used in experiments (National Research Council 2011). This can be achieved through adaptations and improvements, e.g., in husbandry, care, handling, anesthesia, analgesia, and postoperative care (Flecknell 2002; Goecke et al. 2005).

Although the 3Rs are implemented through European and German law, Franco et al. showed that in 2014/ 2015 just over half of the questioned researchers that worked with animals, or planned to, were aware of the 3Rs before attending a training in laboratory animal science (Franco et al. 2018). The implementation and possible expansion of the 3Rs are, furthermore, a continuous discourse within the scientific community (Strech and Dirnagl 2019). In detail, some authors have proposed to extend the 3Rs by further Rs such as *Responsibility* or *Remembering* (Iliff 2002; Klein and Bayne 2007). In 2019, Strech and Dirnagl suggested a 6R principle for ethical animal research, that complements the 3R principle with the three additional principles of *Robustness*, *Registration* and *Reporting* to increase the validity, reproducibility and scientific value of animal experiments (Strech and Dirnagl 2019).

2.2 Refinement in animal experiments with specific focus on mice

In the European Union, animal research is regulated and harmonized under Directive 2010/63/10 of the European Parliament and of the council on the protection of animals used for scientific purposes. In 2013, the legal requirements were implemented into national law in each EU Member State leading to the adaptation of the following German legal acts:

- Animal Rights Law [Tierschutzgesetz (TierSchG)].
- Regulation for the protection of animals used in animal experiments or for other scientific purposes [Verordnung zum Schutz von zu Versuchszwecken oder zu anderen wissenschaftlichen Zwecken verwendeten Tieren (TierSchVersV)]

These regulations provide a legally binding framework to implement suitable *Refinement* measures. Therefore, it is advised to evaluate potential *Refinement* measures that can help to improve animal welfare and well-being, without compromising scientific objectives, before conducting experiments that may cause some degree of discomfort, pain, stress or suffering to the used animals (Auer et al. 2007; Hawkins et al. 2015). The identification and selection of suitable *Refinement* measures depends on different factors, such as species used, specific animal-models, scientific objectives, and experimental readouts. Some *Refinement* measures can be applied to most species used in animal experiments; others, like certain handling techniques, are more species-dependent. Hawkins et al. also suggests that the used *Refinement* approaches should always consist of proactive and evidence-based measures as opposed to a reactive management (Hawkins et al. 2015).

General *Refinement* measures in rodents, for example, include acclimation periods after transportation and the provision of adequate time to habituate the animals to their new environment and group compositions before the start of the experiment (Tuli et al. 1995; Conour et al. 2006; Obernier and Baldwin 2006; Auer et al. 2007; Arras et al. 2020). Improvements in housing and care, such as effortless access to food and water in the form of (wet) food or fluid gels on the floor after health-compromising interventions, are other examples

for general *Refinement* in rodents (Arras et al. 2020). Captive and non-enriched environments often limit or prevent the expression of natural behaviors and may thereby lead to boredom, frustration, (di)stress, depression, and behaviors indicative of poor animal welfare such as abnormal repetitive behaviors (Würbel and Garner 2007; Mieske et al. 2022; Ratuski and Weary 2022). In a systematic review on the impact of environmental enrichment on the welfare of laboratory mice and rats, Mieske and colleagues showed that both species benefit from enriched living environments (Mieske et al. 2022). However, the term enriched housing has changed greatly over the last decades and the concept of enrichment remains a relatively vague notion (Newberry 1995; Mieske et al. 2022). There are various interpretations and limited evidence on the actual impact of enrichment opportunities on animal welfare from the animals' perspective (Newberry 1995; Hobbiesiefken et al. 2021; Mieske et al. 2022; Ratuski and Weary 2022). It has also been shown that training of the animals and the use of non-invasive handling techniques such as tunnel handling or cupping instead of tail-handling in mice can help to minimize anxiety and reduce handling-induced stress, while simultaneously improving their performance in behavioral tests (Conour et al. 2006; Hurst and West 2010; Rasmussen et al. 2011; Gouveia and Hurst 2013; Ghosal et al. 2015; Gouveia and Hurst 2017; Ueno et al. 2020). The use of model-specific score sheets, the strict definition of humane endpoints and the identification of indicators for pain or reduced well-being further support the *Refinement* of animal-based studies (Schuppli et al. 2004; Hawkins et al. 2015; Lang et al. 2016). More in detail, general guidelines suggest that score sheets should include multiple robust species-specific and model-specific clinical (pain) parameters, as well as general indicators of well-being, to ensure an objective severity assessment (Arras et al. 2020). Moreover, recommendations further indicate that (surgical) procedures should be carried out as minimally invasive as possible. A thoughtfully planned pre-, intra and post-operative pain management as well as adequate anesthesia protocols also help to reduce pain and suffering (Auer et al. 2007; Jirkof 2017; Arras et al. 2020). Even though analgesia in animal studies is not a *Refinement* measure but is required by law (TierSchG; TierSchVersV; European Parliament and the Council of the European Union 2010), the use of model-adapted, stress-free, and thoughtfully planned pre-, intra and post-operative (multimodal) analgesic protocols can be considered a further *Refinement* measure.

2.3 Pain assessment and management in laboratory mice

According to the German legislation, pain-relieving substances or procedures are prescribed in experiments on vertebrates and cephalopods to ensure the avoidance and reduction of pain and suffering to the lowest possible level (§ 17 Schmerzlinderung und Betäubung (TierSchVersV)), while pain, suffering and harm inflicted on the animals must be limited to an indispensable extent (§ 7 (1) (TierSchG)). It has also been shown that untreated or

inadequately treated pain can alter different mechanical, physiological and immunological mechanisms and could, therefore, influence the overall health and responsiveness of the animals as well as the scientific results (Morton and Griffiths 1985; Page 2003; Stasiak et al. 2003; Jirkof 2017; Peterson et al. 2017). Thus, adequate pain management is essential not solely for ethical, humane and legal reasons, but also for scientific considerations (Jirkof 2017). Nevertheless, the concern that the use of certain analgesics may interfere with the used animal model is an ongoing concern in the scientific community (Jirkof 2017; Mogil et al. 2020).

To evaluate whether a pain management protocol is adequate or not, a reliable identification and evaluation of the occurring pain is of utmost importance. However, pain assessment in laboratory animals is an ongoing challenge and can be highly species-specific (Stasiak et al. 2003). No gold standard for assessing pain in animals exists and as pain in animals cannot be measured directly, pain assessment in laboratory rodents is built around tools to observe indicators of pain and (reduced) well-being (Auer et al. 2007; Turner et al. 2019). Factors such as strain, sex, age, and various environmental stressors are known to influence pain perception and response to analgesic compounds in mice (Mogil 1999; Auer et al. 2007; Bodnar and Kest 2010; Jirkof 2017; Smith 2019). As prey animals, mice tend to conceal signs of disease or weakness for as long as possible. Thus, it has been reported that mild or moderate pain is harder to observe, and clear changes often only become noticeable when the animal is already experiencing greater pain (Stasiak et al. 2003; Arras et al. 2007). Miller and Leach also described the ability of mice to hide behaviors like grimacing when in sight of an observer (Miller and Leach 2015), which makes the pain assessment in mice even more difficult. Moreover, several studies have indicated that routine handling of mice is also often associated with additional stress for the animals, which can increase or decrease the perception of pain, possibly resulting in stress-induced analgesia (Watkins and Mayer 1982; Amit and Galina 1988; Carstens and Moberg 2000). Therefore, different authors suggest that the sole use of very simple assessment methods may not be sensitive enough to pick up on low to moderate pain and that different measures should be combined in a composite pain measurement scheme for the most reliable pain assessment (Jirkof et al. 2019b; Turner et al. 2019; Aulehner et al. 2022).

2.3.1 Pain assessment: clinical, behavioral, and physiological measurements

For decades, algesiometric assays, such as the tail-flick test, hot plate test, paw withdrawal test and von Frey test have been widely used to assess pain and analgesic efficacy in mice (Baumans 1994; Mogil 2009; Barrot 2012; Deuis et al. 2017; Turner et al. 2019). These tests assess nociception and evoke response reflexes following a noxious mechanical, thermal, or chemical stimulus (Turner et al. 2011). However, nociception and pain are not synonymous and can each occur without the other (Loeser and Treede 2008; Raja et al. 2020). It has

therefore been proposed that these tools might represent an oversimplification when assessing pain and analgesic efficacy (Turner et al. 2019).

Thus, several clinical and behavioral parameters have been put forward to provide a more reliable pain assessment in mice. Those measurements are largely based on the absence or presence of clinical signs indicative of pain and include the assessment of deviations from normal behavior, activity and appearance, the analyses of food and water intake as well as body weight developments (Jirkof 2017; Turner et al. 2019). Other commonly used indicators comprise the analyses of exploratory behaviors and nesting behavior, the analyses of normal and abnormal behavior and the assessment of the mouse grimace scale or a composite score (Stasiak et al. 2003; Arras et al. 2007; Wright-Williams et al. 2007; Langford et al. 2010; Jirkof 2017; Jirkof et al. 2019a; Turner et al. 2019). When trying to assess pain or reduced well-being, the data collection of baselines (naïve animals) has also shown to be essential to identify deviations from baseline values and behavior (Kilkenny et al. 2010; Langford et al. 2010; Hawkins et al. 2015).

Furthermore, many studies report the effect of anesthesia and analgesia itself on parameters used for post-operative/-procedural pain assessment (Liles and Flecknell 1992; Goecke et al. 2005; Langford et al. 2010; Miller and Leach 2015; Hohlbaum et al. 2018; Oliver et al. 2018). In detail, indicators such as food and water intake, and consequently body weight, are not only negatively affected by pain but can also be affected by analgesia and/or anesthesia (Liles and Flecknell 1992; Goecke et al. 2005; Oliver et al. 2018). The mouse grimace scale (MGS), first described by Langford et al. has been shown to be a helpful tool to measure pain through changes in facial expression and focuses on orbital tightening, bulging of the nose and cheeks, as well as the positions of the ears and whiskers (Langford et al. 2010). Yet, anesthesia, analgesia, drugs, sex, and strain have also been reported to affect the mouse grimace scale (Miller et al. 2015; Hohlbaum et al. 2018; Mogil et al. 2020). Moreover, Miller and Leach showed differences in the MGS scoring between live and retrospective scoring (Miller and Leach 2015). Alternatively, composite scores have been proven to be reliable tools to assess pain and reduced well-being in mice (Arras et al. 2007; Jirkof et al. 2015; Jirkof et al. 2019a). The composite score combines facial expressions and overall appearance of the animals, for instance posture and coat condition, and can be adapted to the used experimental model (Arras et al. 2007; Jirkof et al. 2019a). However, the composite score can also be affected by anesthesia- and analgesia-associated effects (Jirkof et al. 2019a; Wolter et al. 2023a). In addition, different studies have described that pain can lead to decreased activity and decreased normal, mostly exploratory behaviors, such as rearing, digging, and sniffing as well as increased abnormal behaviors (Wright-Williams et al. 2007; Miller et al. 2012). Evidence from the literature suggests that locomotor activity and behavior can also be affected by drugs

including different analgesics, for example opioids (Kuribara and Tadokoro 1991; Jirkof et al. 2015; Sauer et al. 2016; Oliver et al. 2018). The analyses of other intrinsically motivated behaviors such as burrowing, nest building and the time-to-integrate-into-nest-test have also been successfully used as valuable tools to identify reduced well-being and possibly pain, as mice experiencing pain are less inclined to burrow or build and maintain proper nests (Gaskill et al. 2013; Jirkof et al. 2013b; Rock et al. 2014; Oliver et al. 2018; Turner et al. 2019). Nevertheless, anesthesia–analgesia-associated effects on nesting in the absence of surgical pain have been reported by different authors (Jirkof et al. 2013b; Oliver et al. 2018). The value and use of grimace scales and behavioral parameters such as burrowing, and nest building have been described by various studies for the detection of pain. Nevertheless, the still mostly unknown validity, robustness and reliability as well as the subjectiveness, diverging interobserver reliability and the multitude of confounding factors of these parameters highlight the need for the combination of multiple assessments for an adequate pain assessment (Turner et al. 2019; Aulehner et al. 2022).

Today, different manual and automatic systems can be used to measure the frequency and/or the duration of certain behaviors, movements and general activity as well as other pain-related indicators (Roughan et al. 2009; Tappe-Theodor et al. 2019; Grieco et al. 2021; Segalin et al. 2021; Klein et al. 2022). While manual assessment is highly time-consuming and only allows for short observations of a few minutes, automatic assessments require less expenditure of time and allow for distinct longer observations (Roughan et al. 2009; Klein et al. 2022). Therefore, Roughan et al. suggests that the highly time-consuming nature of manual analysis leads to the assessment of merely a limited assortment of behaviors in short periods (Roughan et al. 2009). Depending on the set-up, analyses can take place in the home cage, through different home cage observation systems or in designated observation cages, boxes, or apparatuses (Grieco et al. 2021; Klein et al. 2022). Klein et al. suggests that compared to observations outside the home cage, the non-invasive home cage analyses might increase standardization and reduce influences of handling and other stress inducing factors (Klein et al. 2022).

As pain has effects on the endocrine-, cardiovascular-, sympathetic nervous-, and digestive system, it thereby also leads to physiological and biochemical changes (Morton and Griffiths 1985; Baumans 1994; Stasiak et al. 2003). The analyses of physiological parameters such as heart and respiratory rate, blood pressure and body temperature have, therefore, been successfully used in several studies to identify pain (Baumans 1994; Carstens and Moberg 2000; Stasiak et al. 2003; Jirkof 2017). However, Baumans et al. reported that changes in those visceral reactions are not a sole evidence of pain and can also be induced by stress or nociceptive stimuli during deep anesthesia (Baumans 1994). While the analyses of the heart

rate and heart rate variability may be able to detect mild or moderate pain that otherwise could be missed by more simplistic observation methods, a greater and more invasive effort is often needed to assess those parameters (Arras et al. 2007).

2.3.2 Pain management in laboratory mice

As elaborated above, adequate pain management is important for ethical and legal reasons, but also for scientific reasons. Several studies indicate that insufficiently treated pain can affect and alter different mechanical, physiological and immunological responses, which influence data and result in increased variability and reduced reproducibility (Morton and Griffiths 1985; Page 2003; Stasiak et al. 2003; Jirkof 2017; Peterson et al. 2017). Nevertheless, pain-relieving drugs themselves may also bear the potential to modify experimental data (Carbone and Austin 2016; Jirkof 2017). The adequate choice and reporting of the analgesic protocol is therefore of great importance. Different organizations such as the GV-SOLAS and the American College of Laboratory Animal Medicine provide resources on pain management recommendations (Kohn et al. 2007; Carbone 2012; Arras et al. 2020). However, the use and reporting quality of different pain management protocols in animal-based studies still varies markedly (Jirkof et al. 2015; Carbone and Austin 2016; Jirkof 2017).

When conducting surgical interventions or causing other potentially painful injuries or conditions in animal-based studies, an adapted preemptive, intraoperative, and postoperative analgesic treatment should ensure that the expected period of greatest pain is covered. Analgesics such as opioids or non-steroid anti-inflammatory drugs (NSAIDs) are often employed to manage and prevent pain in laboratory rodents (Jirkof 2017). Based on general guidelines, the adequate pain treatment protocol, including the selection of suitable analgesics, the dose, application route, interval and other possible accompanying pain-reducing measures must be thoughtfully planned out before conducting an animal experiment (Arras et al. 2020). Thereby, factors such as the species, strain, interventions and expected level and duration of pain, the research question, study design and experimental read-out parameters are of importance and guide the choice of the appropriate analgesic regimes (Jirkof 2017; Arras et al. 2020). In a review, Smith also summarized occurring strain and sex disparities in pain response and analgesic sensitivity in mice (Smith 2019). Although standardized pain management protocols in animal-based studies are of great value, adequate pain management, therefore, still requires constant re-evaluation and potentially adjustments in future experiments (Jirkof 2017).

Systemic application routes of analgesics include subcutaneous-, intraperitoneal-, intravenous-, intramuscular injections as well as oral und transmucosal ways of application (Arras et al. 2020). In laboratory rodents, subcutaneous and intraperitoneal injections as well

as oral applications are the most commonly used routes for applying analgesics. The advantage of injecting analgesics is the supervised administration in controllable intervals and dosages. However, depending on the half-life of the drug, injections must be repeated in frequent intervals to maintain adequate analgesic efficacy. Several studies have shown that the increased handling and fixation for repeated injections are associated with increased stress for the animals (Balcombe et al. 2004; Hurst and West 2010), which can result in additional pain and can further influence physiological and behavioral parameters (Jirkof et al. 2015; Gouveia and Hurst 2017). Nevertheless, extended intervals between injections can bear the risk of periods with insufficient or no pain relief (Jirkof et al. 2015). Analgesics with a very short half-life are, therefore, proposed as not suitable for repeated injections and should alternatively be applied orally (Lang et al. 2016). A great benefit of orally applied analgesics is the reduction of animal handling. Yet, it has been described that the oral uptake frequency of analgesics, e.g., with the drinking water, in gels or similar, can be negatively affected by the circadian rhythm, possible alterations of taste due to bitter drugs and reduced uptake after anesthesia and/or surgery. Moreover, the first-pass effect may reduce the overall bioavailability and final serum concentration of the analgesics, and the voluntary uptake hinders the individual monitoring of specific analgesic levels, as uptake and timing cannot be controlled (Sauer et al. 2016; Arras et al. 2020). Another alternative to repeated analgesic injections is the usage of subcutaneously implanted micropumps or the injection of sustained-release/depot preparations that provide a continuous pain relief without additional handling of the animals (Jirkof 2017; Arras et al. 2020). Micropumps can be osmotic, or battery-powered and provide constant levels of analgesics or other drugs over a continuous period (Eckenhoff and Yum 1981; Verma et al. 2004; Keraliya et al. 2012; Meng and Hoang 2012; Arras et al. 2020). Parenteral sustained-release/depot preparations slowly release the analgesic over a longer period and can, therefore, provide prolonged analgesia for 24h to 72h (Kendall et al. 2016; Schreiner et al. 2020; Myers et al. 2021). To date, no analgesic sustained-release preparation for mice or rats is commercially available on the European market. However, in the United States of America (USA), different analgesics, for example buprenorphine, already exist as sustained-release/depot preparations for mice and/or rats and are marketed as unapproved but indexed animal drugs for minor species ("*The Index of Legally Marketed Unapproved New Animal Drugs for Minor Species*") (Arras et al. 2020; Schreiner et al. 2020; Food and Drug Administration 2022). To be legally marketed, new animal drugs must undergo pre-market review by the Food and Drug Administration (FDA) and obtain a legal marketing status through an approval, a conditional approval, or an index listing (Food and Drug Administration 2019). However, due to their marketing status, those preparations are not commercially available in Europe (Jirkof et al. 2015; Schreiner et al. 2020). As the absence of analgesic sustained-release formulations on the European market hinders the road to a more stress-free pain

management for many laboratory rodents, different experimental formulations of sustained-release buprenorphine have been developed and tested in Europe in recent years (Jirkof et al. 2015; Schreiner et al. 2020; Schreiner et al. 2021).

The use of multimodal analgesia that combines different analgesics or local anesthetics is a further measure to improve analgesic protocols (National Research Council 2009; Oliver et al. 2018). The National Research Council Committee on Recognition and Alleviation of Pain in Laboratory Animals recommends multimodal analgesia for severely painful procedures; but it can also be considered for treating moderate pain (National Research Council 2009). Although different recommendations for additional local anesthesia exist, their use in e.g., rodents is only scarcely reported and further systematic studies on their benefits and side effects are needed (Durst et al. 2021b). When additionally using local anesthesia in a mouse laparotomy model, Durst et al. did not observe indicators of improved pain management or recovery. However, they also did not observe adverse side effects, concluding that it may be advisable to integrate local anesthesia in future studies (Durst et al. 2021b). Even though Oliver et al. identified significant drug effects of their used multimodal analgesia (combination of buprenorphine s.c. and carprofen in the drinking water) throughout a 48h time period after anesthesia and analgesia, they also found that the used multimodal analgesic regime had the highest analgesic coverage compared to the used single-class analgesics (Oliver et al. 2018).

In addition to analgesics, non-pharmacological measures should accompany the analgesic treatment for better pre-, intra-, and postoperative pain management (Arras et al. 2020). Those measures are versatile and can be adapted to the respective animal experiment. Examples include adequate acclimatization periods, the use of non-invasive handling techniques like tunnel handling or cupping for mice, and training of the animals prior to the experiment (Conour et al. 2006; Obernier and Baldwin 2006; Hurst and West 2010; Gouveia and Hurst 2013), atraumatic or minimal invasive surgical methods by experienced staff, adequate pre-, intra-, and postoperative care and thermal management and adjusted postoperative feeding methods after health-compromising interventions (Arras et al. 2020). To further reduce stress, it is also recommended to let the animals wake up in familiar surroundings and to not subject them to a freshly changed cage directly after surgery (Jirkof et al. 2013a; Jirkof 2015; Arras et al. 2020). Cage enrichment that could potentially lead to self-inflicted injuries should be removed from the cages and reinstalled at given times.

2.4 Research on fracture healing

Fractures commonly occur in humans and animals. Despite the fact that bone has the particular capability to regenerate with no scarring, up to 5-10% of the human patients suffer from bone healing disorders and insufficient fracture healing outcomes (Haas 2000; Dimitriou

et al. 2005; Schmidt-Bleek et al. 2014; Peric et al. 2015; Pfeiffenberger et al. 2021). Those adverse outcomes and decelerated repair can be a significant burden to the patient's life and thereby reduce their quality of life (Pfeiffenberger et al. 2021). Although different *ex-vivo* and *in-vitro* methods can be used to delineate the underlying mechanisms of bone regeneration, comorbidities, and risk factors as well as to develop potential new treatment strategies (Bigham Sadegh and Oryan 2014; Schmidt-Bleek et al. 2014; Lang et al. 2016; Haffner-Luntzer et al. 2019; Pfeiffenberger et al. 2021), animal models still remain the gold standard in fracture healing research today (Auer et al. 2007; Pfeiffenberger et al. 2021).

2.4.1 Fundamental process of fracture healing

Fractures can be classified as closed fractures with no skin break, or as open fractures, including a skin break, with more severe damage to the surrounding tissues and periosteum (Bigham Sadegh and Oryan 2014). Open fractures usually have a higher risk of infection and non-union formation, thereby leading to a higher probability for insufficient fracture healing outcomes (Bigham Sadegh and Oryan 2014). Risk factors and comorbidities, such as metabolic disorders, osteoporosis, ischemia, disturbed vascularization, and injuries to the surrounding soft tissue, can further impair the highly complex process of bone regeneration (Gaston and Simpson 2007; Haffner-Luntzer et al. 2019).

Fracture healing can occur through direct (primary) or through the more common indirect (secondary) bone repair, depending on inter-fragmentary stability and fixation stiffness (Auer et al. 2007; Bigham Sadegh and Oryan 2014). Realigned fracture ends, high inter-fragmentary stability, and compression through a rigid fixation, for example with plates, encourages direct bone repair (intramembranous bone formation) whereas fixation with external fixators or intramedullary nails and higher inter-fragmentary movement lead to indirect bone repair via endochondral ossification (Phillips 2005; Auer et al. 2007; Schmidt-Bleek et al. 2014). Thus, direct bone healing results in no or only minimal callus formation, while indirect healing commonly leads to a more prominent callus formation (Phillips 2005; Auer et al. 2007).

Secondary fracture healing by endochondral bone regeneration can be categorized into four stages that occur successively and overlap to certain extents (Fig. 1) (Bigham Sadegh and Oryan 2014; Einhorn and Gerstenfeld 2015). Inflammatory processes and hypoxia are followed by the soft callus formation, the mineralization of the soft callus (endochondral ossification) and lastly bone remodelling (Gerstenfeld et al. 2003; Bigham Sadegh and Oryan 2014; Morgan et al. 2014; Einhorn and Gerstenfeld 2015). The initial phase lasts 3-4 days with the formation of a hematoma in the fracture gap, which includes various inflammatory cell populations in an oxygen- and nutrient-deficient environment (Kolar et al. 2011; Bigham Sadegh and Oryan 2014; Morgan et al. 2014). This phase may be extended depending on the strength of the

traumatic event that caused the fracture (Bigham Sadegh and Oryan 2014). In the second phase, the hematoma is transformed into granulation tissue and a fibrocartilaginous callus is formed that gradually gets vascularized, driving further downstream bone formation (Bigham Sadegh and Oryan 2014; Morgan et al. 2014). Afterwards the fibrocartilaginous callus is resorbed and substituted by primary bone (endochondral ossification), beginning at the fracture end, and proceeding towards the center (Bigham Sadegh and Oryan 2014; Morgan et al. 2014). Lastly, the secondary bone formation takes place. The bone gets entirely remodeled, including the regeneration and reestablishment of the bone marrow space and its hematopoietic tissues (Gerstenfeld et al. 2003; Morgan et al. 2014; Einhorn and Gerstenfeld 2015). Compromised vascularization and/or excessive interfragmentary movements may lead to atrophic non-unions, hypertrophic non-unions or pseudoarthrosis (Bigham Sadegh and Oryan 2014). Lienau et al. also showed that increased interfragmentary shear movements in a tibial osteotomy in sheep fixed with a semirigid fixator were associated with a reduced initial vascularization and led to a less optimal fracture healing compared to the rigid fixated group (Lienau et al. 2005).

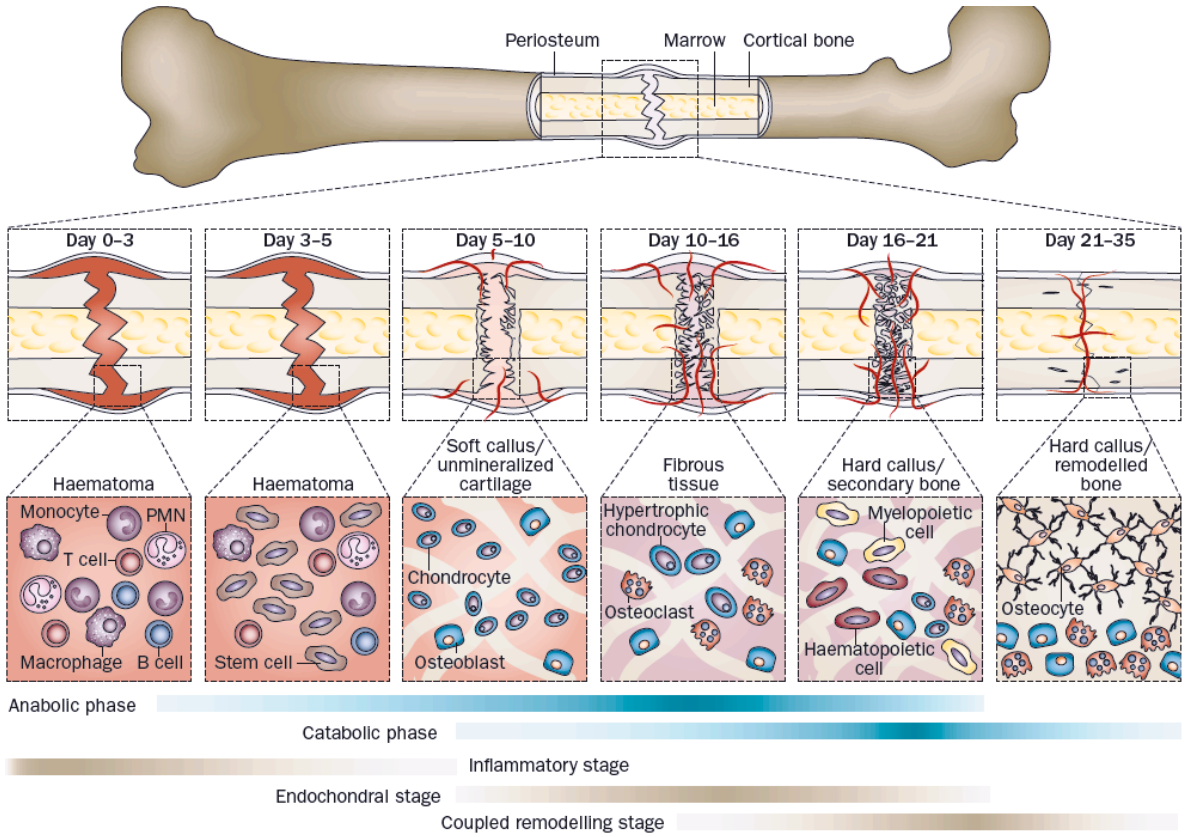

Figure 1 *Schematic depiction of the timeline of the metabolic biological stages during the process of fracture healing and the corresponding main cell populations. Modified and reproduced from Einhorn and Gerstenfeld 2015.*

2.4.2 Animal models in fracture healing studies

Different animal models are used to study mechanism of bone healing in the context of fundamental research questions or translational aspects including the testing of new therapeutic approaches such as medication and biomaterials (Auer et al. 2007; Lang et al. 2016; Pfeiffenberger et al. 2021). In order to answer different research questions, laboratory rodents, e.g., mice and rats, but also zebrafish, rabbits, dogs, sheep, goats, or pigs are often used (Aerssens et al. 1998; Martini et al. 2001; Bundesinstitut Für Risikobewertung 2021).

Analyzing the use of different species in orthopedic research between 1970 and 2001, Martini et al. showed that the greatest percentage of those animals were laboratory rodents (26% mice, 36% rats) (Martini et al. 2001). O'Loughlin et al. also examined studies using animal fracture models in 6 different orthopedic and musculoskeletal journals over 10 years and found that the use of rats (35%), rabbits (19%), and mice (15%) was most common (O'loughlin et al. 2008). Since then, the use of mouse models in bone healing research has continued to grow. According to the yearly report of the Bundesinstitut für Risikobewertung, a total of 15135 animals were used for basic research purposes on the musculoskeletal system in Germany in 2020. Of these, 11213 animals were mice, thereby distinctly being the most utilized species for this research purpose (Bundesinstitut Für Risikobewertung 2021). In terms of translational and applied research on human musculoskeletal disorders, a total of 1242 animals were used in 2020, with pigs, rats and mice being the three most used species (pigs: 420, rats: 352, mice: 350 animals) (Bundesinstitut Für Risikobewertung 2021). In 2019, the European Commission published the first statistical report on the use of animals for scientific purposes in the Member States of the European Union between 2015 and 2017 (The Commission to the European Parliament and the Council 2019). Out of the 9.58 million used animals (first use and any subsequent reuse) in the member states of the European Union in 2017, approximately 1.04% (~100.000) animals were used for basic research purposes on the musculoskeletal system and about 0.42% (~40.000) animals were utilized for translational and applied research on human musculoskeletal diseases (The Commission to the European Parliament and the Council 2019; Pfeiffenberger et al. 2021).

Mice are often used in bone healing research as they require less space and have lower acquisition, breeding and housing costs, are easier to handle, and genetically modified strains are available in large numbers (O'loughlin et al. 2008; Histing et al. 2011; Haffner-Luntzer et al. 2016; Grün et al. 2019). Mice also show an accelerated aging process due to a shortened life span when compared to other animals and a fast-breeding cycle. Mice and rats are often utilized to study molecular pathways and cellular signaling processes during fracture healing (Haffner-Luntzer et al. 2019). In contrast, advantages of larger animals comprise the larger body and bone size, similar loading characteristics, a higher body weight, similar bone

structures and the ability to use equipment also used in the clinic (Reinwald and Burr 2008; Grün et al. 2019). Dogs, for instance, have been shown to have a resembling bone composition to humans, including the existence of a Haversian system (Martin et al. 1981; Aerssens et al. 1998; Auer et al. 2007), while laboratory rodents have a more primitive bone structure, a remaining open growth plate throughout early adulthood that closes between the age of 14-16 weeks and no Haversian system (Gomes and Fernandes 2011; Histing et al. 2011).

2.4.3 Mouse models in fracture healing studies

Despite their small size, models of long bone fractures of the femur or tibia in mice are well established (Röntgen et al. 2010; Histing et al. 2011; Haffner-Luntzer et al. 2016; Lang et al. 2016). To study bone healing in mice, fracture models and osteotomy models can be distinguished. Fracture models are characterized by a mechanical break of the bone while an osteotomy refers to cutting the bone. Both models are likewise used to investigate the process of bone regeneration in preclinical studies. In detail, fracture models closer mimic the clinical circumstances and allow for a fracture without invasive surgical exposure of the bone (Klein and Bayne 2007). However, a disadvantage of this method is that methodological standardization is a great challenge and uncontrollable confounding factors can result in higher dropout rates (Klein and Bayne 2007). In contrast, osteotomies can be highly standardized and allow the creation of more precise gap sizes. On the other hand, osteotomies require more complex and invasive surgical techniques including the exposure of the corresponding bone (Klein et al. 2015; Dabis et al. 2017). The mode of fracture should, therefore, always be adapted to the research question.

To fixate long bone fracture- and osteotomy models in mice, intramedullary pins, external fixators, or plates are often used (Holstein et al. 2007; Gröngröft et al. 2009; Röntgen et al. 2010; Histing et al. 2011; Lang et al. 2016; Wolter et al. 2021). Further fixation systems include intramedullary compression screws, locking nails, pin-clip devices, and mouse nails (Fig. 2) (Histing et al. 2011; Wolter et al. 2021). Standardized and commercially available systems have helped to decrease failure rates, while increasing the reproducibility in animal models (Lang et al. 2016). All available fixation systems have different advantages and disadvantages. External fixators in varying stiffnesses can stabilize long bone fractures in a rigid or semi-rigid, more compliant, approach (Röntgen et al. 2010). Moreover, external fixators provide axial and rotational stability and have been proven to be highly reliable (Histing et al. 2011; Lang et al. 2016). However, as the fixing pins act as a connection between the inside and outside, infections are a common complication of external fixation in humans (Ferreira and Marais 2012; Kazmers et al. 2016) and also have to be considered in laboratory rodents (Lang et al. 2016). Furthermore, the fixator weight is rather heavy compared to the weight of the mouse (Histing et al. 2011). The open surgery approach is more invasive but allows for more

standardized gap sizes and fracture geometries (Röntgen et al. 2010). In contrast, intramedullary pins are lighter, less expensive and allow for a less invasive surgery. However, these closed fractures, induced by blunt trauma, can result in heterogeneous fracture geometries and varying gap sizes (Röntgen et al. 2010; Haffner-Luntzer et al. 2016). Disadvantages of intramedullary pins also include the absence of axial and rotational stability and the great damage to the medullary cavity, bone marrow and endosteum (Röntgen et al. 2010; Histing et al. 2011; Haffner-Luntzer et al. 2016; Lang et al. 2016). Although pin fixation does not provide rotational and axial stability, it is still considered a stable fixation. Alternatively, locking nails provide similar advantages and disadvantages as the pin fixation, but their flattened ends offer more stability than the intramedullary pin (Röntgen et al. 2010; Histing et al. 2011). Yet, this fixation is still limited in the rigidity that can be provided to the fracture site (Histing et al. 2011). Plates allow for a rigid or a less rigid fixation while maintaining the integrity of the bone marrow but bear the risk of damaging the periosteum (Lang et al. 2016; Hauser et al. 2018). Despite the availability of a variety of commercially manufactured fixation systems that can enhance the standardization of procedures and the subsequent healing process, researchers still often opt to use self-made devices such as cannulas as intramedullary pins. (Wolter et al. 2021).

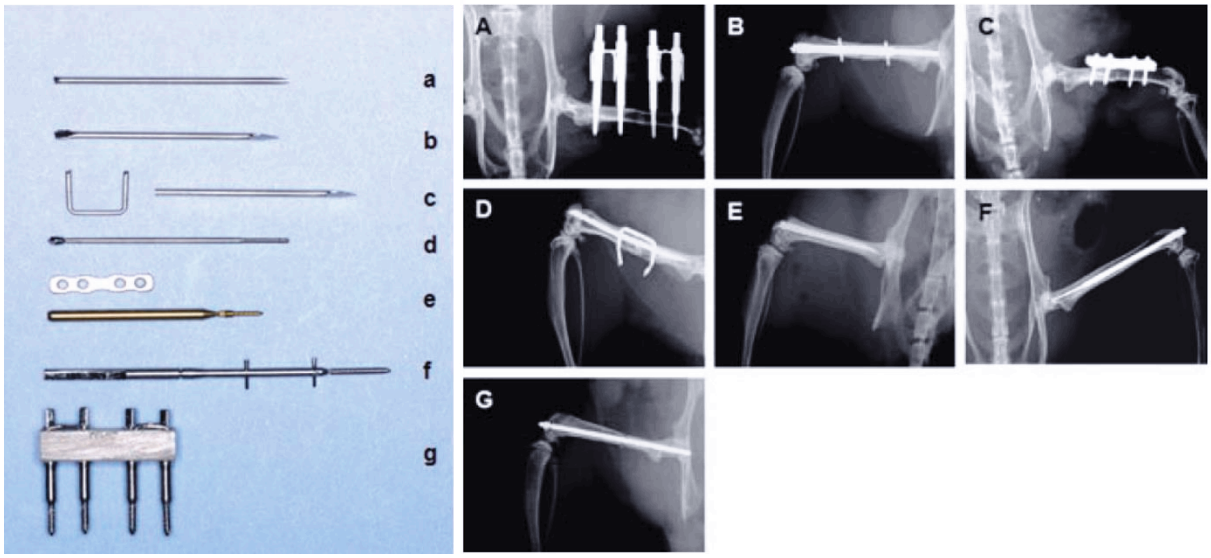

Figure 2 *Overview of existing fixation implants for long bone fractures in mice (a-g) and corresponding in vivo radiographs (A-G). (a) and (E) conventional pin (here: 24-gauge needle); (b) and (F) locking nail (proximal and distal end flattened); (c) and (D) pin-clip device (distally flattened pin and an extramedullary clip); (d) and (G) intramedullary screw (cone-shaped head and a proximal thread); (e) and (C) locking plate (plate and four screws); (f) and (B) mouse nail (proximal thread and two holes for locking pins); (g) and (A) external fixator (fixator block and four screws). Adapted and reproduced from Histing et al. 2009.*

2.4.3.1 Pain assessment in mouse fracture models

Orthopedic surgeries with effective stabilization and adequate analgesia are considered as a moderate severity according to Annex VIII of the Directive 2010/63/EU (European Parliament and the Council of the European Union 2010). For stable osteotomies, Lang et al. also proposed a moderate severity for the first 72h post-operative and a low severity for the following 7 to 10 days (Lang et al. 2016). The application of analgesics and a close monitoring after stable osteotomy is, therefore, often conducted for 72h post-surgery (Lang et al. 2016; Jirkof et al. 2019a; Haffner-Luntzer et al. 2021). Unstable fractures are classified as severe according to Annex VIII of Directive 2010/63/EU (European Parliament and the Council of the European Union 2010).

As no standardized score sheets for osteotomy models in rodents exist, different general and model-specific parameters are often used to monitor well-being and pain after osteotomy. Lang et al. proposed an example score sheet (grades 0-3), including general and model-specific indicators. General indicators include the general condition and appearance of the animals, body weight, facial expression, posture, food- and water intake and defecation/urination (Lang et al. 2016). Model specific indicators used for score sheets can include the state of the skin suture, general activity, standing- and walking behavior as well as load on the operated limb (Lang et al. 2016). Very few orthopedic related studies also describe the analyses of rearing behavior as an additional model-specific assessment tool in the first 3 days after surgery (see e.g., (Koewler et al. 2007; Majuta et al. 2017; Jirkof et al. 2019a)). Rearing is a normal behavior with an ambulatory component where the mice stand on their hind legs while the forelimbs are raised from the ground. It is shown by a variety of mammals, including rodents (Lever et al. 2006), and its analysis has been used in a wide variety of rodent models for several decades (Van Abeelen et al. 1973; Lever et al. 2006). In some studies, it has also been utilized to detect pain and reduced well-being in mice, e.g., after vasectomy or laparotomy (Wright-Williams et al. 2007; Roughan et al. 2014; Kendall et al. 2016). Overall, it has been suggested that the combination of general and model-specific parameters seems to be a useful approach for pain assessment in mouse osteotomy models and other orthopedic related models (Stasiak et al. 2003; Jirkof et al. 2019a).

2.4.3.2 Pain management in mouse fracture models

As stable osteotomies are categorized as moderately painful interventions, a reliable pain management is essential for the first 72h after surgery. Untreated or inadequately treated pain in fracture models may result in only limited use of the operated limb, thereby potentially impacting a successful and reproducible study execution as reduced loading can interfere with bone healing, a process promoted by mechanical stimuli and increased loading of the limb (Claes et al. 1998; Connolly et al. 2003; Schwarz et al. 2013; Birkhold et al. 2014; Jirkof 2017).

As inhibitors of cyclooxygenase 2, nonsteroidal anti-inflammatory drugs (NSAIDs) can have adverse effects on the initial proinflammatory phase of bone healing which limits the use of analgesics in bone healing models (Phillips 2005; Radi and Khan 2005; Cottrell and O'connor 2010). As an alternative to NSAIDs, opioids, in particular buprenorphine and tramadol, are frequently utilized for pain management in mouse femur fracture models (Lang et al. 2016; Wolter et al. 2021). As potent analgesics, opioids are frequently used for pain management in various painful rodent models (Kendall et al. 2016; Jirkof 2017). Pharmacological effects of opioids are achieved through their action on different opioid receptors at which they act as agonists, partial agonists, or antagonists and which are widespread within the gastrointestinal tract and central nervous system (Al-Hasani and Bruchas 2011; Pathan and Williams 2012; Jirkof 2017). Known side effects of opioid application in mice and rats include respiratory depression, blood pressure reduction, nausea and constipation, urinary retention, changes in locomotor activity and behavior, pica behavior as well as the development of addiction, tolerance, and hyperalgesia (Kuribara and Tadokoro 1991; Benyamin et al. 2008; Aronson 2009; Sauer et al. 2016; Jirkof 2017).

Despite the widespread usage of opioids for analgesic treatment in research, the literature and data on effective dosage recommendations and reapplication intervals in mice and rats varies greatly (Clark et al. 2014; Lang et al. 2016; Jirkof et al. 2019a; Wolter et al. 2021). Buprenorphine is mainly administered s.c. or i.p., but can also be administered orally, e.g., with the drinking water. Dosages for injections often vary between 0.05 - 0.75 mg/kg and reapplication intervals range from every 6h to 12h or even 24h (Jirkof et al. 2015; Jirkof et al. 2019a; Wolter et al. 2021). However, due to the half-life of buprenorphine, administration intervals of 4 to 6h are nowadays recommended (Jirkof et al. 2015; Arras et al. 2020). Due to its even shorter half-life (1 to 2h), it has been suggested in several studies that tramadol should only be applied with the drinking water or via other oral routes (Matthiesen et al. 1998; Evangelista Vaz et al. 2018a; Evangelista-Vaz et al. 2018b) and that repeated injections of tramadol are not suitable for adequate pain relief (Lang et al. 2016). The used dosages for tramadol in the drinking also vary greatly (0.025 mg/ml - 1 mg/ml) (Jirkof et al. 2019a; Wolter et al. 2021). Depending on the severity of pain, it has been suggested to consider additional injections during the light phase when applying analgesia via the drinking water, as water intake during this period is reduced due to the circadian rhythmic of mice (Sauer et al. 2016; Jirkof et al. 2019a; Arras et al. 2020).

The use of depot/sustained-release formulations of buprenorphine is another promising possibility for pain management in mouse fracture models, but its use in these models is still rare (Wolter et al. 2021). As of today, no corresponding product is available on the European market and the use of commercially available sustained-release/depot formulations of

buprenorphine in mouse femoral fracture models is, therefore, currently limited to USA-bound studies (see e.g. (Coates et al. 2019; Mckenzie et al. 2019; Ren et al. 2020)).

3 Aims and objectives

Reporting is often inadequate in many areas of (animal-based) biomedical research, resulting in potential scientific, ethical, and economic implications. Nevertheless, good, and transparent reporting is crucial to generate reliable, reproducible, and comparable data and to establish and maintain high scientific standards (Kilkenny et al. 2010; Avey et al. 2015; Strech and Dirnagl 2019; Percie Du Sert et al. 2020). However, different species and the great variety of experimental models in animal-based research do not allow for a universal one-fits-all approach in terms of reporting. Therefore, the first aim of this thesis was to evaluate the status quo of reporting accuracy of different general and model-specific parameters, including the respective analgesia regimes in studies using adequately stabilized mouse femoral fracture models.

Objective 1: Assessment of the reporting accuracy of experimental details in publications using mouse femoral fracture models.

A variety of analgesics such as NSAIDS and opioids are used to cover the expected period of greatest pain in laboratory rodents (Jirkof and Potschka 2021). In fracture healing studies, however, this choice is mostly reduced to the use of opioids such as buprenorphine and tramadol (Wolter et al. 2021). Opioids are potent analgesics, but due to their half-life in rodents, they must be injected repeatedly or given orally, e.g., via the drinking water. Besides oral applications, the stress associated with handling and repeated injections can also be avoided with the use of depot formulations, which ensure analgesic supply for up to 24h to 72h after injection. However, no such product is available on the European market. Different experimental formulations have, therefore, been developed in recent years to potentially fill this gap in Europe. Thus, in a next step, the analgesic efficacy of a promising experimental sustained-release formulation was evaluated in two mouse osteotomy models and compared with the analgesic capacities of the already established protocol of tramadol application via the drinking water.

Objective 2: Evaluation of a new experimental microparticulate depot formulation of buprenorphine for sustained post-surgical analgesia in two mouse osteotomy models.

For legal, ethical, and scientific reasons, pain in animal experiments must be reduced to a minimum. To establish adequate pain management, it is also crucial to reliably identify painful conditions in laboratory animals. However, pain assessment in laboratory rodents is an ongoing challenge and relies on tools and parameters aimed at observing the presence or absence of indicators of pain and (reduced) well-being. For a reliable pain assessment, multiple general and model-specific clinical, behavioral, and physiological parameters are

usually combined. In mouse fracture models, (manual) gait analysis is often used as a model-specific parameter. Therefore, the third part of this thesis aimed to examine whether the analyses of rearing behavior might be an additional helpful model-specific indicator for the identification of pain in this model.

Objective 3: Assessment of rearing behavior as a potential additional model-specific indicator for pain in two mouse osteotomy models.

4 Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models

Title: Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models

Authors: **Angelique Wolter***[a,b], Anna E. Rapp*[a,b], Mattea S. Durst[c], Laura Hildebrand[d], Max Löhning[a,b], Frank Buttgereit[a,b], Katharina Schmidt-Bleek[d,e], Paulin Jirkof[c,f], Annemarie Lang[a,b]

*these authors contributed equally

[a]Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Department of Rheumatology and Clinical Immunology, Berlin, Germany

[b]German Rheumatism Research Centre (DRFZ) Berlin, A Leibniz Institute, Pitzer Laboratory of Osteoarthritis Research, Berlin, Germany

[c]Division of Surgical Research, University Hospital Zurich, University Zurich, Switzerland

[d]Berlin Institute of Health at Charité, Universitätsmedizin Berlin, Center for Regenerative Therapies, Berlin, Germany

[e]Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Julius Wolff Institute, Berlin, Germany

[f]Office for Animal Welfare and 3Rs, University of Zurich, Switzerland

Peer reviewed publication in: *Bone*
November 2021; PMID: 34175502; DOI: 10.1016/j.bone.2021.116088

Citation: Wolter, A., A. E. Rapp, M. S. Durst, L. Hildebrand, M. Löhning, F. Buttgereit, K. Schmidt-Bleek, P. Jirkof and A. Lang (2021): Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models. Bone 152: 116088. DOI: 10.1016/j.bone.2021.116088.

Own contribution (shared first author)**:** Investigation, formal analysis, visualization, writing – original draft.

> Investigation – literature research, definition of detailed criteria for selection, verification of inclusion/exclusion criteria, development of the scoring system for the total score; formal analysis – abstract screening, systematic review of all papers, preparation and management of the summary table; visualization – designing the figures; writing – original draft.

Contribution of Anna E. Rapp (shared first author)**:** Investigation, formal analysis, visualization, writing – original draft.

Contributions of co-authors: Conceptualization: A.L., K.S.-B., P.J.; methodology: A.L., M.D., P.J., investigation: M.D., L.H.; writing – original draft: A.L; writing – review & editing: M.L., F.B., K.S.-B., P.J.; project administration: A.L.

You have to purchase this part online.

https://doi.org/10.1016/j.bone.2021.116088

5 A Buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72h in mouse femoral fracture models

Title: A Buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72h in mouse femoral fracture models

Authors: **Angelique Wolter**[1,2,3], Christian H. Bucher[4,5], Sebastian Kurmies[1,2], Viktoria Schreiner[6], Frank Konietschke[5,7], Katharina Hohlbaum[3,8], Robert Klopfleisch[9], Max Löhning[1,2], Christa Thöne-Reineke[3,5], Frank Buttgereit[1,2], Jörg Huwyler[6], Paulin Jirkof[10,11], Anna E. Rapp[1,2,12]*, Annemarie Lang[1,2,13]*

*these authors contributed equally

[1]Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Rheumatology and Clinical Immunology, Berlin, Germany

[2]German Rheumatism Research Centre (DRFZ), a Leibniz Institute, Berlin, Germany

[3]Institute of Animal Welfare, Animal Behavior and Laboratory Animal Science, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

[4]Julius Wolff Institute, Charité-Universitätsmedizin Berlin, Berlin, Germany

[5]Berlin Institute of Health Center for Regenerative Therapies (BCRT), Charité-Universitätsmedizin Berlin, Berlin, Germany

[6]Division of Pharmaceutical Technology, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland

[7]Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Biometry and Clinical Epidemiology, Berlin, Germany

[8]German Federal Institute for Risk Assessment (BfR), German Centre for the Protection of Laboratory Animals (Bf3R), Berlin, Germany

[9]Institute of Veterinary Pathology, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

[10]Office for Animal Welfare and 3Rs, University of Zurich, Zurich, Switzerland

[11]Dr. Rolf M. Schwiete Research Unit for Osteoarthritis, Department of Orthopedics (Friedrichsheim), University Hospital Frankfurt, Goethe University, Frankfurt, Germany

[12]Departments of Orthopaedic Surgery and Bioengineering, University of Pennsylvania, PA, United States

Peer reviewed publication in: *Scientific Reports*
March 2023; PMID: 36882427; DOI: 10.1038/s41598-023-30641-9

Citation: Wolter, A., C. H. Bucher, S. Kurmies, V. Schreiner, F. Konietschke, K. Hohlbaum, R. Klopfleisch, M. Löhning, C. Thöne-Reineke, F. Buttgereit, J. Huwyler, P. Jirkof, A. E. Rapp and A. Lang (2023): A buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72 h in mouse femoral fracture models. Scientific Reports 13: 3824 (2023). DOI: 10.1038/s41598-023-30641-9.

Own contribution: Study conduction, data collection, analysis and interpretation, drafting manuscript.

> Study conduction/execution – planning of the experimental design and schedule, training of the animals, conduct of all behavioral experiments and acquisition of all animal-related data, biological sample collection; data collection – processing of all collected data, processing of biological samples; analysis and interpretation – extensive analysis and interpretation of all collected data, statistical analysis, visualization; drafting manuscript – concept development, drafting the manuscript, designing the figures; correspondence with the reviewers.

Contributions of co-authors: Study conception: A.L., A.R., P.J., J.H.; study conduction: A.L., A.R., V.S., C.B., K.H., S.K., R.K.; data collection: A.L., A.R., S.K.; analysis and interpretation: A.L., A.R., C.B., F.K.; drafting manuscript: A.L., A.R.; revising manuscript: C.B., K.H., M.L., C.T.-R., F.B., J.H., P.J.

scientific reports

Check for updates

OPEN

A buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72 h in mouse femoral fracture models

Angelique Wolter[1,2,3]✉, Christian H. Bucher[4,5], Sebastian Kurmies[1,2], Viktoria Schreiner[6], Frank Konietschke[5,7], Katharina Hohlbaum[3,8], Robert Klopfleisch[9], Max Löhning[1,2], Christa Thöne-Reineke[3], Frank Buttgereit[1,2], Jörg Huwyler[6], Paulin Jirkof[10], Anna E. Rapp[1,2,11,13] & Annemarie Lang[1,2,12,13]✉

Adequate pain management is essential for ethical and scientific reasons in animal experiments and should completely cover the period of expected pain without the need for frequent re-application. However, current depot formulations of Buprenorphine are only available in the USA and have limited duration of action. Recently, a new microparticulate Buprenorphine formulation (BUP-Depot) for sustained release has been developed as a potential future alternative to standard formulations available in Europe. Pharmacokinetics indicate a possible effectiveness for about 72 h. Here, we investigated whether the administration of the BUP-Depot ensures continuous and sufficient analgesia in two mouse fracture models (femoral osteotomy) and could, therefore, serve as a potent alternative to the application of Tramadol via the drinking water. Both protocols were examined for analgesic effectiveness, side effects on experimental readout, and effects on fracture healing outcomes in male and female C57BL/6N mice. The BUP-Depot provided effective analgesia for 72 h, comparable to the effectiveness of Tramadol in the drinking water. Fracture healing outcome was not different between analgesic regimes. The availability of a Buprenorphine depot formulation for rodents in Europe would be a beneficial addition for extended pain relief in mice, thereby increasing animal welfare.

Animals—especially mice—are still widely used and required in fundamental and translational research to study the complexity of biological and pathophysiological processes. Therefore, the active implementation of the 3R principle (Replace—Reduce—Refine), with a particular importance of *Refinement* forms the indispensable basis for a humane approach to conduct animal experiments. Thus, adequate pain assessment and medication in animals before, during and after the experimental procedure are crucial to decrease suffering and ensure data quality. Insufficiently treated pain and handling-induced stress can affect animal behavior and physiological

[1]Department of Rheumatology and Clinical Immunology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. [2]German Rheumatism Research Centre (DRFZ), Leibniz Institute, Berlin, Germany. [3]Institute of Animal Welfare, Animal Behavior and Laboratory Animal Science, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany. [4]Julius Wolff Institute, Charité-Universitätsmedizin Berlin, Berlin, Germany. [5]Berlin Institute of Health Center for Regenerative Therapies (BCRT), Charité-Universitätsmedizin Berlin, Berlin, Germany. [6]Division of Pharmaceutical Technology, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland. [7]Institute of Biometry and Clinical Epidemiology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. [8]German Centre for the Protection of Laboratory Animals (Bf3R), German Federal Institute for Risk Assessment (BfR), Berlin, Germany. [9]Institute of Veterinary Pathology, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany. [10]Office for Animal Welfare and 3Rs, University of Zurich, Zurich, Switzerland. [11]Dr. Rolf M. Schwiete Research Unit for Osteoarthritis, Department of Orthopedics (Friedrichsheim), University Hospital Frankfurt, Goethe University, Frankfurt, Germany. [12]Departments of Orthopaedic Surgery and Bioengineering, University of Pennsylvania, Philadelphia, PA, USA. [13]These authors contributed equally: Anna E. Rapp and Annemarie Lang. ✉email: a.wolter@fu-berlin.de; Annemarie.lang@pennmedicine.upenn.edu

responses, especially in an immunological context, leading to potential bias in the scientific outcomes and reduced reproducibility[1–6]. However, evidence-based data on individual pain management efficiencies in surgical mouse models is still rare[6,7] and the reporting quality of the used analgesic protocols is often insufficient[8,9].

After surgical intervention, the potent opioid Buprenorphine is often used for pain relief in rodents[10,11]. The plasma half-life of Buprenorphine in mice has been reported to be 3 h after i.v. injection[12] and 3-5 h after s.c. injection[13]. In addition, several studies indicated that approx. 4 h after s.c. injection, plasma concentration in mice were lower than the therapeutic effective threshold in plasma (1 ng/ml)[14–16]. Although it has been described that Buprenorphine shows higher exposure in the brain when compared to the plasma concentration in mice 12 h after s.c. injection, pain alleviation measured by thermal sensitivity could not be achieved at that time point[16]. Therefore, frequent injections are required, resulting in repeated handling of the animals. However, the commonly reported application intervals of Buprenorphine of every 8–12 h can lead to pain peaks due to insufficient analgesic coverage[9,17] and recent guidelines, therefore, suggest application intervals of 4–6 h[18]. The parenteral application of other opioids such as Morphine, Tramadol and Fentanyl is less suitable for pain alleviation in rodents, as their half-life is even shorter[13,19,20].

To reduce handling-associated stress and to ensure continuous analgesic coverage, an alternative application route in form of administration of Buprenorphine or Tramadol via the drinking water has been routinely used e.g., in orthopedics-related mouse models[21–24]. However, as the uptake of analgesics via the drinking water is dependent on the drinking frequency and intake amount, the overall effectiveness of this treatment strategy might be highly influenced by e.g., reduced activity and water intake after anesthesia/surgery and circadian activity[25]. A drug formulation that extends the analgesic effect by sustained parenteral drug release can overcome such challenges and serve as a powerful tool to further refine today's analgesic regimens in animal experiments. However, current depot/sustained-release formulations of Buprenorphine for mice and rats are either only available for dedicated research purposes or are only available in the United States of America (USA), e.g., Buprenorphine ER-LAB (ZooPharm) or Ethiqa XR (indexed by U.S. Food and Drug Administration; Fidelis Animal Health). Attempts to import these products to Europe have failed due to missing approval through the European Medical Evaluation Agency (EMEA).

Schreiner et al. successfully developed a poly-lactic-co-glycolic acid (PLGA) based microparticulate drug formulation for sustained drug release of Buprenorphine in mice[16,26]. In a proof-of-concept study, they observed therapeutic-relevant drug levels of the sustained-release Buprenorphine (BUP-Depot) in the plasma for 12–24 h and in the brain for more than 24 h, an antinociceptive effect in the hot plate test, and pain relief after a minor abdominal surgery in female C57BL/6J mice for at least 72 h[16].

This present study, therefore, aims at exploring the analgesic capacities of the newly developed BUP-Depot and its potential to improve animal welfare in a wider range of mouse models in Europe. To test the effectiveness of the developed BUP-Depot in a preclinical setting of surgical interventions, we here compared the analgesic capacities of the BUP-Depot to the established application of Tramadol via the drinking water. Both pain management protocols were examined for their analgesic efficacy and adverse effects on experimental readouts in two femoral osteotomy models using rigid and flexible external fixators. To consider potential sex-dependent differences in response to the analgesic protocol, male and female mice were included. We monitored (i) general parameters of well-being e.g., body weight, food and water intake, nest building and explorative behavior, composite score, and (ii) model-specific pain parameters including walking behavior (limping score) and CatWalk analysis. In addition, fracture healing outcomes were examined at the end of the study to exclude negative influences on the regeneration process.

Results

To investigate the analgesic efficacy of the BUP-Depot, we chose an integrative study design to (i) generate intra-individual controls for the assessments and (ii) reduce animal numbers used in this study (Fig. 1). In brief, animals underwent a first intervention consisting of isoflurane anesthesia and administration of the assigned analgesics. Assessments were performed at 12 h, 24 h, 48 h and 72 h (referred to in the following as "post-anesthesia"). 14 days after the first intervention, the same animals underwent a second intervention including isoflurane anesthesia, administration of the respective analgesics and an additional osteotomy of the left femur. The same assessments as post-anesthesia were performed at 12 h, 24 h, 48 h and 72 h (in the following referred to as "post-osteotomy") (Fig. 1b). Of note, mice did not undergo osteotomy during the first intervention (post-anesthesia), but were already assigned to the respective fixation, leading to group descriptions of "rigid fixation" or "flexible fixation" even after anesthesia.

Body weight, food and water intake are affected post-anesthesia and post-osteotomy independent of analgesic regime and fixation. To assess general indications for well-being, the body weight was monitored at 12 h, 24 h, 48 h and 72 h post-anesthesia (initial body weight at 0 h—males: 25.96 ± 2.1 g; females: 21.41 ± 1.3 g) and post-osteotomy (initial body weight at 0 h—males: 27.61 ± 1.9 g; females: 22.75 ± 1.3 g). The body weight showed a statistically significant reduction in the range of 5% in all groups at 12 h and 24 h regardless of intervention, sex, fixation, and analgesic regime (Fig. 2). At 48 h post-anesthesia, body weight normalized in all groups and even exceeded the pre-intervention weight (Fig. 2a,b). After osteotomy, we found that male mice showed prolonged body weight loss over 48 h, when compared to post-anesthesia (Fig. 2a). Female mice had less body weight loss at 24 h post-osteotomy compared to post-anesthesia, but a similar recovery at 48 h (Fig. 2b). The body weight increase at 72 h compared to the initial value was higher post-anesthesia than post-osteotomy in all groups independent of sex, analgesia, and fixation. Until osteotomy/euthanasia, body weight was measured every other day with comparable weight development at any time point (Fig. S1).

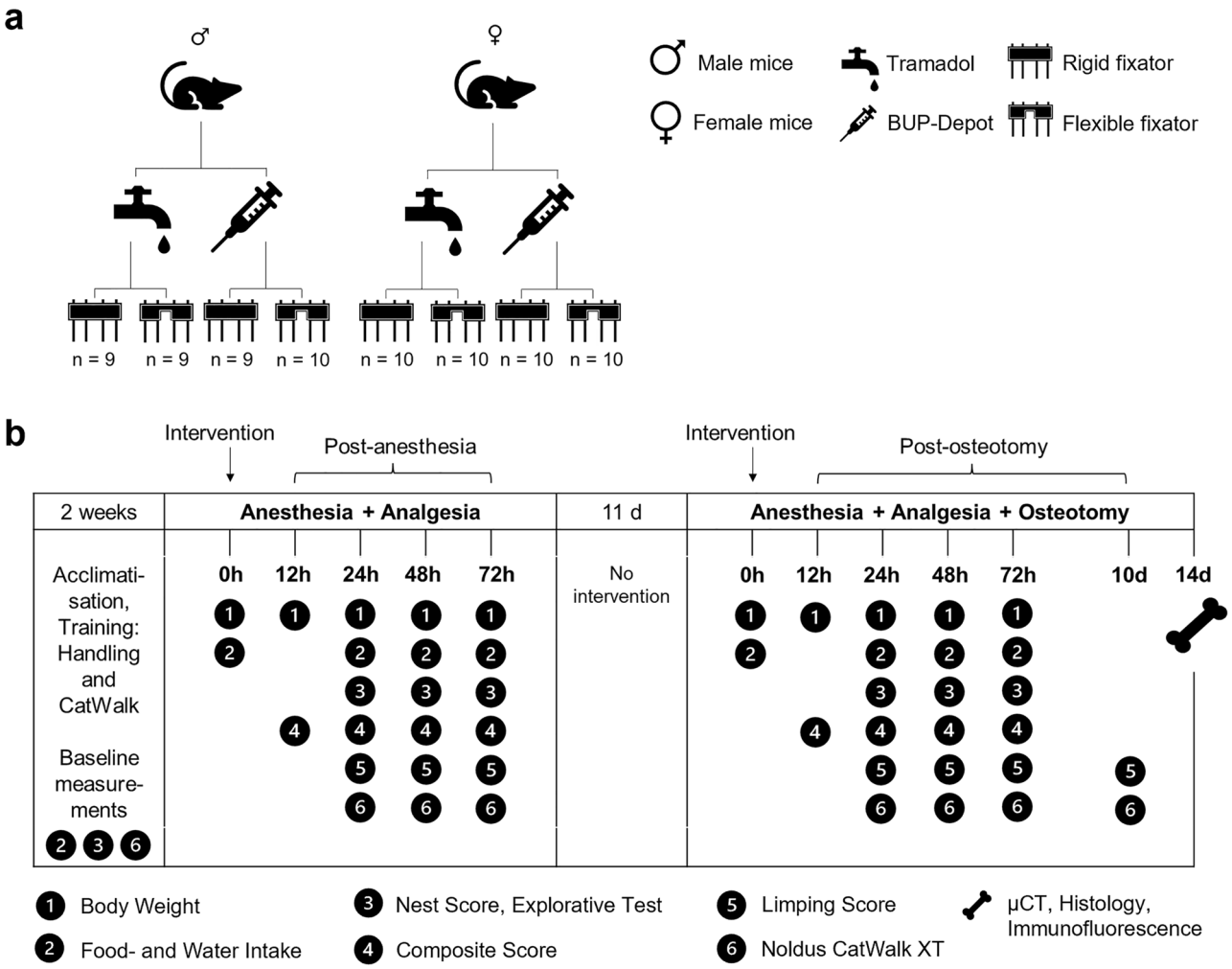

Figure 1. Group assignment and study design. Overview on (**a**) the group assignment, and (**b**) time points and measurements of the different parameters. To assure analgesic coverage during surgery, each animal received a single s.c. dose of Temgesic (1 mg/kg) at the beginning of each intervention. Depending on the assigned analgesic protocol, mice additionally received Tramadol (0.1 mg/ml) via the drinking water (provided one day before and for three consecutive days after the interventions) or a single dose of sustained-release BUP-Depot (1.2 mg/kg s.c) was administered at the end of the two interventions.

Food and water intake per cage were assessed at 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. Initial values at 0 h covering the previous 24 h per cage were as follows: post-anesthesia—males 8.1 ± 0.7 g (food) and 9.2 ± 1.5 ml (water); females 7.8 ± 1.2 g (food) and 9.2 ± 1.2 ml (water); post-osteotomy—males 8.7 ± 1.0 g (food) and 9.9 ± 1.5 ml (water); females 8.4 ± 0.9 g (food) and 9.5 ± 0.9 ml (water). The lowest food and water intake (approximately 50% reduction to initial values) across all groups and sexes was measured 24 h after each intervention and reached the level of the initial values at 48 h (post-anesthesia) or 72 h (post-osteotomy) (Fig. 2c–f). No significant main effects were detected between treatment groups (Tables S2, S3). To rule out any constipating adverse effects of the BUP-Depot, all groups were closely monitored for defecation during the assessments, as constipation is a known-side effect of chronic-opioid usage[17,27]. However, a reduction in defecation was only noticeable at 12 h after both interventions but showed no differences between the Tramadol and BUP-Depot groups. The reduction in defecation was considerably more pronounced in female than in male mice (Fig. S2).

Nest building and explorative behaviors are not influenced by the analgesic regime. To detect model independent changes in spontaneous behavior, monitoring of nest complexity scores and explorative behaviors was performed at 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. Analyses showed no differences in the nest building performance between the Tramadol and BUP-Depot groups post-anesthesia and post-osteotomy as medians ranged between scores of 4.5 to 5 (Fig. 3). Explorative behavior was present in all cages with male mice post-anesthesia while being reduced (no exploration) in some cages at 24 h post-osteotomy (8/20) independent of analgesic regime or fixation. In the female mice, 3 out of 20 cages (Tramadol flexible and BUP-Depot flexible) showed no explorative behavior at 24 h after both interventions (Fig. S3).

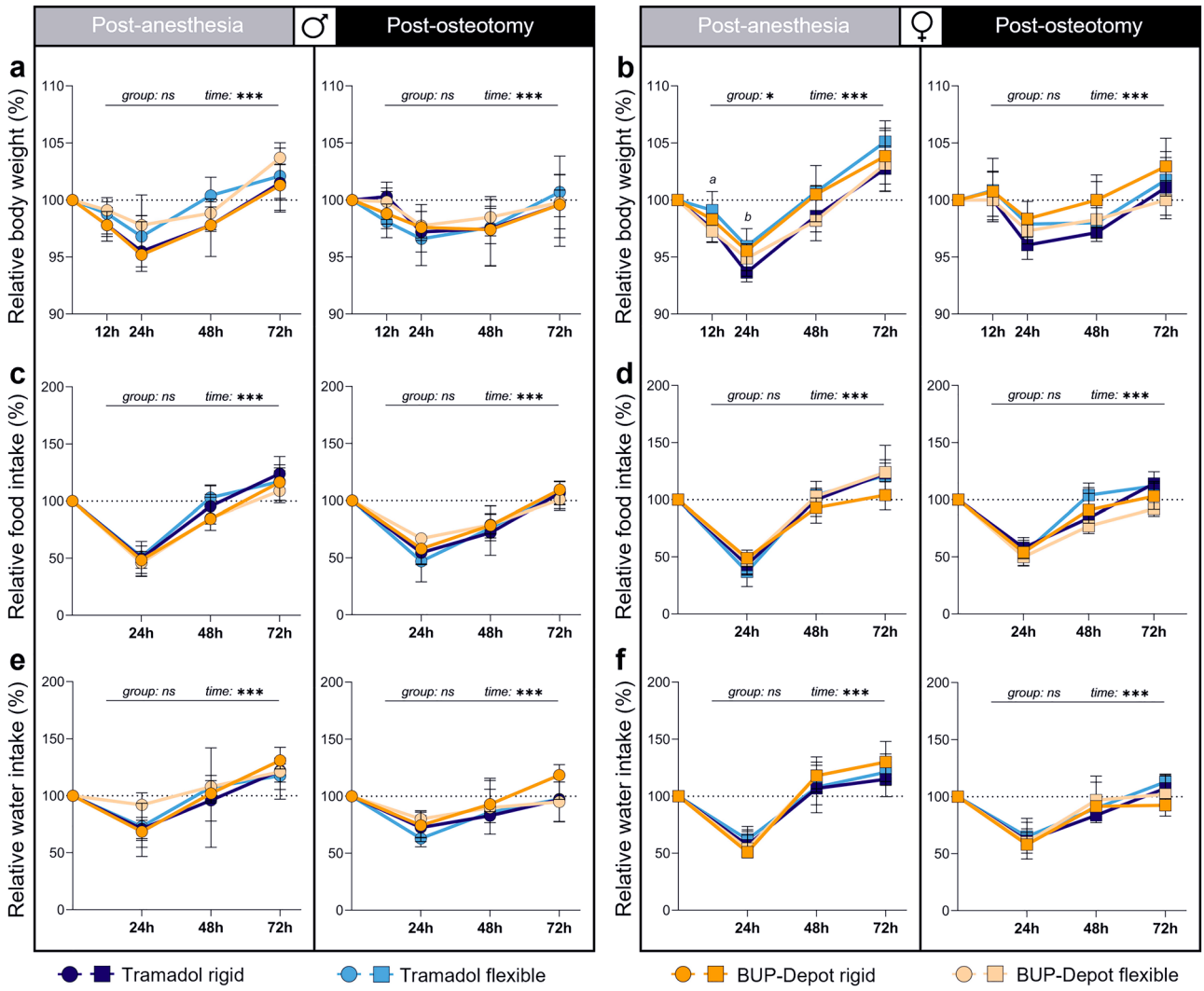

Figure 2. Reduction of body weight as well as food and water intake can be observed at 24 h and 48 h post-anesthesia and post-osteotomy. (**a,b**) Body weight was measured at 24 h, 48 h and 72 h; (**c–f**) and food/water intake was measured at 12 h, 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. Body weight and food/water intake were normalized to the initial value (0 h = 100%). Of note, mice did not undergo osteotomy during the first intervention (= post-anesthesia). However, they were already assigned to their respective groups post-osteotomy. All graphs show median with interquartile range for n = 9–10 (body weight) and n = 4–5 (food/water intake). Non-parametric ANOVA-type test—main effects of time and of group are represented in the graphs; exact *p*-values are listed in Table S1–S3; $*p < 0.05$, $***p < 0.001$. To determine group differences Kruskal–Wallis test and Dunn's post hoc test with Bonferroni correction were performed. (**a**) Significant difference Tramadol flexible vs. BUP-Depot flexible; (**b**) significant difference Tramadol rigid vs. Tramadol flexible.

Delta composite pain score indicates limited analgesic capacity of Tramadol in male mice with flexible fixation. The composite pain score combines parameters of facial expression (mouse grimace scale) and overall appearance and was assessed at 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. Since we observed an influence of the anesthesia and analgesic regime alone on the composite score (Fig. S4; Table S4–S4.3), we corrected the individual scores post-osteotomy for the respective scores post-anesthesia to obtain a delta composite pain score, that depicts the isolated effect of the osteotomy without the interfering effects of anesthesia and analgesia. The delta composite pain score was highest in all groups at 12 h and 24 h after osteotomy, and declined after 48 h and 72 h, reaching lowest scores in all groups after 72 h (significant main time effect in both sexes, $p < 0.001$; Fig. 4a; Table S5). Female mice showed comparable score developments between treatment and fixation groups with the highest median scores (1–1.5) at 12 h. In male mice, we found a significantly higher delta composite pain score in the flexible fixation group treated with Tramadol at 24 h (median score = 2) and 48 h (median score = 1.5) (Kruskal–Wallis test of all groups at 24 h: $p = 0.031$ and at 48 h: $p < 0.001$; Table S5.1) when compared to the other groups (Dunn's post hoc test for Tramadol flexible vs. BUP-Depot flexible, BUP-Depot rigid and Tramadol rigid, respectively: 24 h $p = 0.077, 0.042, 0.032$; 48 h $p = 0.004, < 0.001, 0.018$; Fig. 4a; Table S5.2–S5.3).

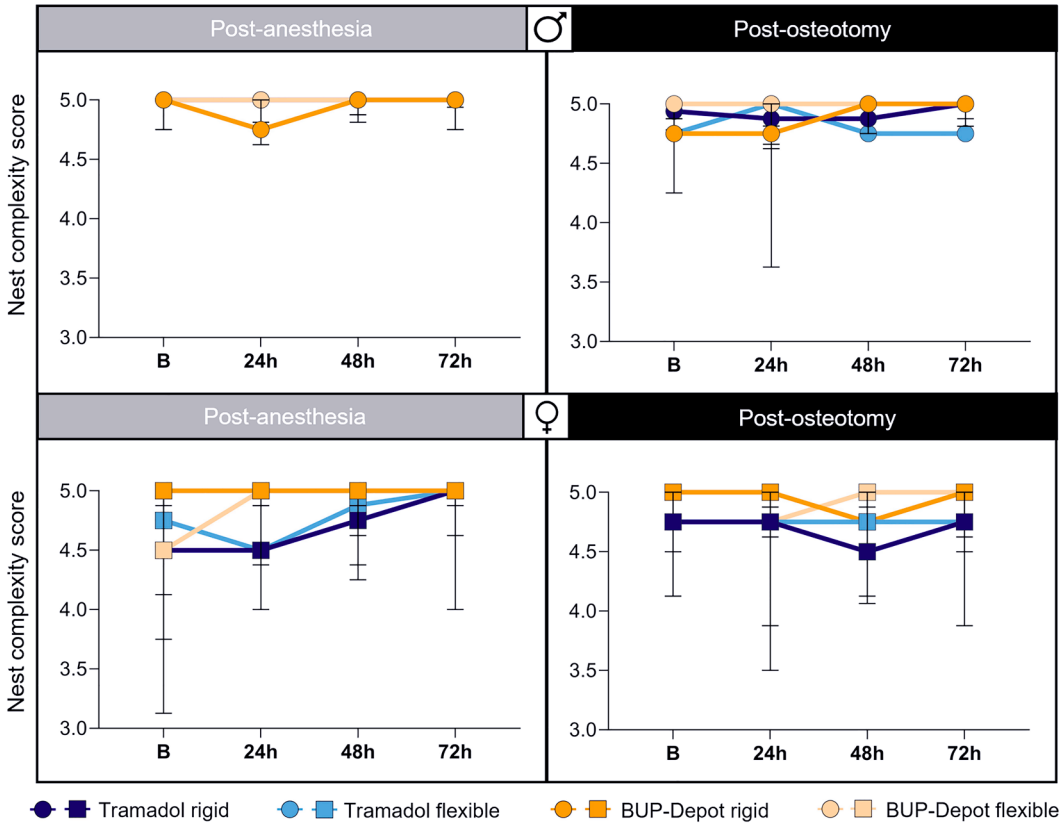

Figure 3. Nest building behavior remains largely unaffected post-anesthesia and post-osteotomy. Nest building was monitored per cage at 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. All graphs show median with interquartile range for n = 4–5 based on cages (pair housing). *B* baseline measurement.

Limping indicates model-related alterations in walking behavior with only slight differences between groups. Walking behavior was assessed by (i) using a metric limping score applied to individual 3 min videos and (ii) analyses of gait and locomotion using the Noldus CatWalk XT at 24 h, 48 h, 72 h, and 10 days post-osteotomy. Walking behavior was also videotaped at the respective time points post anesthesia, but limping was only considered post-osteotomy, as mice did not show any alterations in walking post-anesthesia during routine monitoring and videotaping.

As expected, limping was observed at all time points post-osteotomy until 10 days with a general improving trend (significant main time effect in males $p < 0.001$ and females $p = 0.018$; Fig. 4b; Table S6). In general, median scores ranged between 0 and 1 at 24 h and 48 h independent of sex and analgesics, while higher variations were seen at 72 h (male BUP-Depot rigid; female BUP-Depot flexible; Fig. 4b). Slight alterations in limping were still visible after 10 days in some individual mice in the female BUP-Depot groups and male Tramadol groups (all = score 1; sporadic limping; up to two animals, respectively).

Gait and locomotion analysis indicates alterations in walking behavior and velocity with differences between groups. Analyses of gait and locomotion by the CatWalk XT exhibited a clear reduction in velocity after osteotomy in all groups (significant main time effect post-osteotomy for both sexes $p < 0.001$; Fig. 5a,b; Table S7) with an upward trend towards 10 days. At 24 h and 48 h post-osteotomy, males with flexible fixation and Tramadol treatment showed a significantly lower relative velocity compared to males with rigid fixation and BUP-Depot (non-parametric ANOVA-type test for group $p = 0.034$; Kruskal–Wallis test 24 h $p = 0.052$ and 48 h $p = 0.038$; Dunn's post hoc test—$p < 0.05$ when compared to BUP-Depot rigid; Fig. 5a; Table S7.1–S7.3). In females, significant differences in velocity were also evident at 24 h post osteotomy, as mice with rigidly stabilized osteotomies exhibited elevated velocity compared to all other groups with female mice (non-parametric ANOVA-type test for group $p = 0.035$; Kruskal–Wallis test 24 h $p = 0.009$; Fig. 5b; Table S7.4–S7.5). With respect to the relative mean intensity (a measure for load bearing) and relative stand duration, we did not observe significant group effects independent of sex and analgesics (Fig. 5c–f), while time-dependent reductions at 24 h, 48 h and 72 h improved over time (relative mean intensity: non-parametric ANOVA-type test $p < 0.001$ in both sexes; relative stand duration $p < 0.001$ in males and $p = 0.005$ in females; Tables S8 and S9). However, the male mice with flexible fixation and Tramadol treatment showed numerically reduced values in relative mean intensity (Fig. 5c) when compared to the other males. In female mice, the relative stand duration only slightly increased over 10 days (median range 10 days: 0.68–0.75; Fig. 5f). The stride length was only markedly modulated over the first 72 h in female mice but then recovered to baseline values at 10 days (Fig. 5g,h;

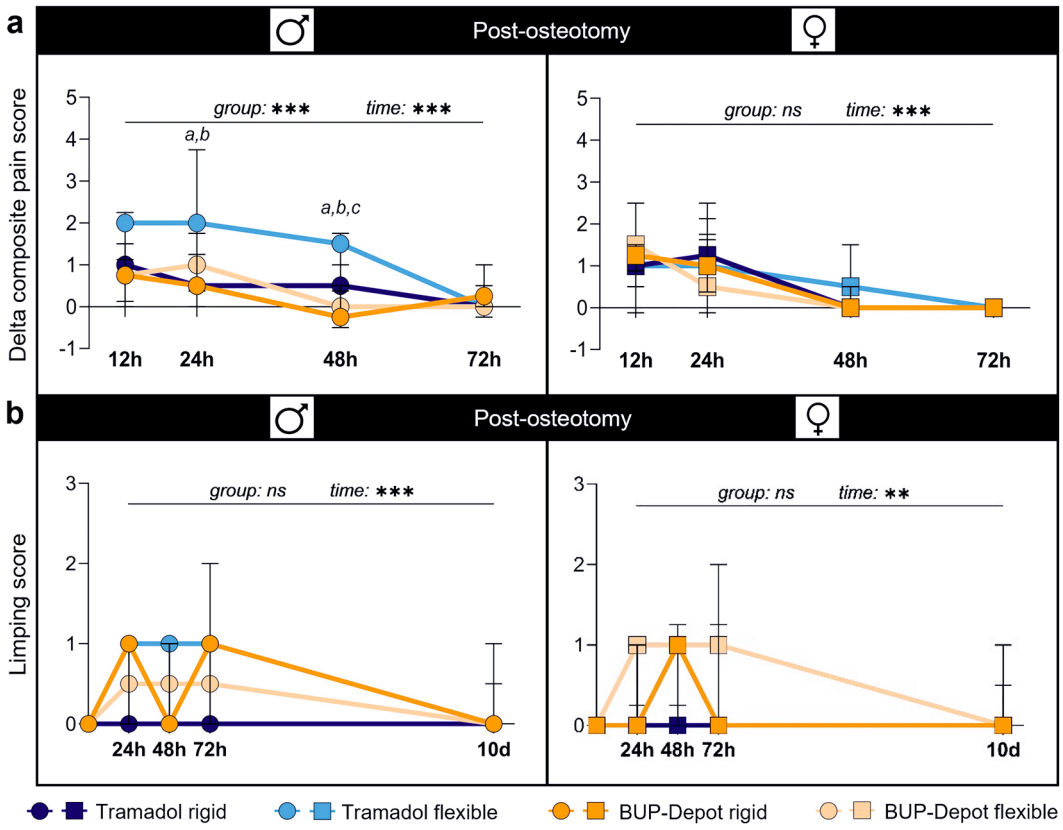

Figure 4. The delta composite pain score suggests adequate pain alleviation in most groups while limping score shows slight differences between groups post-osteotomy. (**a**) Scoring was performed at 12 h, 24 h, 48 h and 72 h post-anesthesia and post-osteotomy. For the delta composite pain score, scores from each individual mouse post-anesthesia were subtracted from their respective scores post-osteotomy. (**b**) The limping score was assessed at 24 h, 48 h, 72 h and 10 days post-osteotomy. All graphs show median with interquartile range for n = 8–10 (delta composite pain score) and n = 9–10 (limping score). Non-parametric ANOVA-type test—main effects of time and of group are represented in the graphs; exact *p*-values are listed in Table S5–S5.3 and Table S6; $^{*}p < 0.05$, $^{***}p < 0.001$. To determine group differences Kruskal–Wallis test and Dunn's post hoc test with Bonferroni correction were performed. (**a**) Significant difference Tramadol rigid vs. Tramadol flexible; (**b**) significant difference Tramadol flexible vs. BUP-Depot rigid; (**c**) significant difference Tramadol flexible vs. BUP-Depot flexible.

Table S10). Male mice with flexible fixation and Tramadol treatment also showed shortened stride length which was significant at 24 h compared to the BUP-Depot rigid group (non-parametric ANOVA-type test $p = 0.024$; Kruskal–Wallis test 24 h $p = 0.044$; Dunn's post hoc test—$p < 0.05$ when compared to BUP-Depot rigid; Fig. 5g; Table S10–S10.2), and did not recover over 10 days. To evaluate the influence of the velocity on mean intensity, stand duration and stride length, we performed Spearman correlation analyses (Fig. S5a–c) which indicated that relative mean intensity (males r = 0.602, females r = 0.336; Fig. S5a) and relative stride length (males r = 0.801, females r = 0.522; Fig. S5c) correlated with the relative velocity (all $p < 0.001$) after osteotomy. Spearman correlation analyses showed only a very week correlation between relative stand duration and relative velocity (males r = −0.15, females r = 0.046) (both $p < 0.001$). Interpretation of those parameters must, therefore, be considered in context of the overall velocity.

Fracture healing outcome is affected by fixation stability but not by analgesic regime. To evaluate potential effects of the BUP-Depot on fracture healing outcomes, we performed *ex-vivo* µCT (3D), histomorphometric analysis (2D) and vessel staining at day 14 post-osteotomy (Fig. 6). We observed a numerically lower BV/TV when comparing flexible fixation to rigid, except for males treated with Tramadol. However, the observed numerical differences did not reach statistical significance (BV/TV; Fig. 6a,b; Fig. S6; Table S11). The differences in BV/TV were slightly more pronounced in the BUP-Depot groups (male: median rigid = 25.87% vs. median flexible = 17.80%; female: median rigid = 26.12% vs. median flexible = 20.58%) than in the Tramadol groups (male: median rigid = 21.89% vs. median flexible = 21.52%; female: median rigid = 30.32% vs. median flexible = 23.31%).

Histomorphometric analysis revealed comparable differences between the rigid and flexible fixation in relative bone and cartilage fraction. As expected, bone formation was reduced while cartilage formation was elevated

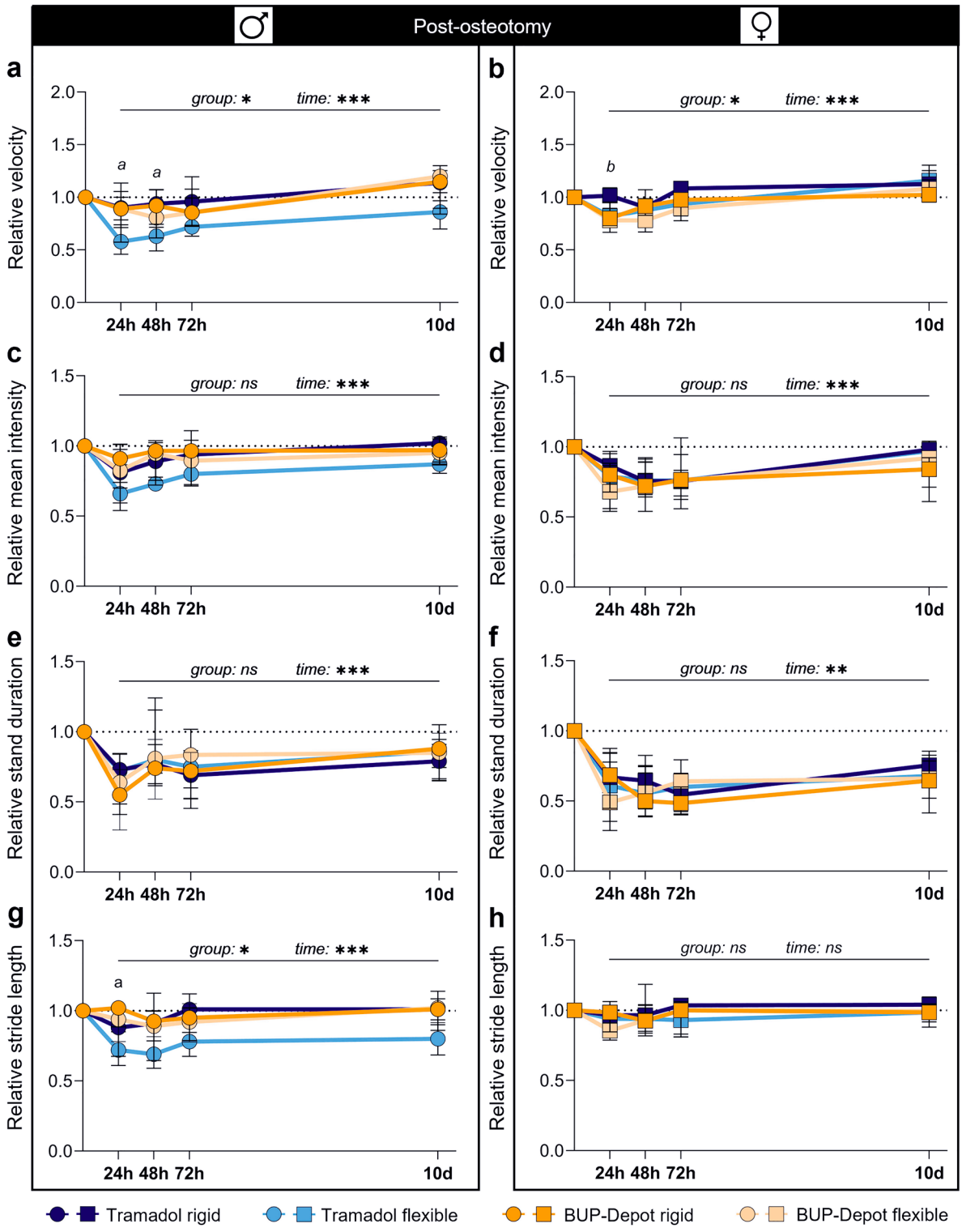

Figure 5. Gait analyses show model-related alterations in walking behavior specifically pronounced in male mice with flexible fixation and Tramadol treatment. CatWalk analysis was conducted at 24 h, 48 h, 72 h and 10 days post-osteotomy and normalized to the initial mean baseline value focusing on (**a,b**) relative velocity, (**c,d**) relative mean intensity, (**e,f**) relative stand duration and (**g,h**) relative stride length. All graphs show median with interquartile range for n = 8–10. Non-parametric ANOVA-type test—main effects of time and of group are represented in the graphs; exact *p*-values are listed in Table S7–S10.2; $*p < 0.05$, $***p < 0.001$. To determine group differences Kruskal–Wallis test and Dunn's post hoc test with Bonferroni correction were performed. (**a**) Significant difference Tramadol flexible vs. BUP-Depot rigid; (**b**) significant difference Tramadol rigid vs. BUP-Depot flexible.

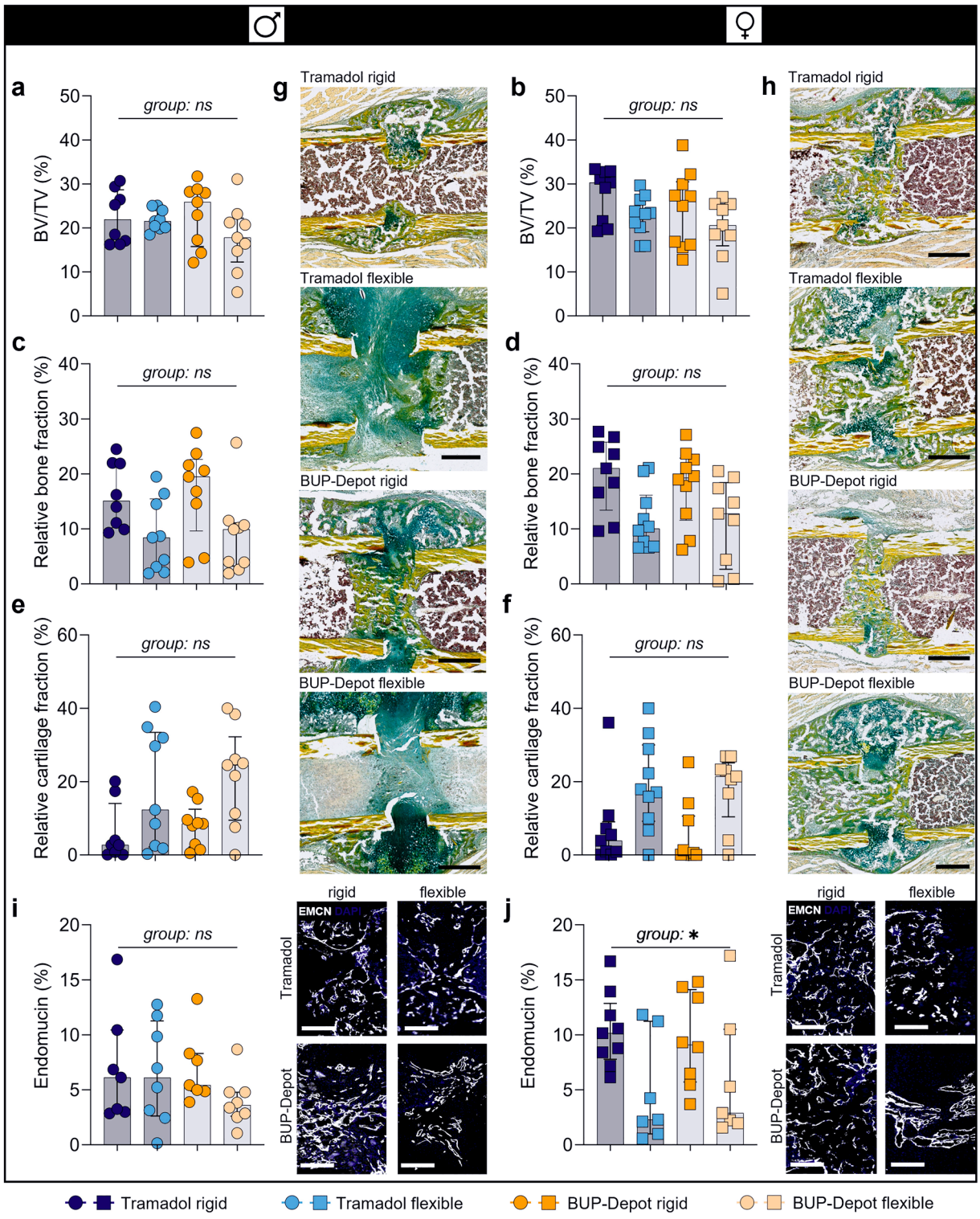

Figure 6. Analgesic regimes did not negatively affect fracture healing outcome at day 14, while the different fixations led to differences in new bone formation. (**a,b**) Relative bone volume (BV/TV) (%), (**c,d**) relative bone fraction (%) and (**e,f**) relative cartilage fraction (%). (**g,h**) Exemplary images of the Movat's pentachrome staining: yellow = mineralized bone, green = cartilage, magenta = bone marrow; scale bar 500 µm. (**i,j**) Immunofluorescence staining of vessel formation (Endomucin) including quantification and exemplary images; scale bar 200 µm. All graphs show median with interquartile range for n = 8–10 (**a–h**) and n = 6–9 (**i,j**). To determine group differences, Kruskal–Wallis test and Dunn's post hoc test with Bonferroni correction were performed; exact *p*-values are listed in Table S11–S13.1; $^{*}p < 0.05$.

in the flexible groups compared to rigid fixation in both sexes and treatment groups (Fig. 6c–h; Table S12). Analysis of vessel formation (Endomucin/Emcn staining) within the callus area did not show differences in the male groups (Fig. 6i; Fig. S6; Table S13). Differences in vessel formation (Emcn) were shown between fixation in female mice (Kruskal–Wallis test $p = 0.045$; Dunn's post hoc test $p > 0.05$ between all groups; Fig. 6j; Table S13.1) with flexible fixation groups showing less relative Emcn$^+$ areas when compared to the rigid fixation, independent of analgesia. Furthermore, DAPI staining indicated lower cellularity in the female mice with flexible fixation compared to rigid, also independent of analgesic treatment (Fig. S6).

Discussion

In the present study, we evaluated the analgesic efficacy and possible side-effects of a newly developed sustained-release Buprenorphine (BUP-Depot) in comparison to an already established protocol, Tramadol in the drinking water, in two mouse-osteotomy models with different fixation stiffnesses[16,22,26,28]. Due to individual differences in behavioral changes and pain perception[22,29], we chose a consecutive study design to analyze the effect of anesthesia and analgesia alone and in combination with an osteotomy in the same animal. The BUP-Depot delivered reliable pain relief over 72 h post-surgical without side effects on fracture healing outcome.

General clinical parameters such as body weight, and food and water intake were noticeably reduced post-anesthesia and post-osteotomy. Reduction of food intake and, therefore, negatively influenced body weight development are known side-effects post-surgical but can also be related to anesthesia or Buprenorphine/Tramadol administration[17,30–32]. Reduction of body weight observed after osteotomy was quite low (in the range of 5%) and similar to post-anesthesia, which is indicative of an appropriate pain management and well-being in all animals. Lowest values in body weight and food intake were reached after 24 h in all groups, with an increase over 48 h and 72 h post-anesthesia and post-osteotomy. Interestingly, body weight loss at 24 h was less pronounced after osteotomy than after anesthesia. The mice were 10 weeks of age during the first intervention and 12 weeks old at osteotomy. Mice show a rapid body weight development until skeletal maturity between approximately 10–12 weeks, which also includes the formation of more body fat[33,34]. Since interventions entailing anesthesia also result in short-term starvation during the recovery period, it can be speculated that mice at 10 weeks lost body weight more rapidly due to limited body fat reserves when compared to more mature mice (12 weeks). In general, the BUP-Depot showed similar effects on the body weight development as well as food and water intake as the established Tramadol treatment. We did not find differences in water intake between groups, excluding a negative effect of the Tramadol-containing water on the overall drinking amount and ensuring a continuous uptake of medication as shown previously[22,35]. Although the measurement of the overall 24 h water intake does not allow to conclude on sufficient water intake during the first hours after surgery, we have previously shown that the drinking frequency is indeed reduced, but the intake of Tramadol remains sufficient over 48 h post-osteotomy[22]. To improve the food intake, food can be alternatively provided on the cage floor to prevent the animals from having to stand on their hind legs[36]. However, this was not possible in this study. The use of high-caloric dietary gels could also be a valid alternative ensuring the consumption of food and liquids and can be also used as a route for oral analgesic administration[37].

Changes in nest building and the willingness to explore foreign objects can indicate alterations in well-being in laboratory mice[22,38–41]. In this study, nest building behavior was only scarcely influenced by anesthesia and osteotomy in all groups, independent of the analgesic regime, sex or fixation stiffness. This is in line with other studies suggesting that nest complexity scoring might not be sensitive enough to differentiate between minor pain and other potential stressors or reduced well-being after surgery[39]. However, these findings are contrary to our previous study where we reported a reduction in the nest building performance after osteotomy[22]. Technical variances (amount of nesting material) or individual differences in scoring might explain the variations in our findings and underline the necessity for more objective approaches. The explorative behavior in male mice seemed to be negatively impacted by osteotomy, especially at 24 h post-surgery, when compared to the post-anesthesia and female mice. Hohlbaum et al. also showed that female C57BL/6JRj mice exhibited shorter latency to explore than male mice 1 day after the last anesthesia in a repeated inhalation anesthesia trial, indicating a sex-specific difference that is in accordance with our findings with respect to explorative behavior[42].

A composite pain score was used to combine the assessment of facial expressions (parts of the mouse grimace scale)[43], and clinical appearance post-anesthesia and post-osteotomy. Based on our consecutive study design, we were able to calculate a delta for each individual mouse representing numerically the actual osteotomy effect. In line with other studies, the composite pain score was already slightly impacted by anesthesia and analgesia alone, indicating that some components of the composite pain score might not only be influenced by pain but rather also depict stress or discomfort[22,44,45]. The highest delta composite scores were reached 12 h and 24 h after osteotomy suggesting the pain/discomfort peak due to the surgical procedure. However, median delta scores varied around 1, indicating only limited residual pain and/or discomfort during the first 24 h, which constantly declined till 72 h post-surgical. The BUP-Depot provided comparable and sufficient alleviation of pain signs to the Tramadol treatment and our data indicate that pain relief can be achieved over 72 h after a single BUP-Depot injection.

As model-specific parameters, we assessed limping behavior and locomotion using gait analysis. After osteotomy, limping was observed over 72 h post-osteotomy in all groups, irrespective of sex, fixation, and analgesic, with no significant differences between the groups. A more detailed gait analysis using the Noldus CatWalk XT system revealed reduced velocity and altered gait patterns over 72 h and up to 10 days. While velocity, mean intensity and stride length ameliorated over 10 days in almost all groups (except male mice with flexible fixation and Tramadol analgesia), stand duration remained reduced in all groups. When analyzing CatWalk data, it needs to be considered that most gait parameters correlate to velocity and failing to address possible changes in velocity can affect the outcome of gait-related data[46–48]. In this study, for example, relative velocity in male mice correlated strongly with relative mean intensity and stride length, but only weakly with stand duration which

explains the comparable improvements over time. However, as the stand duration seems to be independent of the velocity, a limited functionality, especially with regard to the full restoration of the musculature, might be a plausible explanation. As the type of surgery performed here requires splitting of the muscle and transection of the muscular insertion at the trochanter major, a certain degree of the observed gait alterations might be due to the not yet fully restored muscular function. In addition, limited mobility and changed gait are also common in human patients with e.g., proximal femur fractures of the femoral diaphysis and are not directly related to pain[49,50]. Thus, we propose that the gait alteration in terms of stand duration over 10 days was likely caused by an unfinished functional restoration rather than pain or discomfort which is also supported by the absence of any additional pain-indicative signs at 72 h. As the relative velocity was markedly reduced in male mice with the more flexible fixation and Tramadol as their analgesics, the reduced values in the relative mean intensity and the relative stride length in this group are most likely explained by the reduced velocity. This group also displayed the highest delta composite pain scores after 24 h and 48 h indicating a potential clinically relevant level of discomfort or pain in individual animals. An explanation could be that the effective Tramadol dose was not achieved by the application of 0.1 mg/g Tramadol due to the higher body weight of male mice. Evangelista et al. showed that male mice had lower serum concentrations of Tramadol than female mice when applying Tramadol (0.2 mg/ml) via the drinking water for up to 30 h, although the analgesic effective M1 metabolite was similar between male and female mice[35]. This is in line with a previous study performed in rats[51]. In contrast, other studies report lower sensitivity of female mice or rats to Tramadol[52,53]. With respect to the influence of the body weight on the local degree of interfragmentary movement, which might cause discomfort when too high, Röntgen et al. characterized two configurations of the external fixator, the rigid one (18.1 N/mm) and a very flexible one (0.82 N/mm), calculating the interfragmentary strain in a 0.5 mm osteotomy gap for a 25 g mouse with 2.8% and 61%, respectively[54]. However, they did not find differences in body weight, ground reaction force and locomotion in female mice during 18 days, indicating sufficient analgesia even in the presence of higher local strains[54]. These observations highlight potential sex-specific differences in pain perception but also the response to analgesic medication[55]. Sex-specific adaptions of pain management regimes are, therefore, advisable in future studies. Our findings underline further, that pain management in animal experiments requires a constant reevaluation of the chosen protocol as well as the consideration of strain, sex, interindividual differences in animals, procedure, options to reduce re-injections and more[6]. As not only sex but also genetics influence experienced pain and response to analgesia in mice[55,56], further studies using different strains are needed in the evaluation of commonly used analgesic regimens in laboratory rodents.

In terms of fracture healing outcome analyzed by *ex-vivo* µCT and histomorphometry, we found no differences between the analgesic regimens, indicating safe use of the newly developed BUP-Depot in the analyzed models. A more flexible fixation allows pronounced interfragmentary movements and, therefore, promotes cartilage rather than bone formation as well as the formation of a larger periosteal callus[54], as seen in the histomorphometric analysis. Staining for endomucin, representing vessels, revealed no difference between fixations or analgesics in male mice. However, female mice with more flexible stabilized osteotomies showed a reduced Emcn-positive area as well as lower cellularity, regardless of their analgesic regime. Besides mechanical hindrance of revascularization due to higher strains, the higher proportion of cartilage observed in female mice with flexible stabilized osteotomies might have prevented revascularization due to the intrinsic anti-angiogenic nature of cartilage e.g., due to chondromodulin-1[57,58]. This reduced vascularization is in accordance with earlier observations in sheep[59]. The different patterns of Emcn staining in males and females might also indicate a more advanced callus remodeling and, therefore, a more rapid healing progression in males than females[60].

An effective sustained-release Buprenorphine in Europe would not only be a conceivable alternative to the application of Tramadol with the drinking water as investigated in this study, but also a potential alternative to repeated injections of Buprenorphine, which are still most frequently used for pain management in femur fracture models[9]. Assessment of the analgesic effect of Buprenorphine is most often based on measurements of plasma or blood serum levels. Studies specified therapeutic effective concentrations of Buprenorphine in plasma at a threshold of around 1 ng/ml or 1 ng/g in mice and rats[10,61,62]. However, Buprenorphine works through the µ-, κ- and δ-opioid receptors in the brain[63,64] and thus, reliable pain alleviation is more reliant on specific binding concentration values in the brain than specific plasma concentrations[16], as exemplary demonstrated by a correlation between analgesic effects and specific binding concentrations of Buprenorphine in the brain of rats[65]. Schreiner et al. found that the BUP-Depot showed effective concentration for up to 72 h in murine brains[16]. Moreover, they contemplated that specific binding concentrations of 5 ng/g (as observed 24 h after injection of the BUP-Depot) in the brain might be needed for reliable pain relief in mice. Binding concentrations of less than 3 ng/g at 48 h post-injection of BUP-Depot—a concentration comparable to levels observed 12 h after Temgesic injection—still resulted in high, but not significantly increased withdrawal latencies compared to a single injection with Temgesic or NaCl[16]. They, therefore, suggest that alleviation of strong pain through the BUP-Depot might require the administration every 24 h[16]. Based on our assessment of the clinical, behavioral, and model-specific parameters, we can postulate that the analgesic properties of the BUP-Depot are sufficient for 72 h post-operative analgesia in our specific mouse osteotomy model of moderate severity. Nonetheless, the potential use and the possible need for re-application of the BUP-Depot in other, more painful models still need to be critically assessed and evaluated.

Taken together, our assessment of clinical, behavioral, and model-specific parameters suggest that the analgesic properties of the BUP-Depot were sufficient for 72 h post-operative analgesia in male and female C57BL/6N mice after femoral osteotomy stabilized with external fixators. The BUP-Depot, therefore, provides an excellent alternative for extended pain relief in preclinical studies. The availability of such a sustained-release formulation of Buprenorphine in Europe would be substantially beneficial for mouse analgesia in animal experiments.

Animals, material and methods

Ethics and guidelines. All methods were carried out in accordance with relevant guidelines and regulations. In detail, the study was conducted according to the guidelines of the German Animal Welfare Act, National Animal Welfare Guidelines, and was approved by the local Berlin state authority (Landesamt für Gesundheit und Soziales—LAGeSo; permit number: G0044/20). Health monitoring in the animal facility was performed according to the FELASA guidelines (Supplementary Information).

Animals and husbandry. A total of 40 male and 40 female C57BL/6N mice aged 8 weeks were either provided by the Experimental Medicine Research Facilities (Charité—Universitätsmedizin Berlin, Berlin, Germany) or purchased from Charles River Laboratories (Sulzfeld, Germany). Mice underwent the first intervention (anesthesia/analgesia) at 10 weeks (body weight—males: 25.96 ± 2.1 g; females: 21.41 ± 1.3 g) and osteotomy at 12 weeks (body weight—males: 27.61 ± 1.9 g; females: 22.75 ± 1.3 g). Mice were housed in a semi-barrier facility in individually ventilated cages (IVC, Eurostandard Type II, Tecniplast, Milan, Italy). Housing conditions encompassed a 12/12–h light/dark cycle (light from 6:00 a.m. to 6:00 p.m.), room temperature of 22 ± 2 °C and a humidity of 55 ± 10%. Food (Standard mouse diet, Ssniff Spezialdiäten, Soest, Germany) and tap water were available ad libitum.

Mice were randomly divided into groups of two per cage. If mice had to be separated due to aggressive behavior, they were housed in a separated pair housing system in Green Line IVC Sealsafe PLUS Rat GR 900 cages (Tecniplast, Milan, Italy), which were divided in two equally sized compartments by a perforated transparent partition wall. Cages contained wooden chips (SAFE FS 14, Safe Bedding, Rosenberg, Germany), 20 g Envirodri (Shepherd Specialty Papers, USA), and a shredded paper towel as bedding and nesting material, a clear handling tube (Datesand Group, Bredbury, UK) and a mouse double swing (Datesand Group, Bredbury, UK). No houses were provided to allow unimpeded scoring and to reduce the risk of injury after osteotomy. After osteotomy, tunnels and double swings were removed from the cages to reduce the risk of injury. Two single swings (Datesand Group, Bredbury, UK) per cage were reinstalled 5 days after osteotomy. Animals were tunnel handled only and all experimenters performing analyses were female. Animal husbandry and care were in accordance with contemporary best practices.

Study design and experimental timeline. Reporting of this study was carried out in compliance with the ARRIVE 2.0 guidelines, including the "Arrive Essential 10" and most of the "Arrive Recommended Set". The study was pre-registered in the Animal Study Registry (Bf3R, Germany; https://doi.org/10.17590/asr.0000221).

The study included 8 groups with each n = 9–10 mice, comparing male and female mice, rigid and flexible external fixators, and two different pain management protocols: Tramadol via drinking water or sustained-release Buprenorphine (BUP-Depot; s.c. injection) (Fig. 1a). Cages and mice received individual random numbers that did not allow any inferences for the analgesic regimen or fixation group. Experimenters performing the pre- and post-surgical training and investigations were blinded.

After acclimation for 5 days, training was performed by one female experimenter, following a 2-week schedule to accustom the mice to the experimenter, tunnel handling, Noldus CatWalk XT (Noldus, Wageningen, Netherlands) and observation boxes (Ugo Basile, Gemonio, Italy). Baseline measurements were also obtained during this period (Fig. 1b). To correct for individual behavioral and clinical changes induced by anesthesia and analgesia alone, mice were first anesthetized and received their assigned analgesic protocol without any further surgical procedure (first intervention). Then, parameters were assessed at 12 h, 24 h, 48 h, and 72 h post-procedure. 14 days after the first intervention, the same animals were subjected to anesthesia, analgesia, and osteotomy (second intervention) and assessed at the above listed time points as well as at day 10 post-surgical. Mice were euthanized 14 days post-osteotomy to retrieve the osteotomized femur.

Analgesic regimes. Each mouse received one s.c. injection of regular Buprenorphine (1 mg/kg Temgesic, RB Pharmaceuticals, Heidelberg, Germany) at the beginning of each intervention (anesthesia/analgesia alone and osteotomy). Depending on the randomly assigned group, mice either additionally received Tramadol administered in the drinking water (0.1 mg/ml, Tramal Drops, Grünenthal, Stolberg, Germany) or a s.c. injection of the BUP-Depot (1.2 mg/kg). Tramadol was administered in the drinking water one day before and three consecutive days after both interventions. The BUP-Depot was injected once at the end of both interventions. BUP-Depot (RG 502 H-Big) was prepared at the University of Basel, Switzerland, as described previously by Schreiner et al.[16,26]. Four different batches of BUP-Depot were imported in accordance with national regulations for controlled substances (BtM import authorization No. 4679477). Each batch was analyzed prior to shipment for drug content, reconstitution time, and drug release kinetics as described previously[16,26]. The BUP-Depot was stored as a lyophilizate in glass vials at 4 °C. Each vial was reconstituted with physiological saline (0.9% NaCl) immediately before administration.

Anesthesia and osteotomy. Independent of the intervention, all mice were anesthetized with isoflurane (~ 2 to 3%; provided in 100% oxygen; CP-Pharma, Burgdorf, Germany) before being weighed and moved onto a heating pad (37 °C). Anesthesia was maintained at ~ 2% via a nose cone. Eye ointment, physiological saline (0.5 ml, 0.9% NaCl), Clindamycin (45 mg/kg, Ratiopharm, Ulm, Germany) and a single s.c. injection of Buprenorphine were applied. Anesthesia was then upheld for 15 min for the first intervention (anesthesia and analgesia alone). For osteotomy, the left femur was shaved and disinfected with alcoholic iodine solution. The osteotomy was conducted as described previously[22,66,67]. A longitudinal skin incision was made between knee and hip, and the musculus vastus lateralis and musculus biceps femoris were bluntly separated to expose the femur. Two standardized external fixators (rigid: 18.1 N/mm; flexible: 3.2 N/mm, both RISystem, Davos, Swit-

zerland) were used for stabilization. The external bar of the fixator was positioned parallel to the femur and all pins were positioned accordingly. Afterwards, an approximately 0.5 mm osteotomy gap was created between the second and third pin using a Gigli wire saw (0.44 mm; RISystem, Davos, Switzerland) and the gap was flushed with saline. Muscle and skin were closed with two layers of sutures (muscle: 5-0 Vicryl, skin: Ethilon 5-0, both Ethicon, Raritan, USA). For recovery, the mice were returned to their home cages under an infrared lamp and were closely monitored.

Body weight and food/water intake. Animals were weighed before the intervention (defined as 0 h), and at 12 h, 24 h, 48 h and 72 h after both interventions and then every other day until osteotomy/euthanasia. Food/water intake were measured per cage (i.e., two mice) by weighing food and water bottles every 24 h, beginning 1 day prior to each intervention and ending 3 days after. The difference to the previous value was calculated. All measurements were normalized to the respective baseline values at time point 0 h.

Explorative test and nest complexity score. Both scores were assessed before any other assessment or handling of the mice. To examine the motivation of the mice to explore and interact (sniffing, holding with forepaws, or carrying) with a foreign object, we added a Nestlet (Ancare, Bellmore, USA) to the home cages and observed the mice for one minute. The explorative test was scored 1 (interaction) or 0 (no interaction) per cage. An interaction of one animal of the cage was deemed as sufficient for a positive score. Nest complexity scoring was performed following Hess et al.[68] in the home cage assigning scores between 0 and 5.

Composite pain score. A composite pain score was used to combine the assessment of facial expressions and clinical appearance[22,43] (Table 1). The maximal score was 9. At 12 h, 24 h, 48 h, and 72 h post-intervention, the mice were transferred into a clear observation box and individually filmed for 3 min after an acclimatization period of 1 min (Basler Video Recording Software, Ahrensburg, Germany). Video analysis was performed by one blinded observer. As anesthesia and analgesia alone also affected the composite pain score, we calculated the delta composite pain score for each mouse by subtracting the scores from each individual mouse post-anesthesia from their respective scores post-osteotomy. This allowed us to evaluate the effect of the surgical procedure on an individual base without the interference of behavioral or clinical changes induced by the anesthesia and analgesia alone[44].

Walking behavior—limping score. To assess the walking behavior of each mouse, the limping score was assessed adapted from Jirkof et al.[22]. The mice were transferred to conventional type III cages that contained the same type of wooden chips as the home cages. After an acclimation period, a 3 min video was recorded. Walking behavior was examined at time points concurrent with the CatWalk analysis at 24 h, 48 h, and 72 h post-osteotomy as well as 10 days post-osteotomy. Video analysis was performed by two blinded observers and scores from 0 to 4 were assigned (Table 2). If walking seemed to be impaired due to a mechanical problem (e.g., displacement of the patella) one point was subtracted from the assigned score.

CatWalk analysis. Specific gait analysis was performed using the CatWalk XT Gait Analysis system for rodents (Noldus, Wageningen, the Netherlands). Multiple runs per animal were acquired before interventions as baseline measurements and at 24 h, 48 h, 72 h post-anesthesia (data not shown in results) and post-osteotomy, as well as 10 days post-osteotomy. Post-acquisition the runs were screened, and non-compliant runs as well as interrupted runs (e.g., by sniffing, rearing) were excluded. Runs that noticeably differed from the rest of the runs of this trial were also excluded, leaving an average of 4.4 runs per animal and time point for analyses. All runs

Parameter	Specification	Score (0 = not present)
Orbital tightening	Faint narrowing of the orbital area up to a tightly closed eyelid	Range: 0–2 0.5 = minimal 1 = moderately 1.5 = moderately severe 2 = severe
Ear position	Ears pulled back or rotated outwards and/or backwards	Range: 0–2 1 = moderately 2 = severe
Posture	Crouched posture, head and nose positioned towards the ground	Range: 0–1 0.5 = held for less than 10 s 1 = held for more than 10 s
Coat condition	Coat appeared disheveled or unkempt	Range: 0–1 0.5 = only certain body parts 1 = all over the body
	Erected fur, mice appear scruffy	Range: 0–1 0.5 = only on one body part 1 = on more than one body part, generalized
Movement	Apathetic, sedated, tipsy; crawling	Range: 0–1 each 0.5 = minimal 1 = moderately

Table 1. Composite pain score.

Specification	Limping score
Normal use, physiological gait	0
Complete ground contact and sporadic limping or alteration of the gait pattern	1
No complete ground contact or limping, constant alteration of the gait pattern	2
Partial non-use of the limb	3
Complete lack of use	4

Table 2. Limping score.

were classified automatically by the Noldus CatWalk XT software (version XT10.6) and revised for classification errors (i.e., incorrect identification of paws), which were corrected manually. From the obtained data, mean speed (cm/s) (velocity) and the following parameters were analyzed for the osteotomized left hind leg: mean intensity, stand duration (s) and stride length (cm). The two baseline measurements were used to calculate one baseline average value. Values at all time points were then normalized to the respective average baseline (time point measurement divided by average baseline value).

Euthanasia and sample collection. Euthanasia was carried out according to contemporary best practice. At 14 days post-osteotomy, mice were euthanized by cervical dislocation in deep anesthesia. The osteotomized femora were retrieved and fixed in 4% paraformaldehyde (PFA; Electron Microscopy Sciences, Hatfield, USA) at 4 °C for 6–8 h. The femora were then transferred into PBS until *ex-vivo* µCT was completed.

Ex-vivo µCT. To determine bone formation three-dimensionally, femurs were scanned in a SkyScan 1172 high-resolution µCT (Bruker, Kontich, Belgium). Voxel size was set to 8 µm and the bones were scanned with a source energy of 70 kV, 142 µA, a rotation step of 0.2 degrees and an 0.5 mm aluminum filter. Scans were reconstructed using NRecon (Bruker, Kontich, Belgium), applying ring artefact reduction and beam hardening corrections. CT Analyser software (version 1.20.3.0; both Bruker, Kontich, Belgium) was used for 2D and 3D analyses. By excluding the original cortical bone within the callus, the total volume (TV, mm^3), the total bone volume (BV, mm^3) and the bone volume fraction (BV/TV) of the newly formed bone were analyzed in a manually defined volume of interest (VOI)[21].

Histology and immunofluorescence. Following ex-vivo µCT, bones were placed in ascending sugar solutions as cryoprotectant (10%, 20%, 30%) at 4 °C for 24 h each, then cryo-embedded in SCEM medium (Sectionlab, Japan) and stored at −80 °C. Consecutive sections of 7 µm were prepared using a cryotome (Leica, Wetzlar, Germany) and cryotape (Cryofilm 2C(9), Sectionlab, Japan). Sections were fixed onto glass slides, air-dried, and stored at −80 °C until staining. Movat's pentachrome staining comprised the following steps: sections were air dried for 15 min, fixed with 4% PFA (30 min; Electron Microscopy Sciences, Hatfield, USA), pretreated with 3% acetic acid for 3 min, stained 30 min in 1% alcian blue pH 2.5, followed by washing in 3% acetic acid under light microscopic control. Sections were rinsed in H_2O_{dest} and immersed in alkaline ethanol for 60 min, then washed in tap water followed by incubation in Weigert's hematoxylin for 15 min. After washing in tap water for 10 min, sections were stained in crocein scarlet-acid fuchsin for 15 min, treated with 0.5% acetic acid for 1 min, followed by 20 min incubation in 5% phosphotungstic acid, and 1 min in 0.5% acetic acid. The sections were washed three times for 2 min in 100% ethanol, followed by incubation in alcoholic Saffron du Gâtinais for 60 min. The slides were dehydrated in 100% ethanol, cleared shortly in xylene, covered with Vitro-Clud and a cover slip. Imaging was performed on a Leica light microscope using LAS X software (Leica Microsystems GmbH, Wetzlar, Germany) at 10× magnification. Quantitative analyses of the Movat's pentachrome staining were evaluated using an ImageJ macro. All analyses were performed blinded to sex, fixation, and pain management protocol.

Immunofluorescence staining was performed as described previously[66,67] using the following antibody: Endomucin (Emcn) (V.7C7 unconjugated, rat monoclonal, sc-65495, 1:100; Santa Cruz Biotechnology, Dallas, USA), goat anti-rat A647 (1:500; A-21247, polyclonal, Invitrogen, Thermo Fisher Scientific, Waltham, USA) and DAPI (1:1,000; Thermo Fisher Scientific, Waltham, USA). Blocking was performed with 10% FCS/PBS and the staining solution contained 5% FCS and 0.1% Tween20 (Sigma Aldrich, St. Louis, USA). Images were acquired using a Keyence BZ9000 microscope (Keyence, Osaka, Japan). The images were processed and analyzed with ImageJ[69,70]. An area of interest was established and managed via the built-in ROI-Manager, while cell number and signal distribution within the area were determined using the plug-ins Cell-counter and Calculator Plus. Data was processed with the ImageJ plugin OriginPro.

Statistical analysis. The sample size was calculated based on own preliminary data[22] using a nonparametric ranking procedure to analyze longitudinal data. The required number of animals was modeled in R using the package nparLD[71]. Assuming a 20% difference and a power of ~ 80% resulted in n = 10 animals per group.

Statistical analysis was performed using RStudio and graphs were created in GraphPad Prism (V9). To test whether the data from female and male mice are homogenous and would allow for an integrated data analysis, body weight data of both sexes was first representatively compared using the F2-LD-F1 design. Since we found significant interaction regarding sex, both sexes were analyzed separately for all statistical analyses.

Nonparametric analysis of longitudinal data (F1-LD-F1 design; named non-parametric ANOVA-type tests), was used to test for significant differences in the main effect of time and main effect of group, separated by sex. When main group differences $p \leq 0.05$ were detected, group comparison for each time point was performed using the Kruskal–Wallis test[72]. To determine group differences, Dunn's post hoc test with Bonferroni correction was performed for each time point[73,74]. Non-parametric ANOVA-type tests as well as exact *p*-values, chi-squared and df of all analyses are provided in the Supplementary Information. Excluded mice and data are detailed in the Supplementary Information.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information file. Further information is made available by the authors upon request.

Received: 11 August 2022; Accepted: 27 February 2023
Published online: 07 March 2023

References

1. Balcombe, J. P., Barnard, N. D. & Sandusky, C. Laboratory routines cause animal stress. *Contemp. Top. Lab. Anim. Sci.* **43**, 42–51 (2004).
2. Carstens, E. & Moberg, G. P. Recognizing pain and distress in laboratory animals. *Ilar J.* **41**, 62–71. https://doi.org/10.1093/ilar.41.2.62 (2000).
3. Guo, S. & Dipietro, L. A. Factors affecting wound healing. *J. Dent. Res.* **89**, 219–229. https://doi.org/10.1177/0022034509359125 (2010).
4. Page, G. G. The immune-suppressive effects of pain. *Adv. Exp. Med. Biol.* **521**, 117–125 (2003).
5. Peterson, N. C., Nunamaker, E. A. & Turner, P. V. To treat or not to treat: The effects of pain on experimental parameters. *Comp. Med.* **67**, 469–482 (2017).
6. Jirkof, P. Side effects of pain and analgesia in animal experimentation. *Lab. Anim. (NY)* **46**, 123–128. https://doi.org/10.1038/laban.1216 (2017).
7. Flecknell, P. Rodent analgesia: Assessment and therapeutics. *Vet. J.* **232**, 70–77. https://doi.org/10.1016/j.tvjl.2017.12.017 (2018).
8. Carbone, L. & Austin, J. Pain and laboratory animals: Publication practices for better data reproducibility and better animal welfare. *PLoS ONE* **11**, e0155001. https://doi.org/10.1371/journal.pone.0155001 (2016).
9. Wolter, A. *et al.* Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models. *Bone* **152**, 116088. https://doi.org/10.1016/j.bone.2021.116088 (2021).
10. Guarnieri, M. *et al.* Safety and efficacy of buprenorphine for analgesia in laboratory mice and rats. *Lab. Anim. (NY)* **41**, 337–343. https://doi.org/10.1038/laban.152 (2012).
11. Roughan, J. V. & Flecknell, P. A. Buprenorphine: A reappraisal of its antinociceptive effects and therapeutic use in alleviating postoperative pain in animals. *Lab. Anim.* **36**, 322–343. https://doi.org/10.1258/002367702320162423 (2002).
12. Yu, S. *et al.* Pharmacokinetics of buprenorphine after intravenous administration in the mouse. *J. Am. Assoc. Lab. Anim. Sci.* **45**, 12–16 (2006).
13. Gades, N. M., Danneman, P. J., Wixson, S. K. & Tolley, E. A. The magnitude and duration of the analgesic effect of morphine, butorphanol, and buprenorphine in rats and mice. *Contemp. Top. Lab. Anim. Sci.* **39**, 8–13 (2000).
14. Clark, T. S., Clark, D. D. & Hoyt, R. F. Jr. Pharmacokinetic comparison of sustained-release and standard buprenorphine in mice. *J. Am. Assoc. Lab. Anim. Sci. (JAALAS)* **53**, 387–391 (2014).
15. Kendall, L. V. *et al.* Pharmacokinetics of sustained-release analgesics in mice. *J. Am. Assoc. Lab. Anim. Sci.* **53**, 478–484 (2014).
16. Schreiner, V. *et al.* Design and in vivo evaluation of a microparticulate depot formulation of buprenorphine for veterinary use. *Sci. Rep.* **10**, 17295. https://doi.org/10.1038/s41598-020-74230-6 (2020).
17. Jirkof, P., Tourvieille, A., Cinelli, P. & Arras, M. Buprenorphine for pain relief in mice: repeated injections vs sustained-release depot formulation. *Lab. Anim.* **49**, 177–187. https://doi.org/10.1177/0023677214562849 (2015).
18. GV-SOLAS. *Expert Information from the GV-SOLAS Committee for Anaesthesia in Collaboration with Working Group 4 in the TVT—Pain Management for Laboratory Animals*. https://www.gv-solas.de/wp-content/uploads/2021/08/2021-04_Pain_Management_for_laboratory_animals.pdf (2020/2021).
19. Evangelista-Vaz, R., Bergadano, A., Arras, M. & Jirkof, P. D. Analgesic efficacy of subcutaneous-oral dosage of tramadol after surgery in C57BL/6J mice. *J. Am. Assoc. Lab. Anim. Sci.* **57**, 368–375. https://doi.org/10.30802/AALAS-JAALAS-17-000118 (2018).
20. Girón, R. *et al.* Synthesis and opioid activity of new fentanyl analogs. *Life Sci.* **71**, 1023–1034. https://doi.org/10.1016/s0024-3205(02)01798-8 (2002).
21. Bucher, C. H. *et al.* Experience in the adaptive immunity impacts bone homeostasis, remodeling, and healing. *Front. Immunol.* **10**, 797. https://doi.org/10.3389/fimmu.2019.00797 (2019).
22. Jirkof, P. *et al.* Administration of Tramadol or Buprenorphine via the drinking water for post-operative analgesia in a mouse-osteotomy model. *Sci. Rep.* **9**, 10749. https://doi.org/10.1038/s41598-019-47186-5 (2019).
23. Rapp, A. E. *et al.* Induced global deletion of glucocorticoid receptor impairs fracture healing. *Faseb J.* **32**, 2235–2245. https://doi.org/10.1096/fj.201700459RR (2018).
24. Histing, T. *et al.* Obesity does not affect the healing of femur fractures in mice. *Injury* **47**, 1435–1444. https://doi.org/10.1016/j.injury.2016.04.030 (2016).
25. Sauer, M., Fleischmann, T., Lipiski, M., Arras, M. & Jirkof, P. Buprenorphine via drinking water and combined oral-injection protocols for pain relief in mice. *Appl. Anim. Behav. Sci.* **185**, 103–112. https://doi.org/10.1016/j.applanim.2016.09.009 (2016).
26. Schreiner, V., Detampel, P., Jirkof, P., Puchkov, M. & Huwyler, J. Buprenorphine loaded PLGA microparticles: Characterization of a sustained-release formulation. *J. Drug Deliv. Sci. Technol.* https://doi.org/10.1016/j.jddst.2021.102558 (2021).
27. Webster, L. R., Camilleri, M. & Finn, A. Opioid-induced constipation: Rationale for the role of norbuprenorphine in buprenorphine-treated individuals. *Subst. Abuse Rehabil.* **7**, 81–86. https://doi.org/10.2147/SAR.S100998 (2016).
28. Rapp, A. E. *et al.* Analgesia via blockade of NGF/TrkA signaling does not influence fracture healing in mice. *J. Orthop. Res.* **33**, 1235–1241. https://doi.org/10.1002/jor.22892 (2015).
29. Hohlbaum, K. *et al.* Impact of repeated anesthesia with ketamine and xylazine on the well-being of C57BL/6JRj mice. *PLoS ONE* **13**, e0203559. https://doi.org/10.1371/journal.pone.0203559 (2018).
30. Brennan, M. P., Sinusas, A. J., Horvath, T. L., Collins, J. G. & Harding, M. J. Correlation between body weight changes and postoperative pain in rats treated with meloxicam or buprenorphine. *Lab. Anim. (NY)* **38**, 87–93. https://doi.org/10.1038/laban0309-87 (2009).
31. Clark, M. D. *et al.* Evaluation of liposome-encapsulated oxymorphone hydrochloride in mice after splenectomy. *Comp. Med.* **54**, 558–563 (2004).

32. Goecke, J. C., Awad, H., Lawson, J. C. & Boivin, G. P. Evaluating postoperative analgesics in mice using telemetry. *Comp. Med.* **55**, 37–44 (2005).
33. Ferguson, V. L., Ayers, R. A., Bateman, T. A. & Simske, S. J. Bone development and age-related bone loss in male C57BL/6J mice. *Bone* **33**, 387–398. https://doi.org/10.1016/S8756-3282(03)00199-6 (2003).
34. Yang, Y., Smith, D. L. Jr., Keating, K. D., Allison, D. B. & Nagy, T. R. Variations in body weight, food intake and body composition after long-term high-fat diet feeding in C57BL/6J mice. *Obesity (Silver Spring)* **22**, 2147–2155. https://doi.org/10.1002/oby.20811 (2014).
35. Evangelista Vaz, R. *et al.* Preliminary pharmacokinetics of tramadol hydrochloride after administration via different routes in male and female B6 mice. *Vet. Anaesth. Analg.* **45**, 111–122. https://doi.org/10.1016/j.vaa.2016.09.007 (2018).
36. Lang, A., Schulz, A., Ellinghaus, A. & Schmidt-Bleek, K. Osteotomy models—The current status on pain scoring and management in small rodents. *Lab. Anim.* **50**, 433–441. https://doi.org/10.1177/0023677216675007 (2016).
37. Nizar, J. M., Bouby, N., Bankir, L. & Bhalla, V. Improved protocols for the study of urinary electrolyte excretion and blood pressure in rodents: use of gel food and stepwise changes in diet composition. *Am. J. Physiol.-Renal Physiol.* **314**, F1129–F1137. https://doi.org/10.1152/ajprenal.00474.2017 (2018).
38. Gaskill, B. N., Karas, A. Z., Garner, J. P. & Pritchett-Corning, K. R. Nest building as an indicator of health and welfare in laboratory mice. *J. Vis. Exp.* https://doi.org/10.3791/51012 (2013).
39. Jirkof, P. *et al.* Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring. *Lab. Anim.* **47**, 153–161. https://doi.org/10.1177/0023677213475603 (2013).
40. Jirkof, P. Burrowing and nest building behavior as indicators of well-being in mice. *J. Neurosci. Methods* **234**, 139–146. https://doi.org/10.1016/j.jneumeth.2014.02.001 (2014).
41. Rock, M. L. *et al.* The time-to-integrate-to-nest test as an indicator of wellbeing in laboratory mice. *J. Am. Assoc. Lab. Anim. Sci.* **53**, 24–28 (2014).
42. Hohlbaum, K. *et al.* Severity classification of repeated isoflurane anesthesia in C57BL/6JRj mice—Assessing the degree of distress. *PLoS ONE* **12**, e0179588–e0179588. https://doi.org/10.1371/journal.pone.0179588 (2017).
43. Langford, D. J. *et al.* Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **7**, 447–449. https://doi.org/10.1038/nmeth.1455 (2010).
44. Miller, A., Kitson, G., Skalkoyannis, B. & Leach, M. The effect of isoflurane anaesthesia and buprenorphine on the mouse grimace scale and behaviour in CBA and DBA/2 mice. *Appl. Anim. Behav. Sci.* **172**, 58–62. https://doi.org/10.1016/j.applanim.2015.08.038 (2015).
45. Jirkof, P., Arras, M. & Cesarovic, N. Tramadol: Paracetamol in drinking water for treatment of post-surgical pain in laboratory mice. *Appl. Anim. Behav. Sci.* **198**, 95–100. https://doi.org/10.1016/j.applanim.2017.09.021 (2018).
46. Batka, R. J. *et al.* The need for speed in rodent locomotion analyses. *Anat. Rec. (Hoboken)* **297**, 1839–1864. https://doi.org/10.1002/ar.22955 (2014).
47. Cendelín, J., Voller, J. & Vožeh, F. Ataxic gait analysis in a mouse model of the olivocerebellar degeneration. *Behav. Brain Res.* **210**, 8–15. https://doi.org/10.1016/j.bbr.2010.01.035 (2010).
48. Jacobs, B. Y., Kloefkorn, H. E. & Allen, K. D. Gait analysis methods for rodent models of osteoarthritis. *Curr. Pain Headache Rep.* **18**, 456. https://doi.org/10.1007/s11916-014-0456-x (2014).
49. Bizzoca, D. *et al.* Gait analysis in the postoperative assessment of intertrochanteric femur fractures. *J. Biol. Regul. Homeost. Agents* **34**, 345–351 (2020) (**Congress of the Italian Orthopaedic Research Society**).
50. Wong, J. *et al.* Gait patterns after fracture of the femoral shaft in children, managed by external fixation or early hip spica cast. *J. Pediatr. Orthop.* **24**, 463–471. https://doi.org/10.1097/00004694-200409000-00003 (2004).
51. Matthiesen, T., Wöhrmann, T., Coogan, T. P. & Uragg, H. The experimental toxicology of tramadol: An overview. *Toxicol. Lett.* **95**, 63–71. https://doi.org/10.1016/s0378-4274(98)00023-x (1998).
52. Dai, X. *et al.* Gender differences in the antinociceptive effect of tramadol, alone or in combination with gabapentin, in mice. *J. Biomed. Sci.* **15**, 645–651. https://doi.org/10.1007/s11373-008-9252-0 (2008).
53. Liu, H. C., Jin, S. M. & Wang, Y. L. Gender-related differences in pharmacokinetics of enantiomers of trans-tramadol and its active metabolite, trans-*O*-demethyltramadol, in rats. *Acta Pharmacol. Sin.* **24**, 1265–1269 (2003).
54. Röntgen, V. *et al.* Fracture healing in mice under controlled rigid and flexible conditions using an adjustable external fixator. *J. Orthop. Res.* **28**, 1456–1462. https://doi.org/10.1002/jor.21148 (2010).
55. Fillingim, R. B. & Gear, R. W. Sex differences in opioid analgesia: Clinical and experimental findings. *Eur. J. Pain* **8**, 413–425. https://doi.org/10.1016/j.ejpain.2004.01.007 (2004).
56. Smith, J. C. A review of strain and sex differences in response to pain and analgesia in mice. *Comp. Med.* **69**, 490–500. https://doi.org/10.30802/AALAS-CM-19-000066 (2019).
57. Patra, D. & Sandell, L. J. Antiangiogenic and anticancer molecules in cartilage. *Exp. Rev. Mol. Med.* **14**, e10. https://doi.org/10.1017/erm.2012.3 (2012).
58. Shukunami, C. & Hiraki, Y. Role of cartilage-derived anti-angiogenic factor, chondromodulin-I, during endochondral bone formation. *Osteoarthr. Cartil.* **9**(Suppl A), S91-101. https://doi.org/10.1053/joca.2001.0450 (2001).
59. Lienau, J. *et al.* Initial vascularization and tissue differentiation are influenced by fixation stability. *J. Orthop. Res.* **23**, 639–645. https://doi.org/10.1016/j.orthres.2004.09.006 (2005).
60. Haffner-Luntzer, M., Fischer, V. & Ignatius, A. Differences in fracture healing between female and male C57BL/6J mice. *Front. Physiol.* **12**, 712494. https://doi.org/10.3389/fphys.2021.712494 (2021).
61. Yassen, A., Olofsen, E., Dahan, A. & Danhof, M. Pharmacokinetic-pharmacodynamic modeling of the antinociceptive effect of buprenorphine and fentanyl in rats: Role of receptor equilibration kinetics. *J. Pharmacol. Exp. Ther.* **313**, 1136–1149. https://doi.org/10.1124/jpet.104.082560 (2005).
62. Yun, M.-H., Jeong, S.-W., Pai, C.-M. & Kim, S.-O. Pharmacokinetic-pharmacodynamic modeling of the analgesic effect of Buprederm™, in mice. *Health* **02**, 824–831. https://doi.org/10.4236/health.2010.28124 (2010).
63. Villiger, J. W. & Taylor, K. M. Buprenorphine : Characteristics of binding sites in the rat central nervous system. *Life Sci.* **29**, 2699–2708. https://doi.org/10.1016/0024-3205(81)90529-4 (1981).
64. Lutfy, K. & Cowan, A. Buprenorphine: A unique drug with complex pharmacology. *Curr. Neuropharmacol.* **2**, 395–402. https://doi.org/10.2174/1570159043359477 (2004).
65. Ohtani, M., Kotaki, H., Sawada, Y. & Iga, T. Comparative analysis of buprenorphine- and norbuprenorphine-induced analgesic effects based on pharmacokinetic-pharmacodynamic modeling. *J. Pharmacol. Exp. Ther.* **272**, 505–510 (1995).
66. Lang, A. *et al.* Collagen I-based scaffolds negatively impact fracture healing in a mouse-osteotomy-model although used routinely in research and clinical application. *Acta Biomater.* **86**, 171–184. https://doi.org/10.1016/j.actbio.2018.12.043 (2019).
67. Lang, A. *et al.* MIF does only marginally enhance the pro-regenerative capacities of DFO in a mouse-osteotomy-model of compromised bone healing conditions. *Bone* **154**, 116247. https://doi.org/10.1016/j.bone.2021.116247 (2022).
68. Hess, S. E. *et al.* Home improvement: C57BL/6J mice given more naturalistic nesting materials build better nests. *J. Am. Assoc. Lab. Anim. Sci.* **47**, 25–31 (2008).
69. Rasband, W. S. *ImageJ*. https://imagej.nih.gov/ij/ (1997–2018).
70. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675. https://doi.org/10.1038/nmeth.2089 (2012).

71. Noguchi, K., Gel, Y. R., Brunner, E. & Konietschke, F. nparLD: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. *J. Stat. Softw.* **50**, 1–23. https://doi.org/10.18637/jss.v050.i12 (2012).
72. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **47**, 583–621. https://doi.org/10.1080/01621459.1952.10483441 (1952).
73. Dinno, A. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *Stand. Genomic Sci.* **15**, 292–300. https://doi.org/10.1177/1536867x1501500117 (2015).
74. Dunn, O. J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **56**, 52–64 (1961).

Acknowledgements
We would like to thank Dr. Kristina Ullmann, Bianka Verrett and Sylvia Gundelach for their support during the planning and performance of the animal experiments. Special thanks to Darryl Borland and PD Dr. Maxim Puchkov for expert technical support and the preparation of the BUP-Depot. This study was funded by Charite3R, Charité-Universitätsmedizin Berlin (RefineMOMo 2.0; RefineLab) and a grant from Interpharma Basel. This work was further supported in part by the Willy Robert Pitzer Foundation (Pitzer Laboratory of Osteoarthritis Research at DRFZ, Berlin) and the Dr. Rolf M. Schwiete Foundation (Osteoarthritis Research Program). We acknowledge support by the Open Access Publication Fund of the Freie Universität Berlin.

Author contributions
Study conception: A.L., A.R., P.J., J.H.; study conduction: A.W., A.L., A.R., V.S., C.B., K.H., S.K., R.K.; data collection: A.W., A.L., A.R., S.K.; analysis and interpretation: A.W., A.L., A.R., C.B., F.K.; drafting manuscript: A.W., A.L., A.R.; revising manuscript: C.B., K.H., M.L., C.T.-R., F.B., J.H., P.J.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-30641-9.

Correspondence and requests for materials should be addressed to A.W. or A.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

6 Evaluating rearing behavior as a model-specific pain indicator in mouse osteotomy models

Title: Evaluating rearing behavior as a model-specific pain indicator in mouse osteotomy models

Authors: Angelique Wolter[1,2,3], Paulin Jirkof[4], Christa Thöne-Reineke[3], Anna E. Rapp[1,2,5*], Annemarie Lang[1,6* #]

**these authors contributed equally*

[1]Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Rheumatology and Clinical Immunology, Berlin, Germany

[2]German Rheumatism Research Centre (DRFZ), a Leibniz Institute, Berlin, Germany

[3]Institute of Animal Welfare, Animal Behavior and Laboratory Animal Science, Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

[4]Office for Animal Welfare and 3R, University of Zurich, Zurich, Switzerland

[5]Dr. Rolf M. Schwiete Research Unit for Osteoarthritis, Department of Orthopedics (Friedrichsheim), University Hospital Frankfurt, Goethe University, Frankfurt, Germany

[6]Departments of Orthopaedic Surgery and Bioengineering, University of Pennsylvania, PA, United States of America

Peer reviewed publication in: *Laboratory Animals*
September 2023; PMID: 37684025; DOI: 10.1177/00236772231183197

Citation: Wolter A, Jirkof P, Thöne-Reineke C, Rapp AE, Lang A. Evaluating rearing behaviour as a model-specific pain indicator in mouse osteotomy models. Laboratory Animals. 2023;58(1):9-21. Sep 8:236772231183197. doi: 10.1177/00236772231183197. Epub ahead of print. PMID: 37684025.

Keywords: Rearing behavior, Pain assessment, Exploratory behaviour, Refinement, Bone fracture.

Own contribution: Study conception, study conduction, data collection, analysis and interpretation, drafting manuscript.

> Study conception – planning of the experimental design and schedule; study conduction – training of the animals, conduct of all behavioral experiments and acquisition of all animal-related data; analysis and interpretation – extensive analysis

and interpretation of all collected data, statistical analysis, visualization; drafting manuscript – concept development, drafting the manuscript, designing the figures.

Contributions of co-authors: Study conception: A.L., A.R.; study conduction: A.L., A.R.; analysis and interpretation: A.L.; drafting manuscript: A.L., A.R.; revising manuscript: C.B., P.J.

You have to purchase this part online.

https://doi.org/10.1177/00236772231183197

7 Discussion

7.1 Reporting quality in mouse femoral fracture models

Experimental outcomes and data of animal-based research can be impacted by a multitude of confounding factors. Adequate and transparent reporting is, therefore, crucial to further improve reproducibility, reliability, and validity in animal experiments. It is also essential for scientific and ethical reasons (Smith et al. 1997; Kilkenny et al. 2009; Kilkenny et al. 2010; Smith et al. 2018; Percie Du Sert et al. 2020; Gkrouzoudi et al. 2022). Different systematic reviews and surveys have shown that reporting of animal experiments is often inadequate, a problem that causes potential scientific, ethical, and economic implications and reduces reproducibility and transferability of data (Smith et al. 1997; Kilkenny et al. 2009; Kilkenny et al. 2010; Avey et al. 2015; Ting et al. 2015; Carbone and Austin 2016; Gkrouzoudi et al. 2022). The detailed reporting of information on animals, materials and methods is, however, an essential key to enhance the scientific quality of animal-based research, improve animal welfare, reduce the number of animal experiments, and promote translation.

To enhance the reporting and planning of animal experiments, a range of guidelines have been devised. To promote high-quality and comprehensive reporting in animal research, the ARRIVE guidelines (Animals in Research: Reporting In Vivo Experiments) were proposed in 2010 (Kilkenny et al. 2010). They comprise of 20 information elements that should be addressed in manuscripts of studies using animals to maximize their scientific output (Kilkenny et al. 2010; Percie Du Sert et al. 2020). In 2020, the updated ARRIVE guidelines 2.0 were introduced, divided in the *ARRIVE Essential 10* and the *Recommended Set* (Percie Du Sert et al. 2020). The PREPARE guidelines (Planning Research and Experimental Procedures on Animals: Recommendations for Excellence), proposed in 2018 by Smith et al., complement the field with an overarching and adaptable checklist of recommendations for better planning of animal-based studies (Smith et al. 2018).

However, the different species and the great variety of experimental models in animal-based research do not allow for a universal one fits all approach in terms of reporting. The authors of guidelines such as ARRIVE and PREPARE, therefore, propose to provide scientists with essential adaptable elements rather than forcing a strict and imperative standard formula for the planning and reporting of animal-based research (Kilkenny et al. 2010; Smith et al. 2018; Percie Du Sert et al. 2020). According to Percie du Sert et al., reporting in line with *The ARRIVE Essential 10* and *The Recommended Set* serves as best practice, while they propose that at least *The ARRIVE Essential 10* should be included in each manuscript to ensure the reliability of the data (Percie Du Sert et al. 2020).

As good and transparent reporting is crucial to generate reliable, reproducible, and comparable data, and hence establish and maintain high scientific standards, the first objective of this thesis aimed at the evaluation of the reporting accuracy in studies using adequately stabilized mouse femoral fracture models. As fracture healing is a complex and closely regulated mechanism, it can easily be impacted by a variety of elements, such as strain, sex, immunological competence, drugs, fixation methods and more (Phillips 2005; Auer et al. 2007; Glatt et al. 2007; Manigrasso and O'connor 2008; Callewaert et al. 2010; Schwarz et al. 2013; Bucher et al. 2019; Bucher et al. 2022). In this systematic review, the primary focus was set on assessing the reporting accuracy of different general and model-specific parameters, including strain, sex, age, body weight, hygiene monitoring, fixation- and fracturing methods, as well as anesthesia and analgesia protocols in studies using mouse femoral fracture models.

In general, we observed an increase in publications per year within the examined timeframe from 2010 to 2019 from 16 included publications in 2010 to 33 included publications in 2019. This underlines the increasing use of mouse femoral fracture models in animal-based orthopedic research (Wolter et al. 2021). Since the inclusion criteria were strictly focused on mouse femoral fracture models with stable fixation, it must be noted that the overall number of studies using mouse fracture models was distinctly higher than the 254 studies analyzed in this specific review. In the following, the reporting accuracy of general and model-specific parameters as well as the commonly used anesthetic and analgesic regimes in mouse femoral fracture models and the importance of their adequate reporting are discussed more in detail.

General Parameters: This systemic review focused the reporting quality of general parameters according to *The ARRIVE Essential 10* for the reporting on *Experimental animals* including information on species, strain and substrain, sex, age or developmental stage, weight, as well as the health/immune status and genotype (Percie Du Sert et al. 2020). The reporting of the provenance of animals and their use in previous procedures are also listed in the guideline (Percie Du Sert et al. 2020), however, this information was not assessed within the review. We found a high reporting accuracy of the used mouse strain (99.6%) but identified a lack of reporting in terms of substrains of the most commonly used C57BL/6 mouse strain as well as the genetic background of genetically modified mice (Wolter et al. 2021). It has been shown that strain and genetic background of mice impact bone microstructure during development, as well as bone mineral density and elasticity, homeostatic properties, bone formation and fracture healing capacities (Akhter et al. 1998; Li et al. 2001; Akhter et al. 2004; Manigrasso and O'connor 2008; Haffner-Luntzer et al. 2016). Manigrasso and O'Connor reported a significantly faster fracture healing in C57BL/6 mice in comparison to DBA/2 and C3H mice (Manigrasso and O'connor 2008; Haffner-Luntzer et al. 2016), while Sheng et al. found lower osteoblast apoptosis in C3H/HeJ mice compared to C57BL/6J mice (Sheng et al.

2006). The improper reporting of different substrains and genetic background of mice can, therefore, lead to great difficulties in the interpretation and reproducibility of studies. Accurate reporting of the used strain and the genetic background can also be important for animal welfare reasons, as Gkrouzoudi et al. suggests that the more robust inbred, outbred, or mixed strains might withstand surgical interventions better than the more sensitive knock-out/in/down, genetically modified, transgenic, or genetically manipulated mice (Gkrouzoudi et al. 2022).

The sex and age of mice are also known or suspected to influence bone composition and fracture healing capacities (Glatt et al. 2007; Callewaert et al. 2010; Haffner-Luntzer et al. 2016; Haffner-Luntzer et al. 2021). In detail, Glatt et al. showed an early age-related decline in bone volume that continued throughout life and that was less pronounced in male mice compared to female mice (Glatt et al. 2007). Haffner-Luntzer et al. compared fracture healing capacities of female and male C57BL/6J mice and suggested that male mice exhibited a slightly faster fracture healing and a more prominent cartilaginous callus formation (Haffner-Luntzer et al. 2021), while other studies showed no sex-dependent differences in fracture healing (Working et al. 2021) and fracture callus composition (normalized to the body weight) (Wong et al. 2020). Different authors also propose the possibility that differences in tissue composition could be a result of different biomechanical loading properties, due to varying body weight in male and female mice (Histing et al. 2011; Haffner-Luntzer et al. 2021; Working et al. 2021). Although it is assumed that the body weight may play a critical role in bone regeneration, as local tissue strains are an important factor in the development of callus tissue (Haffner-Luntzer et al. 2021), only 22.4% of the analyzed studies reported the body weight of the animals (Wolter et al. 2021). We also found that nearly one quarter of the analyzed studies did not report the sex of the used mice, even though this information is readily available. In studies that reported the sex, male mice were used more frequently than female mice (male: 39.0%, female: 30.7%). Only 6.7% of studies reported the use of both sexes (Wolter et al. 2021). While the use of male and female mice may not be a suitable approach for all studies, the reporting of the used sex should be improved in future publications with respect to potential influences on bone composition and fracture healing properties. Most of the analyzed papers (92.9%) reported the age or at least the developmental stage of the mice, either stating a specific age, an age range or an age category (e.g., “mature”) (Wolter et al. 2021). However, especially the age categories should be defined more precisely in future studies, as the terms “mature” or “adult” represent a wide age range, and age-related changes in fracture healing capacities might impact study outcomes. Different reviews summarize a great variety of mouse studies showing age-related histological and molecular changes within the callus, effects on inflammation, changes in macrophage polarization, decrease in differentiation, proliferation, as well as numbers of different cell types, and the reduction in the expression of different collagen

types (Gibon et al. 2016; Clark et al. 2017), thereby underlining the importance of accurate age reporting in fracture healing studies.

A great lack in reporting was observed in terms of hygiene monitoring. Only 15% of all analyzed studies provided information on hygiene management (Wolter et al. 2021). This lack of reporting has also been shown in another systematic review on the reporting quality in mouse telemetry implantation surgeries (Gkrouzoudi et al. 2022). When hygiene monitoring was reported, we found that statements such as "controlled conditions", "standard conditions", "semi-barrier" or “non-SPF" were used in an attempt to describe the animal’s health, microbial and immune status as well as the husbandry conditions (Wolter et al. 2021). However, these statements are not sufficient without further explanations (Wolter et al. 2021). Specific-pathogen-free (SPF) mice are free of particular pathogens, which is ensured by regular testing. However, the list of monitored pathogens can vary greatly depending on the barrier level of the facility and even between the different animal rooms (Yeadon 2013). Hence, the SPF status only indicates the existence of routine health monitoring but does not depict information on the health or immune status of the mice, if the testing procedures and corresponding results are not provided (Wolter et al. 2021). Bucher et al. showed that bone homeostasis, structural properties, remodeling, and healing is directly impacted by the adaptive immunity and differences in immune composition (Bucher et al. 2019). A more specific reporting of housing conditions in the supplemental information, such as the description of the barrier level of the animal facility and the addition of the health monitoring and the corresponding results according to the FELASA guidelines, might be a realistic approach for future routine reporting. This was also incorporated in the reporting of the following two studies presented within this work. However, a more detailed specification in terms of microbiota is probably not realistic for routine reporting - since these data are not routinely obtained.

Model-specific parameters: Depending on the animal model used, further model-specific parameters can be relevant for adequate reporting. When using fracture or osteotomy models in animal-based research, the fixation method as well as fracturing procedure itself are essential to interpret the study outcomes. Different fracturing and fixation methods lead to different surgical approaches, diverse axial, and rotational stability and interfragmentary movements, and can cause varying damage to the surrounding tissue, the medullary cavity and the bone marrow (Röntgen et al. 2010; Histing et al. 2011; Lang et al. 2016). Encouragingly, the reporting of the used fracturing and fixation methods was high (fracturing methods reported by 96.9% and fixation methods reported by 92.3% of the studies). We found that the most used fixation method was the intramedullary pin fixation, followed by external fixators. Fixation by plates was the least frequently used method (Wolter et al. 2021). The term intramedullary pin included different devices such as “nail”, “pin”, “pin-clip”, “needle”,

“syringe” and “screw”. Unsurprisingly, we found that the used fracturing procedure was highly dependent on the used fixation. While three-point-bending was most often utilized when using a pin fixation, Gigli wires were the most used tool when fixing the fracture with an external fixator (Wolter et al. 2021). Even though a variety of fixation products is commercially available for each of three fixation methods, we noticed that many researchers still used self-made systems, such as syringes or cannulas, when using the intramedullary pin fixation (Wolter et al. 2021). The use of self-produced devices may be a straightforward and cheap solution, but it also increases the variability of these models. Systems from commercial suppliers can help to ensure a greater standardization and subsequent reproducibility of the used models and should, therefore, be used on a regular basis in future studies (Lang et al. 2016).

Anesthetic and analgesic regime: An adequate pain management in stable fixated mouse fracture models is necessary and recommended for the first 72h to cover the period of greatest pain. Thus, the corresponding reporting of the applied anesthesia and analgesia protocols is essential for animal welfare reasons but also for better validity and reproducibility of the acquired data. Inadequate pain management can lead to reduced loading of the operated limb and subsequently impaired or decelerated bone healing, as the process is promoted by mechanical stimuli and increased loading of the limb (Claes et al. 1998; Connolly et al. 2003; Schwarz et al. 2013; Birkhold et al. 2014). Furthermore, certain analgesics, for example NSAIDs, are suspected to interfere with bone healing capacities due to the inhibition of cyclooxygenase 2 (Radi and Khan 2005; Cottrell and O'connor 2010). Long term administration of opioids has also been reported to influence bone homeostasis and the skeletal system (Hirst et al. 2016; Janas and Folwarczna 2017; Ji et al. 2020). However, the literature is inconsistent and recent studies in mice and rats showed no negative effects of opioids on the skeletal system or fracture healing properties (Janas and Folwarczna 2017; Jirkof et al. 2019a).

While analyzing the studies regarding their reporting of the used anesthetic and analgesic regime, we found that 50.4% of the studies did report the use of anesthesia and analgesia, while 32.7% only stated the use of anesthesia and 3.1% only the use of analgesia. 13.8% of the studies did not report the use of anesthesia or analgesia, which, however, does not imply that anesthesia and analgesia were not used, but rather shows a lack of reporting qualities (Wolter et al. 2021). Over 90% of studies that stated the use of anesthesia and/or analgesia specified the anesthetic and/or analgesic regime. Of these, the most often used analgesic was buprenorphine (intra-operative: 55.2%, post-operative: 50.0%); followed by tramadol (intra-operative: 37.9%, post-operative: 37.7%), while carprofen (intra-operative: 3.5%, post-operative: 6.2%) and other analgesics (intra-operative: 3.4%, post-operative: 6.1%) were only rarely used (Wolter et al. 2021). These findings verify the great use of opioids as the standard analgesics in stable fixed mouse femoral fracture models, but also indicated great variances

regarding the dose, application frequency, and duration of the pain medication, which is consistent with the literature (Jirkof 2017). These variances influence the analgesic capacities of the drugs and thus, the load-bearing capacity of the operated limb, which in turn may impact fracture healing outcomes.

Although some analgesic protocols may have been established and used for an extended period of time, pain management should still undergo constant reevaluation and adaption in reference to new scientific developments (Jirkof and Potschka 2021). The pain management should also be adapted to the used fixation method, size of the fracture gap and potentially the body weight of the used animals. External fixators provide axial and rotational stability, while pin fixation does not (Histing et al. 2009). Pin fixations are still considered stable fixations, but they may lead to a greater strain for the animals due to their limited axial and rotational stability. Accordingly, the pain management protocols should be adapted to the stability of the used fixation method.

Reporting accuracy in mouse femoral fracture models in comparison to other fields: Over the last decades, several systematic reviews focused on the reporting accuracy and adherence to the ARRIVE guidelines in various pre-clinical models (Tihanyi et al. 2019). Carbone et al. set a special focus on the completeness of information on anesthesia and analgesia in a variety of different models and species (Carbone and Austin 2016). Adapted general and model specific reporting qualities were also analyzed by Bramhall et al. in animal models of colitis (Bramhall et al. 2015) and by Gkrouzoudi et al. in mouse models of telemetry implantation surgery (Gkrouzoudi et al. 2022). Comparing the reporting qualities on general animal-related items in the systematic review by Gkrouzoudi and colleagues with our results, similar reporting qualities can be observed. In detail, according to Gkrouzoudi et al., out of 110 eligible studies between 2011-2019, 96.4% (106/110) of the studies reported the strain and genetic status of the mice, which is comparable to our results (99.6%) (Wolter et al. 2021; Gkrouzoudi et al. 2022). Reporting quality on age was higher in the femur-fracture studies (92.9%) than in the telemetry studies (71.8%; 79/110), whereas information on the sex of the mice was slightly more commonly reported in the telemetry models (80.9%; 89/110) than in the femur-fracture models (72.4%). Reporting frequency on weight was comparable (telemetry model: 25.5%; 28/110; femoral fracure model: 22.4%) (Wolter et al. 2021; Gkrouzoudi et al. 2022). When comparing the reporting qualities two years before and after the publication of the ARRIVE guidelines, Baker and colleagues observed only slight improvements in the reporting quality and proposed that the guidelines were not effectively imposed by authors, reviewers, editors and journals (Baker et al. 2014). In accordance, Leung et al. assessed and compared the completeness of the reporting of the ARRIVE checklist in 2009 and 2015 in ARRIVE supporting and non-supporting journals and demonstrated that adherence to

reporting guidelines remained low and suggested that journal support for the ARRIVE guidelines did not enhance the reporting quality (Leung et al. 2018). When comparing reporting qualities of studies before 2010 and from 2011 to 2019, Gkrouzoudi et al. also observed only very slight improvements in reporting qualities (Gkrouzoudi et al. 2022). We did not assess the reporting quality of studies published before 2010. However, we evaluated the average reporting accuracy over time (2010 to 2019) in a summarized score of reported items per paper (maximum score of 10) and found only a slight increase in the mean score over the years (2010: 6.3 ± 1.2 and 2019: 7.1 ± 1.9) (Wolter et al. 2021), which is in line with the previously described reports. This further emphasizes that the lack in reporting accuracy is not research field-specific, but rather a systematic problem, that needs to be addressed together with the entire scientific community.

Limitations: Different general recommendations for systematic reviews suggest the use of more than one database for the systematic literature search to identify as many relevant papers as possible (Leenaars et al. 2012). However, only one database (PubMed) was used to retrieve potential studies for analyses in this systematic review. In future systematic reviews, the use of a second or even third database should be considered to allow for an even more comprehensive search approach. However, PubMed is widely used and provides access to MEDLINE, a major biomedical database (Leenaars et al. 2012). It can be assumed that a great number of standardized publications were found through the literature research that was carried out. Strict inclusion and exclusion criteria furthermore resulted in a high number of excluded papers. However, the restriction to adequately fixated mouse femur fracture models and the corresponding exclusion criteria helped to maintain the opportunity for an adequate comparison between the different protocols in such a diverse model.

Outlook: The detailed reporting of information on animals and methods is essential to further enhance quality of animal-based research. Although the reporting quality of studies using adequately stabilized femoral fracture models showed promising groundwork in the reporting of certain animal-based parameters, our work also showed that the reporting quality needs to be improved in future studies. This is in line with other systematic reviews assessing the reporting quality of animal-based studies which also found an overall lack in reporting qualities. Adherence to current guidelines, such as the ARRIVE guidelines, should be pursued collectively by authors, reviewers, editors, and journals and should be treated with the appropriate importance for future reporting of animal-based studies.

7.2 Evaluation of an experimental sustained-release buprenorphine formulation for post-surgical analgesia in mouse femoral fracture models

Opioids are the most used analgesics in adequately stabilized mouse femoral fracture models. However, great variances regarding the dose, the application frequency, and the duration of

the pain medication can be observed (Wolter et al. 2021). Buprenorphine is either injected or given orally while tramadol is mostly applied via the drinking water and rarely through subcutaneous injection (Wolter et al. 2021). Of note, injection intervals of buprenorphine or tramadol every 12h or 24h are not suitable for pain management after bone fracture, as adequate analgesic efficacy cannot be maintained with the previously described half-life of either drug (Jirkof et al. 2015; Evangelista Vaz et al. 2018a). In terms of buprenorphine, shortened injection intervals of 4 to 6h are recommended to maintain adequate analgesia (Jirkof et al. 2015; Arras et al. 2020), which however, are associated with additional handling-induced stress for the animals. Due to its even shorter half-life, repeated injections of tramadol are not suitable for adequate pain relief (Lang et al. 2016). Moreover, it has been suggested that tramadol should only be applied via oral routes (Matthiesen et al. 1998; Evangelista Vaz et al. 2018a; Evangelista-Vaz et al. 2018b).

The primary use of opioids in this model, necessitating frequent re-administration due to their relatively short half-life, highlights the need for analgesic administration options that require less animal handling. This can be achieved by oral analgesia application, the subcutaneous implantation of micropumps or the injection of sustained-release/depot preparations (GV-SOLAS 2020).

While the use of sustained-release buprenorphine formulations could be a promising addition for pain management in the mouse osteotomy model and mouse models in general, no corresponding product is available on the European market as of today. However, multiple sustained-release formulations of buprenorphine for veterinary use are legally marketed in the USA (Schreiner et al. 2020; Food and Drug Administration 2022). To do so, a veterinary drug intended for a minor species must either be approved, conditionally approved, or indexed by the FDA (Food and Drug Administration 2019). Different sustained-release buprenorphine preparations for rodents are indexed and marketed as unapproved animal drugs for minor species in the USA (Arras et al. 2020; Schreiner et al. 2020). As of January 2023, three buprenorphine extended/sustained-release formulations (BupreLab-Rat, BupreLab-Mouse and Ethiqa XR) for the management of post-procedural pain in mice and/or rats are listed in *The Index of Legally Marketed Unapproved New Animal Drugs for Minor Species* (Food and Drug Administration 2022). The product Buprenorphine SR-Lab from ZooPharm (Fort Collins, CO USA) was also listed in *The Index of Legally Marketed Unapproved New Animal Drugs for Minor Species* (Schreiner et al. 2020) and used for pain management in a variety of studies (see e.g., (Foley et al. 2011; Carbone et al. 2012; Clark et al. 2014; Healy et al. 2014; Kendall et al. 2016; Myers et al. 2021)). However, it is currently not listed as of January 2023 (Food and Drug Administration 2022). Even though different sustained-release preparations of

buprenorphine are available in the USA, their import to Europe was not successful in past attempts, likely due to their legal marketing status.

Simbadol (Zoetis, Parsippany, NJ USA), another commercially available option in the USA, is a high-concentration formulation of buprenorphine with a recommended reapplication frequency of 24h, indicated and FDA-approved for the use in cats (Myers et al. 2021). Its effectiveness has also been evaluated in other species, including mice, rats, and Rhesus macaques (Allen and Johnson 2018; Mackiewicz et al. 2019; Myers et al. 2021). However, Myers et al. identified that Simbadol did not provide prolonged serum concentrations in mice (Myers et al. 2021). Moreover, Allen and Johnson concluded that only the highest dosage of Simbadol resulted in short hypoalgesic effects in rats (Allen and Johnson 2018). The use of Simbadol for prolonged analgesia in mouse models, therefore, does not seem to be a feasible option.

The absence of sustained-release buprenorphine formulations on the European market impedes the road to more stress-free pain management options for a great number of laboratory rodents in Europe. Thus, different experimental depot formulations of buprenorphine have been developed in recent years (Jirkof et al. 2015; Schreiner et al. 2020; Schreiner et al. 2021). In detail, Schreiner and colleagues successfully designed and tested an experimental microparticulate sustained-release formulation of buprenorphine, based on poly-lactic-co glycolic acid (PLGA) for prolonged analgesia in mice (Schreiner et al. 2020; Schreiner et al. 2021). Schreiner et al. suggest that this novel depot preparation shows promising properties towards future industrial manufacturing and commercialization (Schreiner et al. 2021) and further concluded that post-surgical pain relief was accomplished after minor surgery (Schreiner et al. 2020). Specific binding concentrations of buprenorphine in the brain showed therapeutic drug levels for at least 24h and a tendency for pain relief for at least 72h (Schreiner et al. 2020). However, further evaluation of the pain-relieving properties of the sustained-release buprenorphine in other, more severe, mouse models is still pending. Schreiner et al. also pointed out that the currently available sustained-release preparations in the USA are not available in Europe and that the available products in the USA might have shortcomings, for example inflammation, erythema and necrosis at the injection site, as well as small injection amounts and greater viscosity of the suspensions (Carbone et al. 2012; Page et al. 2019; Schreiner et al. 2020; Myers et al. 2021; Schreiner et al. 2021).

As discussed above, further evaluations of the pain-relieving properties of the newly developed experimental sustained-release buprenorphine (BUP-Depot) by Schreiner et al. in other and more severe mouse models are currently lacking. The second objective of this thesis, therefore, focused on investigating the analgesic efficacy of the BUP-Depot by Schreiner and colleagues for post-surgical analgesia in two adequately stabilized mouse femoral osteotomy

models (rigid and flexible fixation). It was evaluated whether one injection of the experimental sustained-release buprenorphine ensured adequate analgesia in both osteotomy models in male and female mice. For this, the pain-relieving effects of the BUP-Depot were compared to the established pain management protocol using tramadol application via the drinking water (Wolter et al. 2023a), which has already been successfully used in mouse osteotomy models (see e.g., (Jirkof et al. 2019a)). Both protocols were evaluated for their analgesic efficacy, side effects, and potential implications on fracture healing in female and male C57BL/6N mice, as sex has been shown to influence pain perception and opioid sensitivity in laboratory mice (Wolter et al. 2023a).

In the following paragraphs, the key findings related to general parameters as well as model-specific parameters and potential side effects on fracture healing outcome are discussed. Several studies showed anesthesia- and analgesia-associated effects on multiple parameters that are often used for pain assessment in mice (Liles and Flecknell 1992; Goecke et al. 2005; Langford et al. 2010; Miller et al. 2015; Oliver et al. 2018; Jirkof et al. 2019a). Therefore, a consecutive and integrative study design was utilized to decrease overall animal numbers and to establish intra-individual baselines for the assessment of different parameters and behavioral assessments regarding anesthesia- and analgesia related effects (Wolter et al. 2023a).

General parameters: The analyses of general parameters of well-being included the assessment of the animals' body weight, the food- and water intake, nest building behavior, explorative behavior, and the assessment of the composite pain score. The reduction of food intake was in line with the general body weight development. The reduction of body weight observed after osteotomy was rather low (range of 5%) and similar to the body weight reduction observed after anesthesia and analgesia alone (Wolter et al. 2023a). This is in line with another study by Jirkof et al. that also reported a rather low impact of the osteotomy surgery on body weight and food intake compared to the effect of analgesia and anesthesia (Jirkof et al. 2019a). Effects of opioid analgesia and multimodal analgesia on body weight without painful stimuli has also been shown by other authors (Jirkof et al. 2015; Oliver et al. 2018). This highlights the need for anesthesia-only and analgesia-only control groups when evaluating pain management protocols or, potentially preferable, the use of consecutive study designs to compensate for intra-individual differences. In all groups, body weight and food intake were lowest after 24h and recovered over 48h and 72h post-anesthesia and post-osteotomy. Group differences were not observed (Wolter et al. 2023a). Even though constipation is a reported side effect of chronic opioid application (Young et al. 2018), reduced defecation was only observed at 12h after both interventions and overall defecation within each cage appeared normal. The overall water intake did not differ between groups, thereby ruling out potential

adverse implications of tramadol in the drinking water on the total water uptake. However, it must be noted that the drinking amount was reduced to around 50% in the first 24h after anesthesia/analgesia-only and after surgery (Wolter et al. 2023a). This can be critical as the analgesic supply is reduced accordingly during this period, which can be especially problematic in more painful models. Sauer et al. also showed that water intake in mice during the light period was less frequent due to circadian rhythms (Sauer et al. 2016). It has, therefore, been suggested that additional repeated injections during the light phase of the day could be beneficial in case of moderate to severe pain (Sauer et al. 2016; Arras et al. 2020). While the handling-free application, and therewith stress-free application, of oral analgesia via the drinking water is a great benefit, disadvantages include inaccurate dosing and lack of uptake control, further perpetuated by the reduction of the overall water intake after anesthesia, analgesia and surgery, altered taste and the circadian rhythm (Sauer et al. 2016; Jirkof et al. 2019a; Arras et al. 2020). In conclusion, the oral application of analgesics is a good addition or alternative to repeated injections. Nevertheless, the availability of analgesic sustained-release formulations would further help to greatly expand pain management options for mice in Europe.

In our study, nest building remained largely unaffected by anesthesia/analgesia and by osteotomy (Wolter et al. 2023a). However, these findings are contrary to the results of a previous study, where the nest building was negatively influenced by osteotomy (Jirkof et al. 2019a). These discrepancies might be explained by general differences between studies, such as the amount of nesting material or individual dissimilarities in scoring (Wolter et al. 2023a). Jirkof et al. proposed that the analyses of nest building behavior might be a helpful indicator to detect reduced well-being but also recognized that it may not be able to assess the effectiveness of pain treatment (Jirkof et al. 2013b). They suggest that reduced nest building qualities might also be the result of physiological stress or motor impairment (Jirkof et al. 2013b). However, a study by Arras et al. showed that nest building qualities were in line with aberrant heart rate and heart rate variability after laparotomy, thereby supporting the hypothesis that this parameter can be evidential of post-surgical pain (Arras et al. 2007). In summary, analyses of body weight, food- and water intake, nest building, and explorative behavior did not show differences between the two analgesic regimes and two fixations in both sexes (Wolter et al. 2023a).

The used composite pain score appeared to be a more sensitive tool. When comparing composite pain score baseline values to composite scores post-anesthesia/analgesia, our results showed that anesthesia and analgesia only already had an impact on the composite pain score (Wolter et al. 2023a). This finding is comparable with results from other studies (Miller et al. 2015; Jirkof et al. 2019a). Subsequently, a delta for each individual mouse was

calculated, in an attempt to represent the actual effect of surgery. The highest delta composite score in all groups was reached at 12h and 24h, and then gradually decreased over 48h and 72h post-osteotomy. Significant group differences were observed at 24h and 48h (Wolter et al. 2023a). Thus, the delta composite pain score seemed to reveal possible shortcomings of the tramadol treatment in male mice with the flexible fixation that were not detectable through the analyses of other general parameters, including body weight development, food-, and water intake and the analyses of nest building and explorative behavior (Wolter et al. 2023a). Although the composite score is also influenced by distress or reduced well-being (e.g., induced by anesthesia or analgesia) (Arras et al. 2007; Jirkof et al. 2019a), it still seems to be a sensitive tool to detect minor (residual) pain that might be not detected by other general indicators for pain. This again highlights the need for the use and combination of a multitude of parameters when attempting to reliably detect signs of pain in laboratory animals. It must also be noted that regular assessment of well-being and the detection of pain in animal experiments using score sheets often consist of the monitoring of body weight development, other general parameters and cage-side measures. However, our results suggest that these measures might not be able to detect subtle residual pain, thereby possibly leading to undetected pain, particularly in established pain management protocols where adequate analgesic coverage is already assumed (Wolter et al. 2023a). This can be especially critical in terms of changes in dosages or applications intervals and the use of different strains or model-specific adaptations such as varying fixation methods, where adequate pain coverage might be expected due to its effectiveness under slightly different circumstances. The critical and regular reevaluation of current pain management protocols and pain assessment tools is, therefore, important to further improve pain assessment and pain management and thereby animal welfare in laboratory animals.

The assessment of all general parameters analyzed within the study indicated a satisfactory efficacy of a single BUP-Depot injection, comparable or possibly even superior to the application of tramadol with the drinking water. Moreover, the delta composite score revealed possible shortcomings of the tramadol treatment in male mice with the flexible fixation (Wolter et al. 2023a).

Model-specific parameters: The assessment of model-specific parameters included a manual gait analysis and an automatic gait analyses system (Noldus CatWalk XT). Both revealed alterations in walking performance after osteotomy, with variations observed between groups. While manual gait analyses revealed descriptive differences in limping between the two differently fixated male groups treated with tramadol, these differences were not significant (Wolter et al. 2023a). The automatic locomotion analyses with the Noldus XT Catwalk, however, resulted in more prominent group differences, revealing deficits in male mice with

the flexible fixation and tramadol as their pain management in the relative velocity, relative mean intensity and relative stride length, compared to other groups. Spearman correlation between relative velocity and relative mean intensity as well as relative stride length showed moderate to strong correlations (Wolter et al. 2023a). Thus, the reduced speed directly impacted relative mean intensity and relative stride length. This illustrates why these two parameters revealed group differences between the male mice, while no group differences were observed with respect to relative stance duration, which correlated only very weakly with relative velocity (Wolter et al. 2023a). The Noldus CatWalk, compared to forced gait analysis on treadmills or running wheels, allows a free and voluntary movement of the animals and therewith varying velocities recordings (Batka et al. 2014; Jacobs et al. 2014; Tappe-Theodor et al. 2019). Our results highlight the need to analyze the results of this automatic locomotion analysis in relation to the relative speed (Wolter et al. 2023a). Batka et al. also underlined the need for speed evaluation in mouse gait analyses by showing that over 90% of parameters assessed by the CatWalk XT were dependent on speed (Batka et al. 2014). They also found marked intraindividual and interindividual differences in average speed on a single day as well as between animals across time (Batka et al. 2014). Failure to account for the potential effects of velocity changes can, therefore, affect the results of gait analysis (Cendelín et al. 2010; Batka et al. 2014; Wolter et al. 2023a).

CatWalk analysis has been used in many different mouse and rat models, however, to our current knowledge, it has only been used once in a fracture model, where Hofman et al. utilized the system to assess the impact of an intramedullary stabilized femur fracture in rats on ambulation and muscle atrophy during the fracture healing process (Hofman et al. 2020). Its use as a model-specific indicator for pain in fracture models is, therefore, a relatively new approach and has not been evaluated after (long-bone) fracture in mice yet. As expected, the CatWalk system allowed for a more sensitive and in-depth analysis of affected gait patterns than the manual gait analyses (limping score). However, its conduct and the following data analyses were also more time-consuming.

Analyses of both model-specific gait parameters were in line with the analyses of general parameters and suggested a good analgesic efficacy of the sustained-release buprenorphine in both fixations and sexes that might have been more sufficient than the tramadol application via the drinking water with respect to male mice with the more flexible fixation (Wolter et al. 2023a).

Overall outcome regarding analgesic efficacy: The combined assessment of general clinical, behavioral, and model-specific parameters used in the study identified that one injection of the BUP-Depot achieved sufficient analgesia for 72h after femoral osteotomy in C57BL/6N mice (Wolter et al. 2023a). The use of a delta composite pain score and an

automatic gait analysis proved to be valuable tools for the evaluation of analgesic efficacy in the mouse osteotomy model. Both parameters potentially identified slight insufficiencies in the pain treatment of tramadol application via the drinking water in male mice with the flexible fixation that were not detectable through other, commonly used, general parameters. Overall, analgesia after one BUP-Depot application appeared to be very good in both sexes and in both fixation stiffnesses, comparable to tramadol in the drinking water, and potentially superior to tramadol in the drinking water in the male group with the flexible fixation (Wolter et al. 2023a).

Fracture healing outcome: Analyses of fracture healing outcome two weeks after osteotomy using *ex-vivo* µCT and histomorphometric analysis as well as vessel staining revealed that the analgesic treatments did not have adverse effects on fracture healing. However, as expected, the two fixations resulted in dissimilar bone formation. Female and male mice from both treatment groups exhibited decreased bone formation and increased cartilage formation with the more flexible fixator as compared to the more rigid fixation (Wolter et al. 2023a). As outlined before, researchers often fear the potential implications of analgesics on their respective models and data (Mogil et al. 2020). The analysis of possible impacts of analgesics is, therefore, also especially important when evaluating new analgesic formulations for their potential use in orthopedic research. *Ex-vivo* µCT and histomorphometric analyses of fracture healing outcome indicated no significant differences between the two analgesic treatments, suggesting that the newly developed BUP-Depot can be safely employed in the mouse models studied here (Wolter et al. 2023a). This is in line with a previous study that showed that the application of buprenorphine as well as low and high doses of tramadol via the drinking water over 72h did not impact fracture healing in a stable fixated mouse osteotomy model (Jirkof et al. 2019a).

Limitations: In the presented study, the analgesic efficacy was evaluated in C57BL/6N mice, but not in other mouse strains. It has been shown that mouse strains can differ in their response to pain and analgesia (Smith 2019). However, as described by Wolter et al., the C57BL/6 mouse strain is the most commonly used mouse strain in adequately stabilized mouse fracture models (Wolter et al. 2021) and it can be proposed that the evaluation of the analgesic properties of the BUP-Depot in a C57BL/6 strain, therefore, had most value for this model. Although the consecutive study design allowed individual analysis of the effects associated with anesthesia and analgesia on the evaluated parameters, and a reduction in total animal numbers, a specific differentiation between anesthesia and analgesia-associated effects was not possible. To do so, both interventions would additionally have had to be performed and observed separately, which would most likely have resulted in a higher total animal number and in a higher cumulative burden on the animals. Furthermore, the injection site of the BUP-Depot was only examined macroscopically, but not histologically. Schreiner et al. observed

that only 2 out of 75 mice which received the sustained-release buprenorphine showed mild local reactions after injection (Schreiner et al. 2020). Although no macroscopic effects were observed in the mice receiving the BUP-Depot in this study, postmortem histological examination of the injection site may have helped to confirm the macroscopic observations and should be considered in future studies. However, as mice were euthanized 2 weeks after the last injection of the BUP-Depot, only long-term effects would have been observable.

Outlook: While the administration of analgesics via the drinking water is already a valuable alternative or addition to repeated injections, the availability of sustained-release formulations such as of buprenorphine would greatly help to further broaden the field of analgesic options in Europe. We found a promising analgesic efficacy of the novel and experimental sustained-release buprenorphine in female and male mice in two mouse osteotomy models. We propose that the availability of the BUP-Depot could have the potential to further refine analgesia in animal-based studies in Europe. However, the evaluated sustained-release buprenorphine is an experimental formulation and is currently not approved. The formulation is, therefore, not commercially available and cannot be regularly used for analgesia in mice. It can currently only be used in animal-based studies if its analgesic efficacy is evaluated within the utilized mouse model (“for research purpose only”). The process to import the BUP-Depot from Basel, Switzerland, to Germany was also highly time-consuming (>6 months) and might be a further hindrance for future testing, at least in Germany. As buprenorphine is listed in Annex III of the German Narcotics Acts [Betäubungsmittelgesetz (BtMG)], its import, storage and distribution are subjected to strict conditions. Nevertheless, further evaluation of the sustained-release buprenorphine in other mouse models is highly advisable, as the formulation showed excellent analgesic efficacy after a single injection in the two analyzed osteotomy models. The availability of a sustained-release buprenorphine formulation for regular use in animal studies would be highly desirable and could represent a major advance in pain treatment in laboratory mice in Europe. However, whether a regulatory approval of experimental formulations in Europe or the import of existing products from the USA might be realistic in the near or far future, is beyond the scope of this thesis.

7.3 Evaluation of rearing behavior as a model-specific indicator in mouse osteotomy models

For pain assessment in mouse fracture models, different general and model-specific indicators are usually combined, integrating varying clinical parameters and spontaneous behaviors. An often-used model-specific parameter in long-bone fracture models is the assessment of walking behavior. As rearing behavior has a high ambulatory component and the animals must put weight on their hind legs to get into an upright position, it is reasonable to suggest that the assessment of rearing behavior might also be a helpful additional model-specific parameter for

pain assessment after a long-bone fracture (Wolter et al. 2023b). Rearing behavior has been analyzed in a variety of contexts for over 50 years (Lever et al. 2006) (see e.g., (Van Abeelen 1970; Simiand et al. 1984; Adamah-Biassi et al. 2013)) and as a spontaneous, exploratory behavior, it is relatively easy to assess and shows high inter-experimenter correlations (Lever et al. 2006; Kendall et al. 2016). It has also been used to detect reduced well-being and possibly pain after (surgical) interventions, such as laparotomy and vasectomy (see e.g., (Wright-Williams et al. 2007; Leach et al. 2012; Kendall et al. 2016)). However, its use in mouse fracture models is still scarce and seems to have only become more common in recent years.

Use of rearing behavior analyses in mouse fracture models: Although some studies that used mouse fracture/osteotomy models analyzed rearing behavior, the analyses differed greatly between studies, including various time points and periods post-surgery and different assessment methods and settings. While some studies used manual assessment methods, resulting in short observation periods, others used automatic assessments that allowed for distinctly longer observations (Koewler et al. 2007; Das et al. 2017; Majuta et al. 2017; Majuta et al. 2018; Jirkof et al. 2019a; Magnusdottir et al. 2021). This makes comparisons between results difficult. Model-specific factors also differed markedly, with divergent fracturing and fixation methods of the femur or tibia. Major differences were also noticeable in the pain management, including vehicle (PBS) only treatment, anti-NGF treatment, a single injection of buprenorphine and application of tramadol or buprenorphine with the drinking water (see e.g., (Koewler et al. 2007; Majuta et al. 2015; Das et al. 2017; Majuta et al. 2018; Jirkof et al. 2019a; Magnusdottir et al. 2021)). It has been shown that rearing can be further influenced by a great variety of environmental and experimental external influences, including anxiety, fear, injected substances, the environment, the set-up in which rearing is measured as well as habituation and familiarization to the set-up, the light/dark phase, noise, stress and more (Lever et al. 2006). Furthermore, intrinsic factors such as sex, strain, and age have also been shown to influence rearing behavior (Van Abeelen 1970; Tang and Sanford 2005; Fahlström et al. 2011). Although some studies analyzed rearing behavior in mouse fracture models, these studies did not specifically focus on the validation of rearing behavior as pain parameter within the first 72h after fracture in mice.

Therefore, the third objective of this thesis was to investigate whether the evaluation of rearing behavior could serve as an additional model-specific indicator for pain in adequately stabilized femoral osteotomy models. For this, data collection was retrospectively performed using video material and scoring results from our own previous study of the tramadol-in-drinking-water groups (see., 5. “A Buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72h in mouse femoral fracture models”) (Wolter et al. 2023b). There is no duplication of results published in the previous study.

Assessment of rearing behavior as a potential model-specific indicator for pain in mouse osteotomy models: The general assessment of absolute and relative rearing time showed that rearing behavior in both sexes was similar between the two baseline measurements and between the three measurements after anesthesia. Thus, rearing behavior did not appear to be affected by anesthesia/analgesia at the measured timepoints (Wolter et al. 2023b). In our previous study, we showed that the impact of anesthesia and analgesia on the food and water intake as well as the body weight development were greatest at 24h (Wolter et al. 2023a), while no impact on rearing behavior was observed 24h post-anesthesia/analgesia (Wolter et al. 2023b). This could be an advantage of rearing assessment compared to the assessment of other parameters, such as food and water intake, and body weight development, especially at the 24h time point. When analyzing the influence of anesthesia and analgesia on rearing behavior at earlier time points, Jirkof et al. observed reduced rearing behavior in C57BL/6J mice at 1h and 6h post isoflurane anesthesia, compared to later timepoints (Jirkof et al. 2019a). While Miller et al. found no significant effects of isoflurane anesthesia on rearing 30 min after recovery compared to baseline measurements in CBA and DBA/2 mice, they did find that buprenorphine application led to significantly reduced rearing 45 min after the injection compared to baseline measurements in CBA mice but not DBA/2 mice (Miller et al. 2015).

Kendall et al. observed the absolute rearing time for 5-minute periods at 0, 1, 3, 6, 12, 24, 48, and 72h post-laparotomy in female Crl:CD1(ICR) mice and showed that the least amount of rearing occurred shortly after surgery and that baseline values were reached or exceeded at 72h post-surgery (Kendall et al. 2016). Miller and colleagues found that, compared to baseline values, rearing frequency in CBA mice was significantly reduced at 30 minutes and 5h post-abdominal vasectomy, but returned to baseline values at 24 hours post-surgery (Miller et al. 2016). In our study, absolute rearing duration was also significantly reduced after osteotomy in both sexes and both fixations, thereby seemingly directly depicting the effect of the surgery on rearing behavior (Wolter et al. 2023b). In contrast to the results of Kendall et. al and Miller et al. (Kendall et al. 2016; Miller et al. 2016), in our study, absolute rearing time in female mice did not recover between 24h and 72h, whereas male mice showed some recovery in absolute rearing time over 72h (Wolter et al. 2023b). However, when assessing the recovery of relative rearing time (baseline/osteotomy), significant differences in both sexes compared to the baseline were still apparent at 72h after osteotomy (Wolter et al. 2023b), therefore, showing less recovery over 72h after osteotomy than described by Kendall and colleagues after laparotomy (Kendall et al. 2016). This is explicable by the different models and the accompanying impairment of the musculoskeletal system in the mouse-osteotomy model. In contrast to other parameters used for the assessment of well-being and pain, such as body weight development, food- and water intake, and the pain composite score, rearing behavior

remained reduced and did not recover to baseline values after 72h (Wolter et al. 2023b). This ongoing reduction of rearing behavior might, therefore, depict expectable surgery-related deficits in the skeletal system and ambulation rather than residual pain (Wolter et al. 2023b). Although the results of our previous study suggested that pain management in the male mice with flexible fixation and tramadol treatment may have led to mild residual pain in this group (Wolter et al. 2023a), rearing behavior between the two male fixation groups did not differ (Wolter et al. 2023b), thereby suggesting that rearing behavior analyses might not be sensitive enough to detect mild residual pain. This is in line with findings from Kendall et al., who showed that although the application of buprenorphine did not appear to provide sufficient analgesia in their used laparotomy model, they also found no differences in absolute rearing duration between the three different treatment groups (Kendall et al. 2016). In contrast, Majuta et al. observed significant differences in rearing behavior in the exploratory phase between vehicle (PBS) treated mice and mice treated with anti-NGF after pin-stabilized femur fracture at 1d and 3d post-surgery (Majuta et al. 2018). It therefore might be possible that the analysis of rearing behavior is able to detect differences when the pain coverage is distinctly unequal.

We found that the relative rearing time was significantly reduced compared to the baseline values independent of sex, fixation, and gait (limping), but overall, our analyses could not identify a coherence between relative rearing time and the presence or absence of limping. In our previous study, we proposed that persistent gait alteration over 72h or 10d may be due to incomplete functional recovery after surgery rather than pain, as there were no additional indications of pain observed at 72h post-surgery (Wolter et al. 2023a). This could, therefore, also be the case for a reduced rearing at 10d post-surgery. It could also be possible that familiarization with the environment and study set-up might play a role in the reduced rearing frequency at 10d post-osteotomy, as other studies provide evidence that familiarization with the environment reduces rearing behavior (Lever et al. 2006; Durst et al. 2021a).

Limitations and outlook: Due to its time-consuming nature, manual assessment of rearing behavior in our study only allowed for short observation periods outside the home cage and during the light phase of the day. Majuta et al. showed that greatest activity and greatest differences between groups were found in first the 30 - 60 minutes after the mice were placed in the observation boxes (exploratory phase) and in the first three hours of darkness (Majuta et al. 2018). Fahlström et al. also showed that the differences in rearing frequency between 3-, 8- and 28-month-old mice were most striking in the first 30 minutes (= explorative phase) (Fahlström et al. 2011). Even though these findings underline the relevance and benefits of observing spontaneous activity such as rearing during the night phase, where nocturnal animals are the most active (Majuta et al. 2018), it also demonstrates that analyses during the exploratory phase shows promising capacities to highlight group differences. Therefore, we

conclude that using the exploratory phase for manual assessment of rearing behavior, as done in our study, seems to be a meaningful approach. In terms of future analyses of rearing behavior, manual and automatic assessments appear to be reasonable approaches depending on the research question. However, especially the time-consuming nature of manual rearing behavior analyses makes it unsuitable for routine real-time monitoring of well-being or pain. It appears to be better suited for a retrospective analysis, as used in this study. In the present study, data were analyzed by manual assessment only. However, the comparison of manual analyses to the results gained by automatic assessments could be an interesting line of inquiry for future studies. In this study, rearing behavior after osteotomy was only assessed in C57BL/6N mice. Strain differences in rearing behavior have been well described. Back in 1970, van Abeelen found that C57BL mice generally reared more than DBA mice, while C57BL also showed higher variances in rearing than DBA mice (van Abeelen 1970). Since then, strain differences in rearing have been reported in multiple studies (Tang and Sanford 2005; Adamah-Biassi et al. 2013; Delprato et al. 2017). Analyses of rearing behavior after osteotomy in relation to pain assessment in another strain, therefore, might lead to differing results. However, as proposed before, the C57BL/6 strain is the most used mouse strain in adequately stabilized femur fracture models (Wolter et al. 2021) and the analyses of rearing behavior in this strain is, therefore, suitable for the proposed research question. We propose that future studies using fracture or osteotomy models of long bones in mice could integrate analyses of rearing to further validate its potential practical and informative value as a model-specific assessment tool in fracture-related mouse models.

8 Summary

Preclinical research using animal models is still imperative in biomedical sciences to address the unmet clinical needs in many areas such as fracture healing disorders. General ethical and legal requirements must be carefully considered to provide scientific quality and integrity in animal-based studies. Sufficient reporting, adequate pain assessment and adapted pain management in animal-based research are crucial to generate reliable, reproducible, and comparable data, in line with the humane use of animals in research.

However, ongoing discussions in the scientific community and various systematic literature analyses have indicated that the overall reporting quality in many areas of animal-based biomedical research is still inadequate. Therefore, the first aim of this thesis was to evaluate the status quo of the reporting accuracy of different general and model-specific parameters and the respective analgesia regimes in studies using adequately stabilized mouse femoral fracture models. High reporting accuracies were identified for the used fixation methods and fracturing procedures as well as the used strain, age, and sex of the mice, while insufficient reporting was found regarding the substrain and genetic background of the mice, body weight, hygiene monitoring, anesthesia, and analgesia. The systematic review also showed marked differences in the route, timepoint and duration of application as well as analgesic dosages, further indicating a need for more evidence-based data on model-adapted pain management protocols.

The adequate pain assessment and an adapted stress-reduced pain management in laboratory rodents are ongoing challenges, yet essential to improve animal welfare and increase scientific validity and reproducibility. To reduce handling-associated stress while covering the period of greatest pain, the application of analgesics via the drinking water or the injection of sustained-release preparations can be used as an addition or an alternative to repeated injections. However, no sustained-release preparation for pain management in mice is currently available on the European market. The introduction of such a sustained-release buprenorphine would, therefore, be of great benefit for extended pain relief in a variety of different mouse models. Thus, the second part of the thesis focused on the evaluation of the analgesic efficacy of a newly developed depot formulation of buprenorphine (BUP-Depot) for prolonged post-operative pain relief in two mouse osteotomy models. Following a single injection, the analyses of various general and model-specific parameters yielded promising pain-relieving properties of the BUP-Depot for up to 72h post-surgical in female and male mice in both models. The availability of such a sustained-release formulation of buprenorphine would greatly help to further broaden the field of analgesic options in Europe and could have the potential to further refine analgesia in animal-based studies in Europe.

In the third part of the thesis, rearing behavior was evaluated to serve as an additional and meaningful model-specific indicator for pain in the first 72h after femoral osteotomy in mice. The analyzed data indicated significantly reduced rearing behavior after osteotomy, independent of sex or fixation, while anesthesia and analgesia alone did not impact rearing behavior. However, there was no clear evidence that the analysis of rearing behavior could serve as a meaningful sole indicator for residual model-specific skeletal pain in the analyzed models. Thus, further studies are needed to evaluate the value and predictive capacity of rearing behavior to assess pain in skeletal-impaired mouse models.

Overall, this thesis demonstrates that the accuracy of reporting in mouse fracture models needs to be further improved and that guidelines such as the ARRIVE guidelines should be more actively addressed in future studies. Furthermore, the data demonstrated safe use and efficient analgesia of a newly developed sustained-release buprenorphine for 72h after femoral osteotomy and underline the need and relevance of the availability of sustained-release analgesics for refined analgesia in laboratory mice in Europe. Lastly, this thesis indicates an ongoing need for further in-depth analysis and evaluation of easy-to-use and reliable indicators for pain in laboratory mice to ensure animal welfare.

9 Zusammenfassung

Schmerzerkennung und –management in Maus-Femur-Fraktur-Modellen

Aktuell bleibt die präklinische Forschung im Tiermodell in den biomedizinischen Wissenschaften nach wie vor unverzichtbar, um den klinischen Bedarf in vielen Bereichen wie zum Beispiel bei Frakturheilungsstörungen zu decken. Um dabei die wissenschaftliche Qualität und Integrität tiergestützter Studien zu gewährleisten, müssen die allgemeinen ethischen und rechtlichen Anforderungen sorgfältig berücksichtigt werden. Ein umfassendes Reporting, eine differenzierte Schmerzbeurteilung und eine angepasste Schmerzbehandlung sind dabei von entscheidender Bedeutung, um verlässliche, reproduzierbare und vergleichbare Daten aus tiergestützten Studien zu generieren.

Anhaltende Diskussionen in der wissenschaftlichen Gemeinschaft und zahlreiche systematische Literaturanalysen haben jedoch gezeigt, dass die Qualität des Reportings in vielen Bereichen der tiergestützten biomedizinischen Forschung noch immer unzureichend ist. Daher wurde in dieser Arbeit zunächst der aktuelle Stand der Reporting-Qualität in Bezug auf verschiedene allgemeine und modellspezifische Parameter sowie die eingesetzten Analgetika-Pläne in Studien mit stabilisierten Maus-Femur-Fraktur-Modellen erfasst. Die Auswertung ergab eine hohe Reporting-Qualität hinsichtlich der eingesetzten Fraktur- und Fixierungsmethoden und sowie in Bezug auf den Stamm, das Alter und das Geschlecht der verwendeten Tiere. Das Reporting der Substämme und des genetischen Hintergrundes, des Körpergewichts, der Hygieneüberwachung, der Anästhesie und Analgesie war hingegen unzureichend. Das systematische Review zeigte außerdem deutliche Unterschiede in Bezug auf den Applikationsweg, -zeitpunkt und die -dauer sowie die Dosierungen der verwendeten Analgetika auf und unterstreicht damit einen weiterhin bestehenden Bedarf an evidenzbasierten Daten für modellspezifische Schmerzbehandlungsprotokolle in der Maus.

Die differenzierte Schmerzbeurteilung und ein angepasstes und stressreduziertes Schmerzmanagement bei Labornagern stellen eine fortwährende Herausforderung dar, sind für die Verbesserung des Tierschutzes und die Erhöhung der wissenschaftlichen Validität und Reproduzierbarkeit jedoch unerlässlich. Methoden wie die Verabreichung von Analgetika über das Trinkwasser oder die Injektion von Depotpräparaten mit verzögerter Wirkstofffreisetzung können genutzt werden, um den Stress durch wiederholte Injektionen zu reduzieren und gleichzeitig den Zeitraum der zu erwartenden Schmerzen abzudecken. Derzeit ist jedoch kein Depotpräparat für die verlängerte postoperative Schmerzlinderung bei Mäusen auf dem europäischen Markt erhältlich und die Einführung eines Buprenorphins mit verzögerter Wirkstofffreisetzung in Europa würde sich daher in einer Vielzahl von Mausmodellen als

überaus nützlich erweisen. Daher befasste sich der zweite Teil der Arbeit mit der Bewertung der analgetischen Wirksamkeit eines neu entwickelten Depot-Buprenorphins (BUP-Depot) in zwei Maus-Osteotomie-Modellen. Die umfassende Auswertung unterschiedlicher allgemeiner und modellspezifischer Faktoren ergab, dass das BUP-Depot bei weiblichen und männlichen Mäusen und in beiden Modellen bis zu 72 Stunden post-operativ vielversprechende schmerzlindernde Eigenschaften aufwies. Die Verfügbarkeit eines solchen Buprenorphins mit verzögerter Wirkstofffreisetzung könnte erheblich zur Erweiterung der verfügbaren analgetischen Optionen beitragen und das Potenzial besitzen, die Analgesie in tiergestützten Studien in der Maus in Europa weiter zu verbessern.

Im dritten Teil dieser Arbeit wurde das Rearing-Verhalten (Aufrichten auf die Hinterbeine) als potenziell aussagekräftiger modellspezifischer Indikator für die Schmerzbeurteilung in den ersten 3 Tagen nach Femur-Osteotomie in der Maus untersucht. Die analysierten Daten wiesen auf ein signifikant reduziertes Rearing-Verhalten nach Osteotomie hin, unabhängig von Geschlecht oder Fixierung, während die Anästhesie und Analgesie allein keinen Einfluss auf das Rearing-Verhalten zu haben schienen. Es gab jedoch keine eindeutigen Hinweise darauf, dass die Analyse des Rearing-Verhaltens als alleiniger Indikator für modellspezifische muskuloskelettale Schmerzen in den untersuchten Modellen hilfreich ist. Zukünftig sind daher weitere Studien notwendig, um den Nutzen und die Aussagekraft des Rearing-Verhaltens zur Beurteilung von muskuloskelettalen Schmerzen in Maus-Osteotomie-Modellen konkret zu bewerten und einzuordnen.

Insgesamt zeigt diese Arbeit, dass die Genauigkeit des Reportings in Maus-Fraktur-Modellen weiter verbessert werden muss und dass Leitlinien, wie die ARRIVE-Guidelines, in Zukunft aktiver berücksichtigt und angewendet werden sollten. Darüber hinaus unterstreichet die Arbeit die Notwendigkeit, Relevanz und Vorteile der Verfügbarkeit von Analgetika mit verzögerter Wirkstofffreisetzung für die Anwendung bei Labormäusen in Europa. Die Auswertungen ergaben eine sichere Anwendung und effektive Schmerzlinderungen eines neu entwickelten Buprenorphin Depot Präparates in männlichen und weiblichen Mäusen für 72 Stunden nach Osteotomie. Weiterhin verdeutlicht die Arbeit die bestehende Notwendigkeit für die detaillierte Analyse und Beurteilung von möglichst verlässlichen und anwendungs-freundlichen Indikatoren zur Erkennung von Schmerzen bei der Labormaus, um so den Tierschutz zu gewährleisten.

10 References

Adamah-Biassi E. B., Stepien I., Hudson R. L. und Dubocovich M. L. (2013):

Automated video analysis system reveals distinct diurnal behaviors in C57BL/6 and C3H/HeN mice.

Behav Brain Res 243: 306-312.

Aerssens J., Boonen S., Lowet G. und Dequeker J. (1998):

Interspecies Differences in Bone Composition, Density, and Quality: Potential Implications for in Vivo Bone Research*.

Endocrinology 139: 663-670.

Akhter M. P., Cullen D. M., Pedersen E. A., Kimmel D. B. und Recker R. R. (1998):

Bone response to in vivo mechanical loading in two breeds of mice.

Calcif Tissue Int 63: 442-449.

Akhter M. P., Fan Z. und Rho J. Y. (2004):

Bone intrinsic material properties in three inbred mouse strains.

Calcif Tissue Int 75: 416-420.

Al-Hasani R. und Bruchas M. R. (2011):

Molecular mechanisms of opioid receptor-dependent signaling and behavior.

Anesthesiology 115: 1363-1381.

Allen M. und Johnson R. A. (2018):

Evaluation of self-injurious behavior, thermal sensitivity, food intake, fecal output, and pica after injection of three buprenorphine formulations in rats (Rattus norvegicus).

American Journal of Veterinary Research 79: 697-703.

Amit Z. und Galina Z. H. (1988):

Stress Induced Analgesia plays an adaptive role in the organization of behavioral responding.

Brain Research Bulletin 21: 955-958.

Aronson J. K. (2009):

Meyler's Side Effects of Analgesics and Anti-inflammatory Drugs. Elsevier Science – ISBN 9780444532732.

Arras M., Becker K., Bergadano A., Durst M., Eberspächer-Schweda E., Fleischmann T., Haberstroh J., Jirkof P., Sager M., Spadavecchia C. und Zahner D. (2020):

Expert information from the GV-SOLAS Committee for Anaesthesia in collaboration with Working Group 4 in the TVT: Pain management for laboratory animals.

Arras M., Rettich A., Cinelli P., Kasermann H. P. und Burki K. (2007):

Assessment of post-laparotomy pain in laboratory mice by telemetric recording of heart rate and heart rate variability.

BMC Vet Res 3: 16.

Auer J. A., Goodship A., Arnoczky S., Pearce S., Price J., Claes L., von Rechenberg B., Hofmann-Amtenbrinck M., Schneider E., Müller-Terpitz R., Thiele F., Rippe K.-P. und Grainger D. W. (2007):

Refining animal models in fracture research: seeking consensus in optimising both animal welfare and scientific validity for appropriate biomedical use.

BMC Musculoskeletal Disorders 8: 72.

Aulehner K., Leenaars C., Buchecker V., Stirling H., Schönhoff K., King H., Häger C., Koska I., Jirkof P., Bleich A., Bankstahl M. und Potschka H. (2022):

Grimace scale, burrowing, and nest building for the assessment of post-surgical pain in mice and rats-A systematic review.

Front Vet Sci 9: 930005.

Avey M. T., Fenwick N. und Griffin G. (2015):

The use of systematic reviews and reporting guidelines to advance the implementation of the 3Rs.

J Am Assoc Lab Anim Sci 54: 153-162.

Baker D., Lidster K., Sottomayor A. und Amor S. (2014):

Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Pre-Clinical Animal Studies.

PLOS Biology 12: e1001756.

Balcombe J. P., Barnard N. D. und Sandusky C. (2004):

Laboratory routines cause animal stress.

Contemp Top Lab Anim Sci 43: 42-51.

Barrot M. (2012):

Tests and models of nociception and pain in rodents.

Neuroscience 211: 39-50.

Batka R. J., Brown T. J., McMillan K. P., Meadows R. M., Jones K. J. und Haulcomb M. M. (2014):

The need for speed in rodent locomotion analyses.

Anat Rec (Hoboken) 297: 1839-1864.

Baumans V., Brain, P.F., Brugere, H., Clausing, P., Jeneskog, T., Perretta, G. (1994):

Pain and distress in laboratory rodents and lagomorphs. Report of the Federation of European Laboratory Animal Science Associations (FELASA) Working Group on Pain and Distress accepted by the FELASA Board of Management November 1992.

Lab Anim 28: 97-112.

Benfenati E., Gini G., Hoffmann S. und Luttik R. (2010):

Comparing in vivo, in vitro and in silico methods and integrated strategies for chemical assessment: problems and prospects.

Altern Lab Anim 38: 153-166.

Benyamin R., Trescot A. M., Datta S., Buenaventura R., Adlaka R., Sehgal N., Glaser S. E. und Vallejo R. (2008):

Opioid complications and side effects.

Pain Physician 11: S105-120.

Bigham Sadegh A. und Oryan A. (2014):

Basic concepts regarding fracture healing and the current options and future directions in managing bone fractures: Bone healing biology.

International wound journal 12.

Birkhold A. I., Razi H., Duda G. N., Weinkamer R., Checa S. und Willie B. M. (2014):

Mineralizing surface is the main target of mechanical stimulation independent of age: 3D dynamic in vivo morphometry.

Bone 66: 15-25.

Bodnar R. J. und Kest B. (2010):

Sex differences in opioid analgesia, hyperalgesia, tolerance and withdrawal: central mechanisms of action and roles of gonadal hormones.

Horm Behav 58: 72-81.

Bramhall M., Flórez-Vargas O., Stevens R., Brass A. und Cruickshank S. (2015):

Quality of methods reporting in animal models of colitis.

Inflamm Bowel Dis 21: 1248-1259.

BtMG.

Betäubungsmittelgesetz in der Fassung der Bekanntmachung vom 1. März 1994 (BGBl. I S. 358), das zuletzt durch Artikel 1 der Verordnung vom 16. Dezember 2022 (BAnz AT 20.12.2022 V1) geändert worden ist.

Retrieved 18.11.2022, from https://www.gesetze-im-internet.de/btmg_1981/BJNR106810981.html.

Bucher C. H., Berkmann J. C., Burkhardt L.-M., Paschke C., Schlundt C., Lang A., Wolter A., Damerau A., Geissler S., Volk H.-D., Duda G. N. und Schmidt-Bleek K. (2022):

Local immune cell contributions to fracture healing in aged individuals – A novel role for interleukin 22.

Experimental & Molecular Medicine 54: 1262-1276.

Bucher C. H., Schlundt C., Wulsten D., Sass F. A., Wendler S., Ellinghaus A., Thiele T., Seemann R., Willie B. M., Volk H. D., Duda G. N. und Schmidt-Bleek K. (2019):

Experience in the Adaptive Immunity Impacts Bone Homeostasis, Remodeling, and Healing.

Front Immunol 10: 797.

Bundesinstitut für Risikobewertung (2021).

Zahlen zu den im Jahr 2020 verwendeten Versuchstieren.

Retrieved 23.12.2022, from https://www.bf3r.de/cm/343/zahlen-zu-den-im-jahr-2020-verwendeten-versuchstieren.pdf.

Callewaert F., Sinnesael M., Gielen E., Boonen S. und Vanderschueren D. (2010):

Skeletal sexual dimorphism: relative contribution of sex steroids, GH–IGF1, and mechanical loading.

Journal of Endocrinology 207: 127-134.

Carbone E. T., Lindstrom K. E., Diep S. und Carbone L. (2012):

Duration of action of sustained-release buprenorphine in 2 strains of mice.

J Am Assoc Lab Anim Sci 51: 815-819.

Carbone L. (2012):

Pain management standards in the eighth edition of the Guide for the Care and Use of Laboratory Animals.

J Am Assoc Lab Anim Sci 51: 322-328.

Carbone L. und Austin J. (2016):

Pain and Laboratory Animals: Publication Practices for Better Data Reproducibility and Better Animal Welfare.

PLoS One 11: e0155001.

Carstens E. und Moberg G. P. (2000):

Recognizing pain and distress in laboratory animals.

Ilar j 41: 62-71.

Cendelín J., Voller J. und Vozeh F. (2010):

Ataxic gait analysis in a mouse model of the olivocerebellar degeneration.

Behav Brain Res 210: 8-15.

Chapman K. L., Holzgrefe H., Black L. E., Brown M., Chellman G., Copeman C., Couch J., Creton S., Gehen S., Hoberman A., Kinter L. B., Madden S., Mattis C., Stemple H. A. und Wilson S. (2013):

Pharmaceutical toxicology: designing studies to reduce animal use, while maximizing human translation.

Regul Toxicol Pharmacol 66: 88-103.

Claes L. E., Heigele C. A., Neidlinger-Wilke C., Kaspar D., Seidl W., Margevicius K. J. und Augat P. (1998):

Effects of mechanical factors on the fracture healing process.

Clin Orthop Relat Res: S132-147.

Clark D., Nakamura M., Miclau T. und Marcucio R. (2017):

Effects of Aging on Fracture Healing.

Curr Osteoporos Rep 15: 601-608.

Clark T. S., Clark D. D. und Hoyt R. F., Jr. (2014):

Pharmacokinetic comparison of sustained-release and standard buprenorphine in mice.

J Am Assoc Lab Anim Sci 53: 387-391.

Coates B. A., McKenzie J. A., Buettmann E. G., Liu X., Gontarz P. M., Zhang B. und Silva M. J. (2019):

Transcriptional profiling of intramembranous and endochondral ossification after fracture in mice.

Bone 127: 577-591.

Connolly C. K., Li G., Bunn J. R., Mushipe M., Dickson G. R. und Marsh D. R. (2003):

A reliable externally fixated murine femoral fracture model that accounts for variation in movement between animals.

J Orthop Res 21: 843-849.

Conour L. A., Murray K. A. und Brown M. J. (2006):

Preparation of animals for research--issues to consider for rodents and rabbits.

Ilar j 47: 283-293. .

Cottrell J. und O'Connor J. P. (2010):

Effect of Non-Steroidal Anti-Inflammatory Drugs on Bone Healing.

Pharmaceuticals (Basel, Switzerland) 3: 1668-1693.

Dabis J., Templeton-Ward O., Lacey A. E., Narayan B. und Trompeter A. (2017):

The history, evolution and basic science of osteotomy techniques.

Strategies in trauma and limb reconstruction 12: 169-180.

Das V., Kroin J. S., Moric M. und Buvanendran A. (2017):

Biochemical and Pharmacological Characterization of a Mice Model of Complex Regional Pain Syndrome.

Reg Anesth Pain Med 42: 507-516.

Deuis J. R., Dvorakova L. S. und Vetter I. (2017):

Methods Used to Evaluate Pain Behaviors in Rodents.

Front Mol Neurosci 10: 284.

Dimitriou R., Tsiridis E. und Giannoudis P. V. (2005):

Current concepts of molecular aspects of bone healing.

Injury 36: 1392-1404.

Durst M., Graf T. R., Graf R., Kron M., Arras M., Zechner D., Palme R., Talbot S. R. und Jirkof P. (2021a):

Analysis of Pain and Analgesia Protocols in Acute Cerulein-Induced Pancreatitis in Male C57BL/6 Mice.

Frontiers in Physiology 12.

Durst M. S., Arras M., Palme R., Talbot S. R. und Jirkof P. (2021b):

Lidocaine and bupivacaine as part of multimodal pain management in a C57BL/6J laparotomy mouse model.

Scientific Reports 11: 10918.

Eckenhoff B. und Yum S. I. (1981):

The osmotic pump: novel research tool for optimizing drug regimens.

Biomaterials 2: 89-97.

Einhorn T. A. und Gerstenfeld L. C. (2015):

Fracture healing: mechanisms and interventions.

Nat Rev Rheumatol 11: 45-54.

European Parliament and the Council of the European Union (2010).

Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes.

Retrieved 18.11.2022, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:en:PDF.

Evangelista-Vaz R., Bergadano A., Arras M. und Jirkof P. D. (2018b):

Analgesic Efficacy of Subcutaneous-Oral Dosage of Tramadol after Surgery in C57BL/6J Mice.

Journal of the American Association for Laboratory Animal Science : JAALAS 57: 368-375.

Evangelista Vaz R., Draganov D. I., Rapp C., Avenel F., Steiner G., Arras M. und Bergadano A. (2018a):

Preliminary pharmacokinetics of tramadol hydrochloride after administration via different routes in male and female B6 mice.

Vet Anaesth Analg 45: 111-122.

Fahlström A., Yu Q. und Ulfhake B. (2011):

Behavioral changes in aging female C57BL/6 mice.

Neurobiol Aging 32: 1868-1880.

Ferreira N. und Marais L. C. (2012):

Prevention and management of external fixator pin track sepsis.

Strategies Trauma Limb Reconstr 7: 67-72.

Flecknell P. (2002):

Replacement, Reduction, Refinement.

ALTEX - Alternatives to animal experimentation 19: 73-78.

Foley P. L., Liang H. und Crichlow A. R. (2011):

Evaluation of a sustained-release formulation of buprenorphine for analgesia in rats.

J Am Assoc Lab Anim Sci 50: 198-204.

Food and Drug Administration (2019).

Unapproved Animal Drugs.

Retrieved 18.04.2023, from https://www.fda.gov/animal-veterinary/compliance-enforcement/unapproved-animal-drugs.

Food and Drug Administration (2022).

The Index of Legally Marketed Unapproved New Animal Drugs for Minor Species.

Retrieved 23.12.2022, from https://www.fda.gov/animal-veterinary/minor-useminor-species/index-legally-marketed-unapproved-new-animal-drugs-minor-species.

Franco N. H., Sandøe P. und Olsson I. A. S. (2018):

Researchers' attitudes to the 3Rs—An upturned hierarchy?

PLOS ONE 13: e0200895.

Gaskill B. N., Karas A. Z., Garner J. P. und Pritchett-Corning K. R. (2013):

Nest building as an indicator of health and welfare in laboratory mice.

J Vis Exp: 51012.

Gaston M. S. und Simpson A. H. (2007):

Inhibition of fracture healing.

J Bone Joint Surg Br 89: 1553-1560.

Gerstenfeld L. C., Cullinane D. M., Barnes G. L., Graves D. T. und Einhorn T. A. (2003):

Fracture healing as a post-natal developmental process: molecular, spatial, and temporal aspects of its regulation.

J Cell Biochem 88: 873-884.

Ghosal S., Nunley A., Mahbod P., Lewis A. G., Smith E. P., Tong J., D'Alessio D. A. und Herman J. P. (2015):

Mouse handling limits the impact of stress on metabolic endpoints.

Physiol Behav 150: 31-37.

Gibon E., Lu L. und Goodman S. B. (2016):

Aging, inflammation, stem cells, and bone healing.

Stem Cell Research & Therapy 7: 44.

Gkrouzoudi A., Tsingotjidou A. und Jirkof P. (2022):

A systematic review on the reporting quality in mouse telemetry implantation surgery using electrocardiogram recording devices.

Physiology & Behavior 244: 113645.

Glatt V., Canalis E., Stadmeyer L. und Bouxsein M. L. (2007):

Age-related changes in trabecular architecture differ in female and male C57BL/6J mice.

J Bone Miner Res 22: 1197-1207.

Goecke J. C., Awad H., Lawson J. C. und Boivin G. P. (2005):

Evaluating postoperative analgesics in mice using telemetry.

Comp Med 55: 37-44.

Gomes P. S. und Fernandes M. H. (2011):

Rodent models in bone-related research: the relevance of calvarial defects in the assessment of bone regeneration strategies.

Lab Anim 45: 14-24.

Gouveia K. und Hurst J. (2013):

Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels.

PloS one 8: e66401.

Gouveia K. und Hurst J. L. (2017):

Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling.

Scientific Reports 7: 44999.

Grieco F., Bernstein B. J., Biemans B., Bikovski L., Burnett C. J., Cushman J. D., van Dam E. A., Fry S. A., Richmond-Hacham B., Homberg J. R., Kas M. J. H., Kessels H. W., Koopmans B., Krashes M. J., Krishnan V., Logan S., Loos M., McCann K. E., Parduzi Q., Pick C. G., Prevot T. D., Riedel G., Robinson L., Sadighi M., Smit A. B., Sonntag W., Roelofs R. F., Tegelenbosch R. A. J. und Noldus L. (2021):

Measuring Behavior in the Home Cage: Study Design, Applications, Challenges, and Perspectives.

Front Behav Neurosci 15: 735387.

Gröngröft I., Heil P., Matthys R., Lezuo P., Tami A., Perren S., Montavon P. und Ito K. (2009):

Fixation compliance in a mouse osteotomy model induces two different processes of bone healing but does not lead to delayed union.

J Biomech 42: 2089-2096.

Grün N., Hahn D., Okutan B., Marek R. und Weinberg A.-M. (2019):

Animal Models in Orthopedic Research: The Proper Animal Model to Answer Fundamental Questions on Bone Healing Depending on Pathology and Implant Material.

Animal Models in Medicine and Biology. E. Tvrdá, S. C. Yenisetti and IntechOpen. 1: 23-42.

ISBN: 978-1-83880-011-6

Haas N. P. (2000):

Callusmodulation – Fiktion oder Realität?

Der Chirurg 71: 987-988.

Haffner-Luntzer M., Fischer V. und Ignatius A. (2021):

Differences in Fracture Healing Between Female and Male C57BL/6J Mice.

Front Physiol 12: 712494.

Haffner-Luntzer M., Hankenson K. D., Ignatius A., Pfeifer R., Khader B. A., Hildebrand F., van Griensven M., Pape H. C. und Lehmicke M. (2019):

Review of Animal Models of Comorbidities in Fracture-Healing Research.

J Orthop Res 37: 2491-2498.

Haffner-Luntzer M., Kovtun A., Rapp A. E. und Ignatius A. (2016):

Mouse Models in Bone Fracture Healing Research.

Current Molecular Biology Reports 2: 101-111.

Hauser M., Siegrist M., Keller I. und Hofstetter W. (2018):

Healing of fractures in osteoporotic bones in mice treated with bisphosphonates - A transcriptome analysis.

Bone 112: 107-119.

Hawkins P., Armstrong R., Boden T., Garside P., Knight K., Lilley E., Seed M., Wilkinson M. und Williams R. O. (2015):

Applying refinement to the use of mice and rats in rheumatoid arthritis research.

Inflammopharmacology 23: 131-150.

Healy J. R., Tonkin J. L., Kamarec S. R., Saludes M. A., Ibrahim S. Y., Matsumoto R. R. und Wimsatt J. H. (2014):

Evaluation of an improved sustained-release buprenorphine formulation for use in mice.

Am J Vet Res 75: 619-625.

Hirst A., Knight C., Hirst M., Dunlop W. und Akehurst R. (2016):

Tramadol and the risk of fracture in an elderly female population: a cost utility assessment with comparison to transdermal buprenorphine.

Eur J Health Econ 17: 217-227.

Histing T., Garcia P., Holstein J. H., Klein M., Matthys R., Nuetzi R., Steck R., Laschke M. W., Wehner T., Bindl R., Recknagel S., Stuermer E. K., Vollmar B., Wildemann B., Lienau J., Willie B., Peters A., Ignatius A., Pohlemann T., Claes L. und Menger M. D. (2011):

Small animal bone healing models: standards, tips, and pitfalls results of a consensus meeting.

Bone 49: 591-599.

Histing T., Holstein J. H., Garcia P., Matthys R., Kristen A., Claes L., Menger M. D. und Pohlemann T. (2009):

Ex vivo analysis of rotational stiffness of different osteosynthesis techniques in mouse femur fracture.

J Orthop Res 27: 1152-1156.

Hobbiesiefken U., Mieske P., Lewejohann L. und Diederich K. (2021):

Evaluation of different types of enrichment - their usage and effect on home cage behavior in female mice.

PLoS One 16: e0261876.

Hofman M., Kolejewska A., Greven J., Andruszkow H., Kobbe P., Tolba R., Hildebrand F. und Poeze M. (2020):

Gait analysis and muscle weight analysis after lower extremity fractures in a small animal model.

Gait & Posture 77: 207-213.

Hohlbaum K., Bert B., Dietze S., Palme R., Fink H. und Thöne-Reineke C. (2018):

Impact of repeated anesthesia with ketamine and xylazine on the well-being of C57BL/6JRj mice.

PLoS One 13: e0203559.

Holstein J. H., Menger M. D., Culemann U., Meier C. und Pohlemann T. (2007):

Development of a locking femur nail for mice.

J Biomech 40: 215-219.

Hurst J. L. und West R. S. (2010):

Taming anxiety in laboratory mice.

Nature Methods 7: 825-826.

Iliff S. A. (2002):

An Additional “R”: Remembering the Animals.

ILAR Journal 43: 38-47.

Jacobs B. Y., Kloefkorn H. E. und Allen K. D. (2014):

Gait analysis methods for rodent models of osteoarthritis.

Curr Pain Headache Rep 18: 456.

Janas A. und Folwarczna J. (2017):

Opioid receptor agonists may favorably affect bone mechanical properties in rats with estrogen deficiency-induced osteoporosis.

Naunyn-Schmiedeberg's archives of pharmacology 390: 175-185.

Ji C. C., Liu B., Shao Y. J., Liang T., Jiang H. Y., Chen G. D. und Luo Z. P. (2020):

Microstructure and mechanical properties of subchondral bone are negatively regulated by tramadol in osteoarthritis in mice.

Biosci Rep 40.

Jirkof P. (2015):

Effects of experimental housing conditions on recovery of laboratory mice.

Lab Anim (NY) 44: 65-70.

Jirkof P. (2017):

Side effects of pain and analgesia in animal experimentation.

Lab Animal 46: 123-128.

Jirkof P., Cesarovic N., Rettich A. und Arras M. (2013a):

Housing of female mice in a new environment and its influence on post-surgical behaviour and recovery.

Applied Animal Behaviour Science 148: 209-217.

Jirkof P., Durst M., Klopfleisch R., Palme R., Thöne-Reineke C., Buttgereit F., Schmidt-Bleek K. und Lang A. (2019a):

Administration of Tramadol or Buprenorphine via the drinking water for post-operative analgesia in a mouse-osteotomy model.

Sci Rep 9: 10749.

Jirkof P., Fleischmann T., Cesarovic N., Rettich A., Vogel J. und Arras M. (2013b):

Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring.

Lab Anim 47: 153-161.

Jirkof P. und Potschka H. (2021):

Effects of Untreated Pain, Anesthesia, and Analgesia in Animal Experimentation.

Jirkof P., Rudeck J. und Lewejohann L. (2019b):

Assessing Affective State in Laboratory Rodents to Promote Animal Welfare-What Is the Progress in Applied Refinement Research?

Animals (Basel) 9.

Jirkof P., Tourvieille A., Cinelli P. und Arras M. (2015):

Buprenorphine for pain relief in mice: repeated injections vs sustained-release depot formulation.

Lab Anim 49: 177-187.

Kazmers N. H., Fragomen A. T. und Rozbruch S. R. (2016):

Prevention of pin site infection in external fixation: a review of the literature.

Strategies Trauma Limb Reconstr 11: 75-85.

Kendall L. V., Wegenast D. J., Smith B. J., Dorsey K. M., Kang S., Lee N. Y. und Hess A. M. (2016):

Efficacy of Sustained-Release Buprenorphine in an Experimental Laparotomy Model in Female Mice.

J Am Assoc Lab Anim Sci 55: 66-73.

Keraliya R. A., Patel C., Patel P., Keraliya V., Soni T. G., Patel R. C. und Patel M. M. (2012):

Osmotic drug delivery system as a part of modified release dosage form.

ISRN Pharm 2012: 528079.

Kilkenny C., Browne W. J., Cuthill I. C., Emerson M. und Altman D. G. (2010):

Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research.

PLOS Biology 8: e1000412.

Kilkenny C., Parsons N., Kadyszewski E., Festing M. F., Cuthill I. C., Fry D., Hutton J. und Altman D. G. (2009):

Survey of the quality of experimental design, statistical analysis and reporting of research using animals.

PLoS One 4: e7824.

Klein C., Budiman T., Homberg J. R., Verma D., Keijer J. und van Schothorst E. M. (2022):

Measuring Locomotor Activity and Behavioral Aspects of Rodents Living in the Home-Cage.

Front Behav Neurosci 16: 877323.

Klein H. J. und Bayne K. A. (2007):

Establishing a culture of care, conscience, and responsibility: addressing the improvement of scientific discovery and animal welfare through science-based performance standards.

Ilar j 48: 3-11.

Klein M., Stieger A., Stenger D., Scheuer C., Holstein J. H., Pohlemann T., Menger M. D. und Histing T. (2015):

Comparison of healing process in open osteotomy model and open fracture model: Delayed healing of osteotomies after intramedullary screw fixation.

J Orthop Res 33: 971-978.

Koewler N. J., Freeman K. T., Buus R. J., Herrera M. B., Jimenez-Andrade J. M., Ghilardi J. R., Peters C. M., Sullivan L. J., Kuskowski M. A., Lewis J. L. und Mantyh P. W. (2007):

Effects of a monoclonal antibody raised against nerve growth factor on skeletal pain and bone healing after fracture of the C57BL/6J mouse femur.

J Bone Miner Res 22: 1732-1742.

Kohn D. F., Martin T. E., Foley P. L., Morris T. H., Swindle M. M., Vogler G. A. und Wixson S. K. (2007):

Public statement: guidelines for the assessment and management of pain in rodents and rabbits.

J Am Assoc Lab Anim Sci 46: 97-108.

Kolar P., Gaber T., Perka C., Duda G. N. und Buttgereit F. (2011):

Human early fracture hematoma is characterized by inflammation and hypoxia.

Clin Orthop Relat Res 469: 3118-3126.

Kuribara H. und Tadokoro S. (1991):

The ambulation-increasing effect of buprenorphine in mice: comparison with the effect of morphine.

Arukoru Kenkyuto Yakubutsu Ison 26: 37-48.

Lang A., Schulz A., Ellinghaus A. und Schmidt-Bleek K. (2016):

Osteotomy models - the current status on pain scoring and management in small rodents.

Lab Anim 50: 433-441.

Langford D. J., Bailey A. L., Chanda M. L., Clarke S. E., Drummond T. E., Echols S., Glick S., Ingrao J., Klassen-Ross T., LaCroix-Fralish M. L., Matsumiya L., Sorge R. E., Sotocinal S. G., Tabaka J. M., Wong D., van den Maagdenberg A. M. J. M., Ferrari M. D., Craig K. D. und Mogil J. S. (2010):

Coding of facial expressions of pain in the laboratory mouse.

Nature Methods 7: 447-449.

Leach M. C., Klaus K., Miller A. L., Scotto di Perrotolo M., Sotocinal S. G. und Flecknell P. A. (2012):

The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale.

PLoS One 7: e35656.

Leenaars M., Hooijmans C. R., van Veggel N., ter Riet G., Leeflang M., Hooft L., van der Wilt G. J., Tillema A. und Ritskes-Hoitinga M. (2012):

A step-by-step guide to systematically identify all relevant animal studies.

Lab Anim 46: 24-31.

Leung V., Rousseau-Blass F., Beauchamp G. und Pang D. S. J. (2018):

ARRIVE has not ARRIVEd: Support for the ARRIVE (Animal Research: Reporting of in vivo Experiments) guidelines does not improve the reporting quality of papers in animal welfare, analgesia or anesthesia.

PLoS One 13: e0197882.

Lever C., Burton S. und O'Keefe J. (2006):

Rearing on hind legs, environmental novelty, and the hippocampal formation.

Rev Neurosci 17: 111-133.

Li X., Gu W., Masinde G., Hamilton-Ulland M., Rundle C. H., Mohan S. und Baylink D. J. (2001):

Genetic variation in bone-regenerative capacity among inbred strains of mice.

Bone 29: 134-140.

Lienau J., Schell H., Duda G. N., Seebeck P., Muchow S. und Bail H. J. (2005):

Initial vascularization and tissue differentiation are influenced by fixation stability.

J Orthop Res 23: 639-645.

Liles J. H. und Flecknell P. A. (1992):

The effects of buprenorphine, nalbuphine and butorphanol alone or following halothane anaesthesia on food and water consumption and locomotor movement in rats.

Lab Anim 26: 180-189.

Loeser J. D. und Treede R. D. (2008):

The Kyoto protocol of IASP Basic Pain Terminology.

Pain 137: 473-477.

Mackiewicz A. L., Salyards G. W., Knych H. K., Hill A. E. und Christe K. L. (2019):

Pharmacokinetics of a Long-lasting, Highly Concentrated Buprenorphine Solution after Subcutaneous Administration in Rhesus Macaques (Macaca mulatta).

J Am Assoc Lab Anim Sci 58: 501-509.

Madden J. C., Enoch S. J., Paini A. und Cronin M. T. D. (2020):

A Review of In Silico Tools as Alternatives to Animal Testing: Principles, Resources and Applications.

Alternatives to Laboratory Animals 48: 146-172.

Magnusdottir R., Gohin S., Ter Heegde F., Hopkinson M., McNally I. F., Fisher A., Upton N., Billinton A. und Chenu C. (2021):

Fracture-induced pain-like behaviours in a femoral fracture mouse model.

Osteoporos Int 32: 2347-2359.

Majuta L. A., Guedon J. G., Mitchell S. A. T., Ossipov M. H. und Mantyh P. W. (2017):

Anti-nerve growth factor therapy increases spontaneous day/night activity in mice with orthopedic surgery-induced pain.

Pain 158: 605-617.

Majuta L. A., Longo G., Fealk M. N., McCaffrey G. und Mantyh P. W. (2015):

Orthopedic surgery and bone fracture pain are both significantly attenuated by sustained blockade of nerve growth factor.

Pain 156: 157-165.

Majuta L. A., Mitchell S. A. T., Kuskowski M. A. und Mantyh P. W. (2018):

Anti-nerve growth factor does not change physical activity in normal young or aging mice but does increase activity in mice with skeletal pain.

Pain 159: 2285-2295.

Manigrasso M. B. und O'Connor J. P. (2008):

Comparison of fracture healing among different inbred mouse strains.

Calcif Tissue Int 82: 465-474.

Martin R. K., Albright J. P., Jee W. S., Taylor G. N. und Clarke W. R. (1981):

Bone loss in the beagle tibia: influence of age, weight, and sex.

Calcif Tissue Int 33: 233-238.

Martini L., Fini M., Giavaresi G. und Giardino R. (2001):

Sheep model in orthopedic research: a literature review.

Comp Med 51: 292-299.

Matthiesen T., Wöhrmann T., Coogan T. P. und Uragg H. (1998):

The experimental toxicology of tramadol: an overview.

Toxicol Lett 95: 63-71.

McKenzie J. A., Maschhoff C., Liu X., Migotsky N., Silva M. J. und Gardner M. J. (2019):

Activation of hedgehog signaling by systemic agonist improves fracture healing in aged mice.

J Orthop Res 37: 51-59.

Meng E. und Hoang T. (2012):

Micro- and nano-fabricated implantable drug-delivery systems.

Therapeutic delivery 3: 1457-1467.

Mieske P., Hobbiesiefken U., Fischer-Tenhagen C., Heinl C., Hohlbaum K., Kahnau P., Meier J., Wilzopolski J., Butzke D., Rudeck J., Lewejohann L. und Diederich K. (2022):

Bored at home?-A systematic review on the effect of environmental enrichment on the welfare of laboratory rats and mice.

Front Vet Sci 9: 899219.

Miller A., Kitson G., Skalkoyannis B. und Leach M. (2015):

The effect of isoflurane anaesthesia and buprenorphine on the mouse grimace scale and behaviour in CBA and DBA/2 mice.

Applied Animal Behaviour Science 172: 58-62.

Miller A. L., Kitson G. L., Skalkoyannis B., Flecknell P. A. und Leach M. C. (2016):

Using the mouse grimace scale and behaviour to assess pain in CBA mice following vasectomy.

Appl Anim Behav Sci 181: 160-165.

Miller A. L. und Leach M. C. (2015):

The Mouse Grimace Scale: A Clinically Useful Tool?

PLoS One 10: e0136000.

Miller A. L., Wright-Williams S. L., Flecknell P. A. und Roughan J. V. (2012):

A comparison of abdominal and scrotal approach methods of vasectomy and the influence of analgesic treatment in laboratory mice.

Lab Anim 46: 304-310.

Mogil J. S. (1999):

The genetic mediation of individual differences in sensitivity to pain and its inhibition.

Proc Natl Acad Sci U S A 96: 7744-7751.

Mogil J. S. (2009):

Animal models of pain: progress and challenges.

Nat Rev Neurosci 10: 283-294.

Mogil J. S., Pang D. S. J., Silva Dutra G. G. und Chambers C. T. (2020):

The development and use of facial grimace scales for pain measurement in animals.

Neurosci Biobehav Rev 116: 480-493.

Morgan E. F., De Giacomo A. und Gerstenfeld L. C. (2014):

Overview of skeletal repair (fracture healing and its assessment).

Methods Mol Biol 1130: 13-31.

Morton D. B. und Griffiths P. H. (1985):

Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment.

Vet Rec 116: 431-436.

Myers P. H., Goulding D. R., Wiltshire R. A., McGee C. A., Dickerson A. B., Comins M. M., Shi M., Kissling G. E., Lih F. B., Deterding L. J., Laber-Laird K. E. und Blankenship-Paris T. L. (2021):

Serum Buprenorphine Concentrations and Behavioral Activity in Mice After a Single Subcutaneous Injection of Simbadol, Buprenorphine SR-LAB, or Standard Buprenorphine.

J Am Assoc Lab Anim Sci 60: 661-666.

National Research Council (2009):

Recognition and Alleviation of Pain in Laboratory Animals.

Recognition and Alleviation of Pain in Laboratory Animals. Washington (DC), National Academies Press (US)

National Research Council (2011):

Guide for the Care and Use of Laboratory Animals: Eighth Edition.

Washington, DC, The National Academies Press – ISBN 978-0-309-15400-0.

Newberry R. C. (1995):

Environmental enrichment: Increasing the biological relevance of captive environments.

Applied Animal Behaviour Science 44: 229-243.

O'Loughlin P. F., Morr S., Bogunovic L., Kim A. D., Park B. und Lane J. M. (2008):

Selection and development of preclinical models in fracture-healing research.

J Bone Joint Surg Am 90 Suppl 1: 79-84.

Obernier J. A. und Baldwin R. L. (2006):

Establishing an appropriate period of acclimatization following transportation of laboratory animals.

Ilar j 47: 364-369.

Oliver V. L., Thurston S. E. und Lofgren J. L. (2018):

Using Cageside Measures to Evaluate Analgesic Efficacy in Mice (Mus musculus) after Surgery.

J Am Assoc Lab Anim Sci 57: 186-201.

Page C. D., Sarabia-Estrada R., Hoffman R. J., Lo C. P. und Gades N. M. (2019):

Lack of Absorption of a Sustained-release Buprenorphine Formulation Administered Subcutaneously to Athymic Nude Rats.

J Am Assoc Lab Anim Sci 58: 597-600.

Page G. G. (2003):

The immune-suppressive effects of pain.

Adv Exp Med Biol 521: 117-125.

Pathan H. und Williams J. (2012):

Basic opioid pharmacology: an update.

British journal of pain 6: 11-16.

Percie du Sert N., Hurst V., Ahluwalia A., Alam S., Avey M. T., Baker M., Browne W. J., Clark A., Cuthill I. C., Dirnagl U., Emerson M., Garner P., Holgate S. T., Howells D. W., Karp N. A., Lazic S. E., Lidster K., MacCallum C. J., Macleod M., Pearl E. J., Petersen O. H., Rawle F., Reynolds P., Rooney K., Sena E. S., Silberberg S. D., Steckler T. und Würbel H. (2020):

The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research.

PLOS Biology 18: e3000410.

Peric M., Dumic-Cule I., Grcevic D., Matijasic M., Verbanac D., Paul R., Grgurevic L., Trkulja V., Bagi C. M. und Vukicevic S. (2015):

The rational use of animal models in the evaluation of novel bone regenerative therapies.

Bone 70: 73-86.

Peterson N. C., Nunamaker E. A. und Turner P. V. (2017):

To Treat or Not to Treat: The Effects of Pain on Experimental Parameters.

Comp Med 67: 469-482.

Pfeiffenberger M., Damerau A., Lang A., Buttgereit F., Hoff P. und Gaber T. (2021):

Fracture Healing Research-Shift towards In Vitro Modeling?

Biomedicines 9.

Phillips A. M. (2005):

Overview of the fracture healing cascade.

Injury 36 Suppl 3: S5-7.

Radi Z. A. und Khan N. K. (2005):

Effects of cyclooxygenase inhibition on bone, tendon, and ligament healing.

Inflamm Res 54: 358-366.

Raja S. N., Carr D. B., Cohen M., Finnerup N. B., Flor H., Gibson S., Keefe F. J., Mogil J. S., Ringkamp M., Sluka K. A., Song X. J., Stevens B., Sullivan M. D., Tutelman P. R., Ushida T. und Vader K. (2020):

The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises.

Pain 161: 1976-1982.

Rasmussen S., Miller M. M., Filipski S. B. und Tolwani R. J. (2011):

Cage change influences serum corticosterone and anxiety-like behaviors in the mouse.

J Am Assoc Lab Anim Sci 50: 479-483.

Ratuski A. S. und Weary D. M. (2022):

Environmental Enrichment for Rats and Mice Housed in Laboratories: A Metareview.

Animals 12.

Reinwald S. und Burr D. (2008):

Review of nonprimate, large animal models for osteoporosis research.

J Bone Miner Res 23: 1353-1368.

Ren J., Han S., Proctor A. R., Desa D. E., Ramirez G. A., Ching-Roa V. R. D., Majeski J. B., Dar I. A., Barber N. E., Forti A. M., Benoit D. S. W. und Choe R. (2020):

Longitudinal 3D Blood Flow Distribution Provided by Diffuse Correlation Tomography during Bone Healing in a Murine Fracture Model.

Photochem Photobiol 96: 380-387.

Rock M. L., Karas A. Z., Rodriguez K. B., Gallo M. S., Pritchett-Corning K., Karas R. H., Aronovitz M. und Gaskill B. N. (2014):

The time-to-integrate-to-nest test as an indicator of wellbeing in laboratory mice.

J Am Assoc Lab Anim Sci 53: 24-28.

Röntgen V., Blakytny R., Matthys R., Landauer M., Wehner T., Göckelmann M., Jermendy P., Amling M., Schinke T., Claes L. und Ignatius A. (2010):

Fracture healing in mice under controlled rigid and flexible conditions using an adjustable external fixator.

J Orthop Res 28: 1456-1462.

Roughan J. V., Coulter C. A., Flecknell P. A., Thomas H. D. und Sufka K. J. (2014):

The conditioned place preference test for assessing welfare consequences and potential refinements in a mouse bladder cancer model.

PLoS One 9: e103362.

Roughan J. V., Wright-Williams S. L. und Flecknell P. A. (2009):

Automated analysis of postoperative behaviour: assessment of HomeCageScan as a novel method to rapidly identify pain and analgesic effects in mice.

Lab Anim 43: 17-26.

Russell W. M. S. und Burch R. L. (1959):

The principles of humane experimental technique. Methuen.

Sauer M., Fleischmann T., Lipiski M., Arras M. und Jirkof P. (2016):

Buprenorphine via drinking water and combined oral-injection protocols for pain relief in mice.

Applied Animal Behaviour Science 185: 103-112.

Schmidt-Bleek K., Petersen A., Dienelt A., Schwarz C. und Duda G. N. (2014):

Initiation and early control of tissue regeneration - bone healing as a model system for tissue regeneration.

Expert Opin Biol Ther 14: 247-259.

Schreiner V., Detampel P., Jirkof P., Puchkov M. und Huwyler J. (2021):

Buprenorphine loaded PLGA microparticles: Characterization of a sustained-release formulation.

Journal of Drug Delivery Science and Technology 63: 102558.

Schreiner V., Durst M., Arras M., Detampel P., Jirkof P. und Huwyler J. (2020):

Design and in vivo evaluation of a microparticulate depot formulation of buprenorphine for veterinary use.

Scientific Reports 10: 17295.

Schuppli C. A., Fraser D. und McDonald M. (2004):

Expanding the three Rs to meet new challenges in humane animal experimentation.

Altern Lab Anim 32: 525-532.

Schwarz C., Wulsten D., Ellinghaus A., Lienau J., Willie B. M. und Duda G. N. (2013):

Mechanical load modulates the stimulatory effect of BMP2 in a rat nonunion model.

Tissue engineering. Part A 19: 247-254.

Segalin C., Williams J., Karigo T., Hui M., Zelikowsky M., Sun J. J., Perona P., Anderson D. J. und Kennedy A. (2021):

The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice.

eLife 10: e63720.

Sheng M. H. C., Lau K. H. W., Mohan S., Baylink D. J. und Wergedal J. E. (2006):

High Osteoblastic Activity In C3H/HeJ Mice Compared to C57BL/6J Mice Is Associated with Low Apoptosis in C3H/HeJ Osteoblasts.

Calcified Tissue International 78: 293-301.

Simiand J., Keane P. E. und Morre M. (1984):

The staircase test in mice: a simple and efficient procedure for primary screening of anxiolytic agents.

Psychopharmacology (Berl) 84: 48-53.

Smith A. J., Clutton R. E., Lilley E., Hansen K. E. A. und Brattelid T. (2018):

PREPARE: guidelines for planning animal research and testing.

Lab Anim 52: 135-141.

Smith J. A., Birke L. und Sadler D. (1997):

Reporting animal use in scientific papers.

Lab Anim 31: 312-317.

Smith J. C. (2019):

A Review of Strain and Sex Differences in Response to Pain and Analgesia in Mice.

Comp Med 69: 490-500.

Stasiak K. L., Maul D., French E., Hellyer P. W. und VandeWoude S. (2003):

Species-specific assessment of pain in laboratory animals.

Contemp Top Lab Anim Sci 42: 13-20.

Strech D. und Dirnagl U. (2019):

3Rs missing: animal research without scientific value is unethical.

BMJ Open Science 3: bmjos-2018-000048.

Tang X. und Sanford L. D. (2005):

Home cage activity and activity-based measures of anxiety in 129P3/J, 129X1/SvJ and C57BL/6J mice.

Physiol Behav 84: 105-115.

Tannenbaum J. und Bennett B. T. (2015):

Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose.

Journal of the American Association for Laboratory Animal Science : JAALAS 54: 120-132.

Tappe-Theodor A., King T. und Morgan M. M. (2019):

Pros and Cons of Clinically Relevant Methods to Assess Pain in Rodents.

Neurosci Biobehav Rev 100: 335-343.

The Commission to the European Parliament and the Council (2019).

2019 report on the statistics on the use of animals for scientific purposes in the Member

States of the European Union in 2015-2017.

Retrieved 23.12.2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0016&from=EN.

TierSchG.

Tierschutzgesetz in der Fassung der Bekanntmachung vom 18. Mai 2006 (BGBl. I S. 1206, 1313), das zuletzt durch Artikel 2 Absatz 20 des Gesetzes vom 20. Dezember 2022 (BGBl. I S. 2752) geändert worden ist.

Retrieved 18.11.2022, from https://www.gesetze-im-internet.de/tierschg/BJNR012770972.html.

TierSchVersV.

Tierschutz-Versuchstierverordnung vom 1. August 2013 (BGBl. I S. 3125, 3126), die zuletzt durch Artikel 1 der Verordnung vom 11. August 2021 (BGBl. I S. 3570) geändert worden ist.

Retrieved 18.11.2022, from https://www.gesetze-im-internet.de/tierschversv/.

Tihanyi D. K., Szijarto A., Fülöp A., Denecke B., Lurje G., Neumann U. P., Czigany Z. und Tolba R. (2019):

Systematic Review on Characteristics and Reporting Quality of Animal Studies in Liver Regeneration Triggered by Portal Vein Occlusion and Associating Liver Partition and Portal Vein Ligation for Staged Hepatectomy: Adherence to the ARRIVE Guidelines.

Journal of Surgical Research 235: 578-590.

Ting K. H., Hill C. L. und Whittle S. L. (2015):

Quality of reporting of interventional animal studies in rheumatology: a systematic review using the ARRIVE guidelines.

Int J Rheum Dis 18: 488-494.

Tuli J. S., Smith J. A. und Morton D. B. (1995):

Stress measurements in mice after transportation.

Lab Anim 29: 132-138.

Turner P. V., Brabb T., Pekow C. und Vasbinder M. A. (2011):

Administration of substances to laboratory animals: routes of administration and factors to consider.

J Am Assoc Lab Anim Sci 50: 600-613.

Turner P. V., Pang D. S. und Lofgren J. L. (2019):

A Review of Pain Assessment Methods in Laboratory Rodents.

Comp Med 69: 451-467.

Ueno H., Takahashi Y., Suemitsu S., Murakami S., Kitamura N., Wani K., Matsumoto Y., Okamoto M. und Ishihara T. (2020):

Effects of repetitive gentle handling of male C57BL/6NCrl mice on comparative behavioural test results.

Scientific Reports 10: 3509.

van Abeelen J. H. F. (1970):

Genetics of rearing behavior in mice.

Behavior Genetics 1: 71-76.

van Abeelen J. H. F., van der Kroon P. H. W. und Bekkers M. F. J. (1973):

Mice selected for rearing behavior: Some physiological variables.

Behavior Genetics 3: 85-90.

Verma R. K., Arora S. und Garg S. (2004):

Osmotic pumps in drug delivery.

Crit Rev Ther Drug Carrier Syst 21: 477-520.

Watkins L. R. und Mayer D. J. (1982):

Organization of endogenous opiate and nonopiate pain control systems.

Science 216: 1185-1192.

Wolter A., Bucher C. H., Kurmies S., Schreiner V., Konietschke F., Hohlbaum K., Klopfleisch R., Löhning M., Thöne-Reineke C., Buttgereit F., Huwyler J., Jirkof P., Rapp A. E. und Lang A. (2023a):

A buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72 h in mouse femoral fracture models.

Scientific Reports 13: 3824.

Wolter A., Jirkof P., Thöne-Reineke C., Rapp A. E. und Lang A. (2023b):

Evaluating rearing behaviour as a model-specific pain indicator in mouse osteotomy models.

Lab Anim: 236772231183197.

Wolter A., Rapp A. E., Durst M. S., Hildebrand L., Löhning M., Buttgereit F., Schmidt-Bleek K., Jirkof P. und Lang A. (2021):

Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models.

Bone 152: 116088.

Wong S. A., Hu D., Shao T., Niemi E., Barruet E., Morales B. M., Boozarpour O., Miclau T., Hsiao E. C., Nakamura M., Bahney C. S. und Marcucio R. S. (2020):

β-catenin Signaling Regulates Cell Fate Decisions at the Transition Zone of the Chondro-Osseous Junction During Fracture Healing.

bioRxiv: 2020.2003.2011.986141.

Working Z. M., Morris E. R., Chang J. C., Coghlan R. F., Johnstone B., Miclau T., 3rd, Horton W. A. und Bahney C. S. (2021):

A quantitative serum biomarker of circulating collagen X effectively correlates with endochondral fracture healing.

J Orthop Res 39: 53-62.

Wright-Williams S. L., Courade J. P., Richardson C. A., Roughan J. V. und Flecknell P. A. (2007):

Effects of vasectomy surgery and meloxicam treatment on faecal corticosterone levels and behaviour in two strains of laboratory mouse.

Pain 130: 108-118.

Würbel H. und Garner J. (2007):

Refinement of rodent research through environmental enrichment and systematic randomization.

NC3Rs 9: 1-9.

Yeadon J. (2013).

The difference between 'germ free' and 'specific-pathogen-free' mice.

Retrieved 28.12.2022, from https://www.jax.org/news-and-insights/jax-blog/2013/may/the-difference-between-germ-free-and-specific-pathogen-free-mice.

Young A., Viswanath A., Kalladka M., Khan J., Eliav E. und Diehl S. R. (2018):

Mouse model demonstrates strain differences in susceptibility to opioid side effects.

Neurosci Lett 675: 110-115.

11 Publications and congress participation

11.1 Publications from the presented work (peer reviewed)

Angelique Wolter*, Anna E. Rapp*, Mattea S. Durst, Laura Hildebrand, Max Löhning, Frank Buttgereit, Katharina Schmidt-Bleek, Paulin Jirkof, Annemarie Lang (2021) Systematic review on the reporting accuracy of experimental details in publications using mouse femoral fracture models. *Bone.* *Contributed equally.

Angelique Wolter, Christian H. Bucher, Sebastian Kurmies, Viktoria Schreiner, Frank Konietschke, Katharina Hohlbaum, Robert Klopfleisch, Max Löhning, Christa Thöne-Reineke, Frank Buttgereit, Jörg Huwyler, Paulin Jirkof, Anna E. Rapp*, Annemarie Lang* (2023) A Buprenorphine depot formulation provides effective sustained post-surgical analgesia for 72h in mouse femoral fracture models. *Scientific Reports.* *Contributed equally.

Angelique Wolter, Paulin Jirkof, Christa Thöne-Reineke, Anna E. Rapp*, Annemarie Lang* (2023). Evaluating rearing behavior as a model-specific pain indicator in mouse osteotomy models. *Laboratory Animals*. *Contributed equally.

11.2 Publications from additional work

Annemarie Lang, Andreas Benn, **Angelique Wolter**, Tim Balcaen, Joseph Collins, Greet Kerckhofs, An Zwijsen, Joel D Boerckel (2023). Endothelial SMAD1/5 signaling couples angiogenesis to osteogenesis during long bone growth. *Preprint.* A*ccepted for publication in Communications Biology.*

Christian H. Bucher, Julia C. Berkmann, Lisa-Marie Burkhardt, Carolin Paschke, Claudia Schlundt, Annemarie Lang, **Angelique Wolter**, Alexandra Damerau, Sven Geissler, Hans-Dieter Volk, Georg N. Duda and Katharina Schmidt-Bleek (2022). Local immune cell contributions to fracture healing in aged individuals - A novel role for interleukin 22. *Experimental & Molecular Medicine.*

Annemarie Lang, Jonathan Stefanowski, Moritz Pfeiffenberger, **Angelique Wolter**, Alexandra Damerau, Shabnam Hemmati-Sadeghi, Rainer Haag, Anja E. Hauser, Max Löhning, Georg N. Duda, Paula Hoff, Katharina Schmidt-Bleek, Timo Gaber, Frank Buttgereit (2022). MIF does only marginally enhance the pro-regenerative capacities of DFO in a mouse-osteotomy-model of compromised bone healing conditions. *Bone.*

11.3 Abstracts and participation in congresses and conferences

In chronological order (newest first):

Deutscher Kongress für Orthopädie und Unfallchirurgie, 2022

Anna E. Rapp, **Angelique Wolter** and Annemarie Lang. *Evaluation of a new depot formulation of buprenorphine for sustained post-surgical analgesia in mouse femoral fracture models*. Oral Presentation by Anna E. Rapp.

DVG-Vet-Congress, 2022

Angelique Wolter, Christian Bucher, Viktoria Schreiner, Frank Buttgereit, Christa Thöne-Reineke, Jörg Huwyler, Paulin Jirkof, Anna E. Rapp, Annemarie Lang. *Verbesserung des Schmerzmanagements im Maus-Osteotomie-Modell*. Oral Presentation by Angelique Wolter.

EUSAAT - European Congress on Alternatives to Animal Testing, 2022

Angelique Wolter, Anna E. Rapp, and Annemarie Lang (2022). *Evaluation of a new microparticulate depot formulation of buprenorphine for sustained post-surgical analgesia*. Oral Presentation by Angelique Wolter.

DVG-Online-Tierschutztagung: Wird Tierschutz dem Geschlecht gerecht?, online, 2022

Angelique Wolter, Christian Bucher, Viktoria Schreiner, Frank Buttgereit, Christa Thöne-Reineke, Jörg Huwyler, Paulin Jirkof, Anna E. Rapp, Annemarie Lang. *Verbesserung des Schmerzmanagements im Maus-Osteotomie-Modell*. Conference Proceedings and Oral Presentation by Angelique Wolter.

Orthopedic Research Society 2022 Annual Meeting, 2022

Annemarie Lang, **Angelique Wolter**, Christian Bucher, Paulin Jirkof, Anna E. Rapp (2022). *Testing a New Microparticulate Depot Formulation of Buprenorphine for Sustained Post-Surgical Analgesia in Mouse Femoral Fracture Models*. Poster Presentation by Annemarie Lang.

11th World Congress on Alternatives and Animal Use in the Life Sciences (WC11), online, 2021

Angelique Wolter, Anna E. Rapp, Mattea Durst, Paulin Jirkof, Annemarie Lang. *Systematic review on the reporting of mouse models for bone healing.* Poster presentation by Angelique Wolter.

JRC Summer School on Non-animal Approaches in Science, online, 2021

Angelique Wolter, Anna E. Rapp, Annemarie Lang. *Evaluating further refinement methods and a newly developed sustained-release Buprenorphine to relief pain in mouse-osteotomy-models - RefineMOMo 2.0.* Poster Presentation by Angelique Wolter.

Angelique Wolter, Anna E. Rapp, and Annemarie Lang. *Evaluating Further Refinement Methods and a Newly Developed Sustained-release Buprenorphine to Relieve Pain in Mouse Osteotomy Models.* Abstract printed in: *Altern Lab Anim*. 2021 Nov;49(6):235-300. doi: 10.1177/02611929211065919.

12 Funding sources

This study was funded by Charite3^{R}, Charité-Universitätsmedizin Berlin (RefineMOMo 2.0; RefineLab) and the Promotionsstipendium I (Abschlussstipendium) of the Charité - Universitätsmedizin Berlin. We acknowledge support by the Open Access Publication Fund of the Freie Universität Berlin (open Access funding enabled and organized by Projekt DEAL).

13 Conflict of interest

There is no conflict of interest through financial support of the work.

14 Anteilserklärung

Alle Anteilserklärungen zu den Veröffentlichungen sind im Detail bei der jeweiligen Veröffentlichung angegeben.

15 Danksagung

Zuerst möchte ich Frau Prof. Dr. Christa Thöne-Reineke für die schöne Zusammenarbeit und die zuverlässige Betreuung meiner Dissertation danken.

Mein größter Dank gilt Dr. Annemarie Lang, PhD, die mir bei jeder Frage und Unwegsamkeit mit Rat und Tat zur Seite stand und mich bis zum Ende dieser Arbeit begleitet hat. Vielen Dank für deine stetige wissenschaftliche und persönliche Unterstützung, Ami.

Weiterhin bedanke ich mich herzlich bei Dr. Anna Rapp, Dr. Christian Bucher und Sebastian Kurmies für die Zusammenarbeit und die fachliche und praktische Unterstützung während der Versuche und der folgenden Probenaufarbeitung.

Bei Prof. Dr. Frank Buttgereit, Dr. Timo Gaber, Dr. Moritz Pfeiffenberger und Dr. Alexandra Damerau von der Charité Universitätsmedizin Berlin bedanke ich mich für die herzliche Aufnahme in die Arbeitsgruppe und die gute Zusammenarbeit.

Ich danke Dr. Kristina Ullmann für ihre Unterstützung bei der Planung der Versuche und Bianka Verrett, Sylvia Gundelach und allen weiteren Tierpflegern und Tierpflegerinnen für die fürsorgliche Betreuung meiner Tiere.

Prof. Dr. Jörg Huwyler danke ich für die zur Verfügungstellung des Sustained-release Buprenorphins und die gelungene Kooperation.

Ich danke auch Marzena Princ und Mario Thiele vom Julius Wolff Institut (JWI) - Center for Musculoskeletal Biomechanics and Regeneration am BIH für Ihre fachkundige Einarbeitung bei der histologischen Aufarbeitung und Analyse der Proben.

Weiterhin gilt mein Dank Dr. Paulin Jirkof und Dr. Katharina Hohlbaum für Ihre fachlich kompetente Unterstützung und den wertvollen Input.

Von ganzem Herzen danke ich insbesondere meinen Eltern, Dagmar und Sami Wolter, und meinen Großeltern, für eure immerwährende Unterstützung und Ermutigung. Danke, dass ihr immer an mich glaubt und mich bei meinen Vorhaben ermutigt! Außerdem bedanke ich mich sehr bei Moritz Vollgraf und meinen Freunden, danke, dass es euch gibt!

16 Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe.

Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen und wurde bisher nicht veröffentlicht.

Berlin, den 13.02.2024

Angelique Wolter

Druck: Mensch und Buch Verlag Berlin