A systematic review of discomfort due to toe or ear clipping in laboratory rodents

Kimberley E. Wever¹, Florentine J. Geessink¹, Michelle A. E. Brouwer¹, Alice Tillema² and Merel Ritskes-Hoitinga¹

Abstract
Toe clipping and ear clipping (also ear notching or ear punching) are frequently used methods for individual identification of laboratory rodents. These procedures potentially cause severe discomfort, which can reduce animal welfare and distort experimental results. However, no systematic summary of the evidence on this topic currently exists. We conducted a systematic review of the evidence for discomfort due to toe or ear clipping in rodents. The review methodology was pre-specified in a registered review protocol. The population, intervention, control, outcome (PICO) question was: In rodents, what is the effect of toe clipping or ear clipping, compared with no clipping or sham clipping, on welfare-related outcomes? Through a systematic search in PubMed, Embase, Web of Science and grey literature, we identified seven studies on the effect of ear clipping on animal welfare, and five such studies on toe clipping. Studies were included in the review if they contained original data from an in vivo experiment in rodents, assessing the effect of toe clipping or ear clipping on a welfare-related outcome. Case studies and studies applying unsuitable co-interventions were excluded. Study quality was appraised using an extended version of SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE)’s risk of bias tool for animal studies. Study characteristics and outcome measures were highly heterogeneous, and there was an unclear or high risk of bias in all studies. We therefore present a narrative synthesis of the evidence identified. None of the studies reported a sample size calculation. Out of over 60 different outcomes, we found evidence of discomfort due to ear clipping in the form of increased respiratory volume, vocalization and blood pressure. For toe clipping, increased vocalization and decreased motor activity in pups were found, as well as long-term effects in the form of reduced grip strength and swimming ability in adults. In conclusion, there is too little evidence to reliably assess discomfort due to toe or ear clipping, compared with no clipping or sham clipping, on welfare-related outcomes. Adequately powered, high-quality studies reporting reliable, relevant outcome measures are needed to accurately assess the impact of these identification techniques. Until more reliable evidence is available, any effect of toe clipping or ear clipping on animal welfare and study results cannot be confirmed or excluded.

Keywords
refinement, toe clipping, ear clipping, systematic review, identification methods

Date received: 14 December 2016; accepted: 29 March 2017

Rodents, especially mice and rats, are the most frequently used laboratory mammals in biomedical research.¹ Within a species, they are often very similar in appearance, and are usually housed in groups for practical and welfare reasons. Individual identification of the animals is often necessary during breeding, daily care or experimental procedures, and several identification methods are in regular use. Selection of the best method of individual identification depends on several factors, including species, age, skin pigmentation, study

¹SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE)
²Medical Library, Radboud University Medical Center, Nijmegen, The Netherlands

Corresponding author:
Dr K E Wever, SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE), Radboud University Medical Center, Geert Grooteplein-Noord 29, Route 231, 6525 GA Nijmegen, The Netherlands, or PO Box 9101, 6500 HB Nijmegen, The Netherlands.
Email: kim.wever@radboudumc.nl.
duration, and the technical expertise available. The ideal identification method should be effective and practical, but should also be minimally invasive in terms of pain and discomfort to the animal, since this can reduce animal welfare and distort the experimental results (discussed in the literature\(^2\)). It is therefore important to assess the effect of identification methods on animal welfare.

Toe clipping is an individual identification method used mostly in mice, and which can be applied in newborn and very young animals. The toe may be clipped at the distal end of the second phalanx to remove the entire nail bed, or a larger segment of the toe may be removed.\(^3\) The removed tissue can be used for genotyping.\(^4\) Ear clipping or punching (notching) is used to identify individual adult rodents (mostly mice and rats). Using a special tool, holes or notches are made in the ear according to a chart/system. The punched or clipped tissue can be used for genotyping.\(^4\)

Possible alternatives for identification by toe or ear clipping include microchipping and ear or tail tattoo; but according to a recent survey,\(^4\) toe clipping and ear clipping are the most commonly used identification methods in neonate and adult mice, respectively. Both methods have a large and lasting impact on an animal’s integrity, and are likely to cause pain and discomfort.\(^2,3\) Both methods also require restraint of animals and may permanently affect their welfare. For instance, toe clipping might impair the rodent’s ability to grip, groom and feed, as well as altering its gait. Ear clipping may interfere with thermoregulation, and clipped ears may be more susceptible to tearing and infection. The ethical justification for performing these methods therefore continues to be a subject of heated debate. Although many guidelines on the subject are available,\(^3\) the evidence for discomfort caused by toe or ear clipping has not been systematically reviewed. Narrative summaries of evidence have significant limitations, because the methods used to identify studies are neither comprehensive nor transparent, and the study quality is usually not assessed. Systematic reviews can overcome many of these challenges, because they are guided by a protocol with explicit methods to identify, select, synthesize (which may include meta-analysis), and appraise all studies relevant to a particular research question. We have therefore conducted a systematic review of the evidence on discomfort due to ear and toe clipping, in order to better inform animal researchers, welfare officers, policy makers and other stakeholders when making decisions on the choice of identification method for rodents. The population, intervention, control, outcome (PICO) question was: In rodents, what is the effect of toe clipping or ear clipping, compared with no clipping or sham clipping, on welfare-related outcomes (e.g. pain, anxiety, physiological impairment, etc.)?

### Materials and methods

#### Review protocol and amendments

The review methodology was pre-specified in a review protocol and registered on http://www.syrcle.nl on 27 October 2015 (see\(^8\) and supplementary material). Minor amendments were made to the review protocol: (1) in addition to Google, we also searched for grey literature in OpenGrey.eu and WorldWideScience.org, using all possible combinations of the search terms ‘mouse’, ‘rat’ or ‘rodent’ with ‘toe’, ‘ear’, ‘phalanx’ or ‘clip’; and (2) studies applying ear tags were included in the review, since the ear is punched in this procedure and ear tagging is also a relevant identification method. However, because of the presence of a tag, results of these studies cannot be pooled with ear clip studies in which no tag is applied.

#### Search and study selection

The full search strategy is presented in Table 1. In brief, we performed a comprehensive search in PubMed, Embase and Web of Science on 5 October 2015, using the search components ‘toe, tail or ear’, ‘discomfort’ and ‘animal’. Studies were included in the review if they met the following inclusion criteria: (1) the study reported original data from an in vivo experiment using rodents; (2) the effect of toe clipping or ear clipping on a welfare-related outcome was reported, compared with a control group undergoing no intervention or sham clipping; (3) the study was not a case study; and (4) no unsuitable co-interventions were applied, e.g. interventions completely unrelated to animal identification. Studies were eligible for inclusion regardless of their use of analgesics or anesthetics. No data or language restrictions were applied. We checked the reference lists of all included studies and relevant reviews for additional references of interest. In addition, we performed a grey literature search (see Study selection below) and contacted the Dutch representatives of the Federation of European Laboratory Animal Science Associations (FELASA), as well as animal welfare officers and professors in laboratory animal science in The Netherlands, with a request to inform us about any published or unpublished data on this topic.

#### Data extraction

Data were extracted by one reviewer (MB or FJG) and checked by a second reviewer (KW), who read each paper in detail to ensure that the extracted data were accurate and that no information was missed. For nearly all papers,\(^9\)–\(^17\) data for one or more outcome measures were extracted from graphs using digital ruler software, because the numerical data were...
not reported. We attempted to contact the authors of eight studies to provide additional information on study characteristics and/or outcome data. We received responses from three authors, who were able to (partially) clarify study characteristics. Additional outcome data were provided by one author.

Risk of bias and quality assessment

For studies using a separate control group, two reviewers (KW and MB) independently assessed the risk of bias and study quality using SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE)'s risk of bias tool. In cases of discrepancies, consensus was reached by discussion between the reviewers, with a third reviewer serving as arbiter if an agreement could not be reached. Selective outcome reporting (item #9) was not assessed, since none of the studies reported the use of a study protocol predefining primary and secondary outcomes. When assessing selection bias (item #3), groups within a study were considered to be similar at baseline if the sex, strain, age and weight of the animals did not significantly differ between groups. For toe clip studies, similar weight and age were required, because the fast development of pups can cause large differences between animals of different ages, and the weight of the pups influences the accuracy with which toe clipping can be performed. For the ear clip studies, similar weight or age was considered to be sufficient.

We also assessed reporting of any randomization, reporting of any blinding, and reporting of a sample size calculation as additional study quality indicators. Because the risk of bias tool was developed for studies using separate control and treated groups, four studies could not be scored due to incompatible study designs (i.e. cross-over design or use of internal controls).
Re-analysis of outcome data

Whenever complete outcome data could be extracted or obtained (i.e. mean, variance and number of animals per group for continuous outcomes, or the number of events and non-events for dichotomous outcomes) we re-analysed the data by calculating the effect size as a standardized mean difference (SMD) or risk ratio (RR), for continuous and dichotomous outcomes, respectively. We aimed to obtain pooled effect estimates of outcome measures reported by three or more studies. However, no single outcome was reported more than twice, and outcome and study characteristics were too heterogeneous to pool various outcomes. We therefore only report the SMD and RR and corresponding 95% confidence intervals for the individual outcomes per study, without pooling data. Effect estimates were calculated using a random-effects model.

Results

Study selection

A flow chart of the study selection process is depicted in Figure 1. We identified a total of 2040 unique references, 12 of which met the inclusion criteria. Two conference abstracts met the inclusion criteria, but were excluded because the data presented appeared to match those in full research articles by the same authors that were also included (and attempts to contact the authors to verify this received no response). An additional 48 potentially relevant references were identified by hand-searching reference lists of included studies and relevant reviews, and grey literature searching. None of these references met the inclusion criteria. One conference abstract appeared to be relevant, but did not contain enough information to assess whether an appropriate control group was used. The authors were contacted for additional data, but these were not supplied.

Study characteristics

The characteristics of the included studies are summarized in Table 2. Out of the 12 studies included, five (42%) were on ear clipping without tagging, two (17%) were on ear clipping with tagging, and five (42%) were on toe clipping. Ten (83%) studies were performed on varying strains of mice, and two studies were performed on rats (one on ear clipping and one on toe clipping). The majority of studies (58%) used both male and female mice, and a number of these studies presented data separately for both sexes. In three studies, the sex of the animals was not reported. For ear clip (including ear tag) studies, the animals were adolescent or adult at the time of intervention (3–12 weeks.
Table 2. Characteristics of the included studies.

| Author + year | Species | Strain | Sex | Age at intervention | Intervention of interest | Site of clipping | Frequency | Control intervention | Outcome [direction of effect*] |
|---------------|---------|--------|-----|---------------------|--------------------------|------------------|----------|----------------------|---------------------------------|
| Cinelli 2007 | Mouse   | B6.Cg-Tg | M/F | 12 wk | Ear clip | Pinna [2 mm punch] | 1 | Restraint | Heart rate [=] Body temperature [=] Motor activity [=] |
| Kasanen 2011 | Rat     | SD/Wistar | NR  | 25 wk | Ear clip | Pinna, right | 2 | Foot microtattoo* | Heart rate 16–24 h post-treatment [↓] Blood pressure 1–24 h post-treatment [↑] |
| Miller 2015  | Mouse   | C57BL/6 | M   | 8 wk  | Ear clip | Pinna | 1 | Restraint | Mouse grimace scale directly post-treatment [=] |
| Rasid 2012   | Mouse   | Balb/c x TCR-HA+/− | M/F | 8–10 wk | Ear clip | Pinna [2 mm band] | 1 | Restraint | Respiratory minute volume [↑] |
| Williams 2008 | Mouse  | B6;12956-StatBb | M/F | PND21–28 | Ear clip | Pinna | 1 | Restraint | Vocalization during treatment [↑] |
| Baron 2005   | Mouse   | FVB/N/ FVB/NCr | NR  | 3–4 wk | Ear tag | Pinna, right base | 1 | Contralateral ear | Tumor incidence [=] Gross histopathology [?] Grading of squamous cell carcinomas [?] |
| Kitagaki 2007 | Mouse  | C57BL/6 | M   | 5 wk  | Ear tag | Pinna, right base | 1 | Contralateral ear | Ear lesions [?] Thickness of ear auricles [↑] Ear metal content [↑] Cytokine levels [↑] Metallothionein expression [INDI [↑] Histopathological changes [NDI [?] |
| Castelhano-Carlos 2010 | Mouse  | C57BL/6J | M/F | PND5 | Toe clip | 1/3 of a toe | 1 | Restraint and subcutaneous puncture | Nest disruption [NDI [=] Rejection by mother [NDI [=] Cannibalism [NDI [=] Urination upon treatment [NDI [=] Vocalization upon treatment [NDI [=] Urination and vocalization upon treatment [=] Skin appearance [NDI [=] Milk spots [NDI [=] Physical activity PND5 [NDI [=] Body weight [=] Anogenital distance [NDI [=] Development of fur [NDI [=] Negative geotaxis [NDI [=] Development of ears, eyes and teeth [=] |

(continued)
| Author + year | Species | Strain     | Sex | Age at intervention | Intervention of interest | Site of clipping | Frequency | Control intervention | Outcome [direction of effect*] |
|---------------|---------|------------|-----|---------------------|--------------------------|-----------------|----------|----------------------|--------------------------------|
| Iwaki 1989\(^{20}\) | Rat     | SD-JCL     | M/F | PND4               | Toe clip                 | 1st joint of a forelimb toe | 1        | Untreated\(^{**}\)    | Postural reflex, grasping reflex, surface righting reflex, wire suspension test and air-righting reflex [ND] [=] Mature walking [ND]? [?]
|               |         |            |     |                    |                          |                 |          |                      | In adults:
|               |         |            |     |                    |                          |                 |          | Body weight [=]       |                                        |
|               |         |            |     |                    |                          |                 |          | Home-cage climbing behavior [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Elevated plus-maze [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Rotarod treadmill constant velocity [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Rotarod treadmill accelerating velocity [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Simplified SHIRPA protocol [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Adrenal gland weight [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Thymus weight [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          |                      | In pups:
|               |         |            |     |                    |                          |                 |          | Survival until weaning [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Erection of pinnae PND7 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Righting reflex PND7 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Opening of eyelids PND17 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Negative geotaxis PND14 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Wire suspension PND21 [ ] |                                        |
|               |         |            |     |                    |                          |                 |          | Swimming ability PND21 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Rotarod treadmill PND28 [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Sexual maturation [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Body weight [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          |                      | In adults:
|               |         |            |     |                    |                          |                 |          | Pregnancy success rate [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Number of viable fetuses [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Offspring fetal weight [=] |                                        |
|               |         |            |     |                    |                          |                 |          |                      | In pups:
|               |         |            |     |                    |                          |                 |          | Vocalization upon treatment [? clipped at PND7; = clipped at PND17] |                                        |
|               |         |            |     |                    |                          |                 |          | Urination upon treatment [? clipped at PND7; = clipped at PND17] |                                        |
|               |         |            |     |                    |                          |                 |          | Struggle upon treatment [? clipped at PND7; = clipped at PND17] |                                        |
|               |         |            |     |                    |                          |                 |          | Dragging limb after treatment [? clipped at PND7; = clipped at PND17] |                                        |
|               |         |            |     |                    |                          |                 |          | Reduced activity directly after treatment [? clipped at PND7; = clipped at PND17] |                                        |
|               |         |            |     |                    |                          |                 |          | Suckling behavior [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Maternal rejection [ND] [=] |                                        |
|               |         |            |     |                    |                          |                 |          | Paw swelling [ND]? |                                        |

(continued)
| Author + year | Species | Strain | Sex | Age at intervention | Intervention of interest | Site of clipping | Frequency | Control intervention | Outcome [direction of effect*] |
|--------------|---------|--------|-----|---------------------|--------------------------|-----------------|----------|---------------------|--------------------------------|
| Schaefer 2010 | Mouse | B6D2F1 | M/F | PND3/7 | Toe clip | 2nd phalanx | 3 [2 fore, 1 hind] | Restraint | Paw erythema (ND) [?] Body weight until 10 wk [=] Neurological reflex development [negative geotaxis, surface righting reflex, grasping reflex, postural reflex, object grasping, wire suspension test, mature walking, air-righting reflex] (ND) [=] In adults: Open-field test (ND) [=] Elevated plus-maze test (ND) [=] Simplified SHIRPA protocol (ND) [=] Balance beam test (ND) [=] Rotarod treadmill (ND) [=] In pups: Paw withdrawal [?] Bleeding [=] Grooming by mother [=] Cannibalism by mother [=] Automutilation [=] Inflammation [=] Milk spots [=] Righting [=] Development of fur, teeth, ears, eyes and walking [=] Body weight [=] Corticosterone levels [=] In adults: Grip strength [† clipped at PND3; = clipped at PND7] Hot plate [= clipped at PND3; = clipped at PND7] Bone regeneration (ND) [=] Bone hypertrophy (ND) [=] In adults: Phalangeal bone length [†] Phalangeal bone width [†] |
| Vachon 1998 | Mouse | C57BL, C3H, B6C3F1 | NR | 2 wk | Toe clip | Distal end of 1st phalanx | 1 | Contralateral toe(s) | NR: not reported; M: male; F: female; wk: weeks; PND: postnatal day; (ND): descriptive, data not shown, or comparison control versus clipped not analyzed. *As described by the authors, indicated as follows: †: clipping increases outcome; ‡: clipping decreases outcome; =: clipping does not affect outcome; ?: the outcome is assessed, but the effect of clipping versus control is not described. **Not reported whether the control group underwent restraint, handling or no intervention at all. |
In toe clip studies, the age at intervention ranged from postnatal day (PND) 3 to PND17 (median PND7). Thus, the age at the time of clipping did not overlap between ear and toe clip studies.

Most studies compared outcomes in the clipped group(s) to a control group undergoing restraint only. Three studies used the unclipped contralateral ear or toe as an internal control. One study described the control group as being ‘untreated’, but whether this included handling or restraint of the animals was not specified. Two studies performed a secondary intervention in the control group, in addition to restraint (namely subcutaneous puncture and microtattoo of the foot), in order to better match the control group with other experimental groups in the study. Importantly, these interventions may have increased the level of discomfort in the control group and therefore interfere with the comparison with the toe or ear clipped group. The site of clipping was reasonably well described, but heterogeneous: the paw chosen for toe clipping differed between studies. All ear clip studies performed ear punching, except for one in which a 2 mm band was clipped off the rim of the ear. The number of ear clips applied was 1–2. The number of clipped toes was 1–3.

### Outcome data

The primary studies report a wide variety of outcome measures related to animal welfare and discomfort. Table 2 lists the outcomes and their direction of effect, as reported by the authors in the primary studies. In many studies, a large number of outcome measures were reported to be assessed, but the outcome data were often not shown, or reported descriptively (‘ND’ in Table 2). For seven papers, one or more outcome measures could not be re-analyzed because the mean, variance and/or number of animals were not reported. In three papers, conservative estimates were made regarding the numbers of animals or the variance of the data.

When an outcome was repeatedly measured in the same animals, we re-analyzed data from the measurement of maximal effect. Thus, Figures 2–6 include the following data from studies with repeated measurements: (1) Castelhano-Carlos 2010: pre-weaning body weight on PND21, post-weaning body weight in...
week 4 for males and week 12 for females; (2) Paluch 2014; body weight data in week 9 for males clipped at PND7 and PND17, week 7 for females clipped at PND7 and week 10 for females clipped at PND17; (3) Castelhano Carlos 2010; rotarod treadmill data 15 rpm velocity; (4) Vachon 1998; phalangeal length data of the 3rd digit and phalangeal width data of the 4th digit; (5) Kitagaki 2007; ear thickness at 26 weeks; and (6) Kasanen 2011; heart rate 16–24 h and blood pressure 4–16 h after ear clipping. Raw data of the forest plots are presented in the supplementary excel file.

**Ear clip studies.** Overall, ear clip studies reported 15 different outcomes (all measured in adults), the majority of which were physiological parameters related to discomfort (e.g., elevated heart rate and inflammation). Two behavioral parameters indicating discomfort or pain were reported (mouse grimace scale and vocalization during treatment).

In mice, they were reported to vocalize more frequently during ear clipping than during restraint only, and their respiratory minute volume was increased. No differences in heart rate, blood pressure, body temperature and scores on the mouse grimace scale were reported. Tagging with metal ear tags was found to increase the metal content of the ear and cause auricular chondritis, as indicated by an increase in ear thickness and elevated cytokine levels. No effect on tumor formation was observed. In rats, blood pressure and heart rate were compared between ear clipping and foot microtattoo, with varying results: blood pressure was increased at various time points after ear clipping, while heart rate was higher in the microtattooed animals.

When re-analyzing the primary data from included studies as SMD (Figure 2) or RR (Figure 3), no additional effects of ear clipping were found.

**Toe clip studies.** Overall, toe clip studies reported nearly 50 different outcome measures (Table 2). Outcomes measured in pups can be divided into parameters related to physical development (e.g., body weight, development of fur and sexual maturation), neurological development (e.g., righting and grasping reflexes), signs of discomfort in pups or their mother (e.g., vocalization during treatment and maternal rejection) and physiological parameters indicating discomfort (e.g., elevated corticosteroid levels and bleeding). Outcomes measured in adult animals mainly cover neurological and neurobehavioral tests (e.g., balance beam and open-field tests).

The majority of outcomes were reported to be unchanged between toe clipped and control animals (Table 2). In rat pups clipped on PND4, performance in the wire suspension test on PND21 was found to be decreased, indicating lower grip strength. Decreased grip strength was also observed in adult mice that were
clipped as pups on PND3, but not in mice clipped on PND7. No regrowth of toes was reported and the thickness of the phalangeal bone was increased in toe stumps.

When re-analyzing the primary data from included studies as SMD or RR, we found five additional significant effects of toe clipping, namely: increased vocalization, reduced motor activity and toe swelling after clipping on PND7 in mice, impaired adult swimming ability in rats clipped on PND4 (all Figure 5), as well as a borderline significant decrease in post-weaning body weight for female pups clipped on PND7 (Figure 4).

Quality assessment

The risk of bias and quality scores from the eight studies using separate control groups are shown in Table 3 (individual scores) and Figure 6 (overall scores). Although randomization of group allocation was mentioned in seven of these studies (87.5%; Figure 6A), no study specified the method of randomization (e.g. use of a random number table). Three out of eight studies (37.5%; Figure 6A) reported that the experimenter performing the assessments was (partially) blinded to treatment, or that he/she assessed the allocation of the animals only after performing the outcome assessment. The other studies did not mention blinding.
during any phase of the experiment. None of the 12 studies included in this review reported a sample size or power calculation.

Because of the poor reporting of bias reduction measures, the majority of items in the risk of bias tool were assessed as ‘unclear’ (Figure 6B). Insufficient reporting of (the method used for) randomization led to an unclear risk of selection, performance and detection bias (items #1, 4 and 6). Baseline characteristics of the animals were adequately reported in two studies, in which we consequently assessed the risk of selection bias to be low (item #2). In all other studies, one or more baseline characteristics were not reported, leading to an unclear risk of bias. As regards blinding, we assessed the risk of performance bias (item #5) to be high in all studies, because ear clipping and toe clipping are in practice impossible to conceal when the intervention is performed, or when the animal is subsequently handled. For this reason, we also assessed the risk of detection bias to be high in all studies (item #7), except for those in which the outcome assessors did not necessarily handle the animals. In the latter studies however, it was unclear if measures had been taken to adequately blind the outcome assessment, leading to an unclear risk of bias. Attrition bias (item #8) was assessed to be high in one study, where the numbers of animals allocated and included in the various outcome assessments could not be matched. Two studies scored correctly reported dropouts, thereby scoring a low risk of attrition bias. In the remaining five studies, the risk of attrition bias was unclear. The risk of other types of bias was considered to be low in all studies (item #9).

Figure 5. Forest plot of dichotomous outcome data from toe clip studies. Effect sizes calculated as risk ratio (RR) and corresponding 95% confidence interval (CI), using a random-effects model. Right-hand side columns indicate events from total in treatment (clipped) and control groups. Note: a RR cannot be computed when there are zero events in both experimental groups. PND: postnatal day; h: hours post-clipping.
Because of their expected impact on animal welfare, toe or ear clipping are generally considered to be controversial techniques and their performance is restricted or even abolished in many animal laboratories. An abundance of guidelines is available on toe clipping and ear clipping, as well as other methods for individual identification and genotyping (examples 3–7). However, these guidelines have, up to now, not been based on a systematic summary of all available evidence. Here, we provide the first systematic review of the evidence for the effect of toe clipping and ear clipping on rodent welfare.

Available evidence and quality

Studies investigating the effects of toe or ear clipping on rodent welfare are in short supply, and highly heterogeneous. This heterogeneity is mainly caused by differences in the population (males, females, various strains) and the intervention (age at time of clipping, number of sites clipped) under investigation, as well as the variety of outcome measures assessed. Most of the reported outcome measures showed no effect of toe or ear clipping on discomfort. Conversely, evidence indicating discomfort is present for ear clipping in the form of increased respiratory volume, vocalization and blood pressure; as well as present for toe clipping in the form of increased vocalization and decreased activity in pups, and reduced grip strength and swimming ability in adults. However, several limitations of the primary studies limit the reliability of both the evidence for and against an effect on discomfort, and hamper their interpretation.

Adequate reporting of methodological details in primary studies is crucial in order to determine the risk of bias in these studies and to assess the quality of a body of evidence. Our risk of bias assessment shows that the laboratory animal science field is no exception to the insufficient reporting of animal studies. We show that poor reporting of various aspects of experimental design resulted in most risk of bias items being assessed as unclear. This is a matter of concern, since evidence from preclinical animal studies indicates that lack of measures to reduce bias can severely influence primary study results. A high risk of performance and/or detection bias is likely to be present in all included studies, and this should be taken into account when interpreting the results.

Figure 6. Risk of bias assessment and reporting of study quality indicators in eight included studies. (a) Reporting of any mention of randomization, blinding or a power calculation. (b) The risks of selection, performance, detection, attrition and other forms of bias were assessed using SYRCLE’s risk of bias tool. Although randomization and blinding were mentioned in several articles, lack of reporting of the method used resulted in an unclear risk of bias for most items. Four studies were excluded from the assessment because their study designs were not compatible with the risk of bias tool.
Specifying the primary outcome in a study prevents changing the primary outcome based on the study results, thereby reducing the risk of bias due to selective outcome reporting. Unfortunately, none of the included studies defined which of the reported outcome measures was the primary outcome measure. In addition, none of the included studies reported a power calculation to support the number of animals used per group, even though this key element of experimental design is mandatory for approval by many animal ethics committees. The sample size calculation should specify the primary outcome measure, its expected mean and variation, and the effect size the authors aim to detect. Doing so prevents the unethical use of animals due to over powering (using more animals than necessary) or under powering (using too few animals, especially relevant in cases where no effect of the intervention is found). Furthermore, knowing the planned sample size is often essential for assessing the correct handling of dropouts and attrition bias. At present, we are unable to assess whether any of the studies were adequately powered to detect differences between the groups in the outcomes under investigation.

In one study, the exact intervention applied in the control group was unclear, making it difficult to interpret the study results. In two other studies, the application of a secondary intervention in the control group may have introduced additional discomfort in these animals, thereby masking the effect of toe or ear clipping. One study used a cross-over design, which can introduce carry-over effects that interfere with the intervention under investigation. In addition, outcome data are incompletely reported for many outcomes. We recognize that these inaccuracies are probably (partly) caused by the fact that some of the included studies were not specifically designed to assess the effect of toe clipping or ear clipping compared with handling or restraint. However, this indicates once again that studies specifically aimed to assess the effects of toe or ear clipping on welfare are very scarce.

As a result of these shortcomings, the effects of toe or ear clipping observed in a particular study cannot be directly generalized to other studies, or the population

Table 3. Individual scores for study quality indicators and risk of bias assessment in eight included studies.

| Study                        | Reporting | Risk of bias |
|------------------------------|-----------|--------------|
|                             | Any randomization | Random Group allocation (selection) | Random housing (performance) | Random intervention (performance)* | Random outcome ass. (detection) | Reporting of dropouts (attrition) | Other biases |
| Castelinho-Carlos 2010*      | Y Y N     | ? L ? ? H ? H ? L |
| Cinelli 2007                  | Y N N     | ? ? ? ? H ? ? L |
| Iwaki 1989                   | Y N N     | ? ? ? ? H ? H ? L |
| Miller 2015                  | Y Y N     | ? L ? ? H ? ? L |
| Paluch 2014                  | Y N N     | ? ? ? ? H ? H ? L |
| Rasid 2012                   | N N N     | ? ? ? ? H ? ? L |
| Schaefer 2010                | Y Y N     | ? ? ? ? H ? H H L |
| Williams 2008                | Y N N     | ? ? ? ? H ? ? L |

Y: yes; N: no; ?: unclear risk of bias; H: high risk of bias; L: low risk of bias; ass: assessment. *Blinding was not possible.
of laboratory rodents in general. Until more reliable evidence is available, any effect of toe clipping or ear clipping on animal welfare and study results can be neither confirmed nor excluded.

**Implications for refinement**

The current body of evidence is too small and too heterogeneous to reliably assess the influence of, for example, species, age, sex, strain or clipping method, on the severity of discomfort. In three studies, outcome measure analysis was performed separately for male and female animals, but the observed effects did not differ between the sexes. Schaeffer and colleagues found that pups undergoing toe clipping on PND3 had a lower grip strength than pups clipped on PND7, which was attributed to the fact that the toes in three-day-old pups are partially fused together, and are too small to accurately clip the distal phalanx only, resulting in too much of the toe being removed. Based on this finding, clipping would not be advisable before PND3, but replication of this result is needed to confirm this. Studies performing clipping on PND4 and PND5 have reported that the procedure was quick and easy to apply, while our re-analysis of data from Paluch and colleagues suggests that clipping causes discomfort in seven-day-old pups (as indicated by increased vocalization and decreased activity), but not in 17-day-old pups. We conclude that the influence of age is presently unclear and further research is needed.

One study tested whether spray-on vapocoolant anesthesia could reduce pain during toe clipping, but concluded that the spray glued the toes together, which increased the risk of incorrectly clipping the distal phalanx of a single toe and increased discomfort due to prolonged handling. Furthermore, the vapocoolant interfered with hemostasis after clipping. Based on these results, application of this analgesic agent would not be advisable, but further research into suitable local anesthetics and analgesics may be worthwhile.

**Implications for laboratory practice**

The advantages and disadvantages of toe clipping, ear clipping, ear tagging and several other identification methods have been extensively described by specialist working/research groups, such as FELASA, the Norwegian Consensus Platform for Replacement, Reduction and Refinement of Animal Experiments (Norecopa), and the joint BVA/AWF/FRAME/RSPCA/UFAW working group. The majority of studies identified in this review were also identified in the 2013 FELASA report, except for Rasid 2012, and the more recent publications by Paluch 2014 and Miller 2015. However, current reports do not present a systematic summary of the study characteristics, including, for example, details on the experimental design, control interventions and co-interventions used. This is unfortunate, especially in view of the highly heterogeneous character of the studies, which has important implications for their external validity. They also do not provide a complete overview of all outcomes assessed, and it is unclear why some outcomes are highlighted and others omitted. Perhaps most importantly, no assessment of methodological quality or risk of bias was performed in any of the previous reports. Our systematic review addresses these limitations, allowing the reader to properly assess the reliability of the available evidence when interpreting this evidence.

Of note, in their 2008 report on toe clipping in mice, Norecopa reported that they had not been able to identify any studies providing electrophysiological or histological evidence on toe(tip) innervation in rodents that would allow for an assessment of the pups’ ability to feel pain at the time of clipping. We did not identify such studies in our systematic search either. Based on the available guidelines, there is international consensus that toe clipping should not be performed after PND7, since pups become increasingly active with age, which amplifies the risk of incorrect clipping and increases the level of restraint needed to correctly perform the procedure. This is independent of the observation that phalangeal ossification is complete around PND18, after which clipping is hypothesized to be more painful, although we found no data supporting this theory directly. By contrast, ear clipping is advised to be performed no earlier than PND14, due to the small size of the ears before PND14. Thus, toe clipping and ear clipping cannot be performed at the same age, and the age at which individual identification is needed is an essential factor in the choice of toe clipping, ear clipping or other alternative method for individual identification. Alternatives include tattooing or microchipping for identification, and hair biopsies or rectal swabs for DNA sampling (an overview of methods is provided). Most of these techniques cannot be used in newborn or very young animals and therefore cannot replace toe clipping. In addition, some of these techniques may cause more discomfort than either toe or ear clipping. Other factors influencing the choice of identification method is whether, and how much, DNA is required for quantitative or qualitative genotyping, and whether the identification should be permanent or temporary.

**Directions for future research**

The present review identifies a number of important shortcomings currently hampering the interpretation of the available evidence: (1) the low number of studies...
dedicated specifically to the assessment of discomfort after toe or ear clipping; (2) the lack of standardization of the chosen outcome measures; (3) insufficient reporting of experimental detail, especially regarding justification of the sample size and measures to reduce bias; and (4) incomplete reporting of outcome data. Thus, further research is needed to provide reliable evidence on the effect of toe clipping or ear clipping on animal welfare. In order for future studies to succeed, all of these issues should be addressed.

Firstly, future studies should be aimed specifically at assessing the effect of toe clipping or ear clipping on discomfort in laboratory rodents. Their design should include appropriate control groups, preferably one receiving no treatment (to provide a baseline) and one receiving sham treatment with handling and restraint only. No co-interventions should be applied apart from the intervention of interest. The animals should be handled and housed under the same circumstances. Furthermore, it is presently unclear whether characteristics such as species, strain, sex and age of the animals influence outcome, and future studies should be specifically designed and powered to reliably address these issues.

Secondly, upon submitting a proposal for new animal studies, researchers should provide rationale for their choice of outcome measures, including details on reproducibility and the optimal time point for outcome assessment. Ideally, the relevance and reproducibility of outcome measures used to assess discomfort and welfare in pups and adult mice and rats should be validated and discussed in the field, so that consensus may be reached and experiments may be standardized accordingly. We hypothesize that a multicenter approach (e.g. the MultiPART initiative; http://www.dcn.ed.ac.uk/multipart/default.htm) may offer the opportunity to increase power and standardization of future experiments.

Thirdly, as we have shown, the urgent need to improve the reporting and methodological quality of (laboratory) animal studies should be recognized. To this end, the ARRIVE guidelines and the Gold Standard Publication Checklist (GSPC) were published in 2010, but the reporting quality of studies included in this review was low regardless of the year of publication. Of note, the ARRIVE guidelines and GSPC do not specify how detailed the reporting of measures to reduce bias should be, and SYRCLE’s risk of bias tool can provide guidance on how to report measures to reduce various forms of bias in various stages of an animal experiment. This is especially important in studies on toe or ear clipping, since these procedures are very difficult to blind, and the risk of biasing the study results is high unless adequate measures are taken. Reporting of a sample size calculation should be mandatory in the publication of future studies, especially since studies on toe or ear clipping need to be powered in order to reliably prove or disprove an effect on outcome. Finally, complete reporting of data for all outcome measures, either in the article, the supplementary material, or through open access data repositories, is essential to reach reliable conclusions in future studies, for the benefit of science and animal welfare.

**Conclusion**

Evidence on any effect of toe or ear clipping on animal welfare is too scarce, too heterogeneous and of insufficient quality to allow for reliable conclusions to be drawn. Studies that do and studies that do not show a welfare effect from toe or ear clipping both suffer from the same limitations: insufficient reporting of experimental detail (especially regarding justification of the sample size and measures to reduce bias); flaws in experimental design; a lack of rationale for, and standardization of, the chosen outcome measures; and incomplete reporting of outcome data. From an ethical, as well as an economical, point of view it is essential that future studies address these limitations. Until such studies are available, we cannot confirm or exclude any effect of toe clipping or ear clipping on animal welfare.

**Acknowledgement**

We thank Professor Adrian Smith for his valuable comments on the manuscript.

**Declaration of Conflicting Interests**

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

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by The Netherlands Organization for Health Research and Development (ZonMW, projects #2015/13447/ZONMW awarded to MR and #2015/18371/ZONMW awarded to KW, http://www.zonmw.nl). The funding body had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Ethical considerations**

This article does not contain any studies involving human participants or animals performed by any of the authors.

**References**

1. European Commission. *Seventh report on the statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union*. Brussels: EC, 2013.
Résumé

Le déphalangeage et le percage des oreilles [ou leur entaillage ou marquage] sont des méthodes utilisées fréquemment pour l’identification individuelle des rongeurs de laboratoire. Ces procédures entraînent potentiellement une gêne qui peut réduire le bien-être de l’animal et déformer les résultats des expériences. Il n’existe toutefois aucun résumé systématique des preuves à ce sujet. Nous avons mené une étude
systématique des preuves de la gêne due au déphalangage ou au perçage des oreilles des rongeurs. La méthodologie de l’étude a été précisée à l’avance dans un protocole d’étude déposé. La question PICO était : chez les rongeurs, quel est l’effet du déphalangage ou du perçage des oreilles par rapport au non-perçage ou au faux perçage, en matière de résultats relatifs au bien-être ? Par le biais d’une étude systématique sur Pubmed, EMBASE, Web of Science et la littérature parallèle, nous avons identifié plusieurs études sur les effets du perçage des oreilles sur le bien-être animal, et cinq études de ce type sur le déphalangage. Les études ont été incluses dans les résultats si elles contenaient des données d’origine issues d’expériences in vivo sur des rongeurs, pour évaluer les effets du déphalangage ou du perçage des oreilles sur les résultats en matière de bien-être. Les études de cas et les études appliquant des co-interventions inappropriées ont été exclues. La qualité de l’étude a été évaluée en utilisant la version étendue de l’outil de risque de partialité SYRCLE pour les études animales. Les caractéristiques de l’étude et les mesures du résultat étaient extrêmement hétérogènes, et il existe un risque imprimé et élevé de partialité de toutes les études. Nous présentons par conséquent une synthèse descriptive des preuves identifiées. Aucune des études ne mentionnait de calcul de la taille de l’échantillon. Sur plus de 60 résultats différents, nous avons trouvé des preuves de gêne suite au perçage de l’oreille sous la forme d’un volume respiratoire supérieur, la vocalisation et la pression sanguine. Concernant le déphalangage, une vocalisation accrue et une activité motrice inférieure a été observée chez les souriceaux ainsi que des effets à long terme sous la forme d’une force réduite de préhension et des capacités de nage inférieure chez les adultes. Pour conclure, il existe trop peu de preuves pour évaluer de manière fiable la gêne due au déphalangage ou au perçage des oreilles et la qualité des preuves disponibles n’est pas précise. Des études de qualité élevée et suffisamment approfondies indiquant des résultats fiables et pertinents sont requises pour évaluer de manière précise l’impact de ces techniques d’identification. Jusqu’à ce que des preuves plus fiables soient disponibles, un effet du déphalangage ou du perçage des oreilles sur le bien-être des animaux et les résultats de l’étude ne peuvent pas être exclus ou confirmés. Cette étude a été financée par l’organisation néerlandaise de la recherche et du développement médical [#114025500].

Abstract

Zehenmarken und Ohrlochung (auch Ohrkerbung, Ohrstanzen) sind häufig verwendete Methoden zur Identifizierung von einzelnen Labornagern. Diese Prozeduren verursachen potenziell starke Beschwerden, die das Tierwohl beeinträchtigen und Versuchsergebnisse verzerren können. Gleichwohl existiert derzeit kein systematischer Überblick über die Evidenz zu diesem Thema. Wir führten eine systematische Bewertung der Evidenz für Beschwerden aufgrund von Zehen- oder Ohrmarken bei Nagern durch. Die Bewertungsmethode war in einem registrierten Bewertungsprotokoll vorgegeben. Die PICO-Fragestellung war: Welche Auswirkung hat Zehen- oder Ohrmarkierung bei Nagern gegenüber keiner Markierung bzw. Scheinmarkierung auf tierwohlbezogene Ergebnisse? Anhand systematischer Recherche in Pubmed, EMBASE, Web of Science und Grauer Literatur identifizierten wir sieben Studien zu Auswirkungen von Ohrmarken auf das Tierwohl sowie fünf derartige Studien zu Zehenmarken. Studien wurden in die Bewertung einbezogen, wenn sie Originaldaten eines In-vivo-Experiments mit Nagern enthielten und die Auswirkung von Ohr- oder Zehenmarken auf einen tierwohlbezogenen Outcome beurteilten. Fallstudien und Studien unter Einsatz ungeeigneter Ko-Interventionen wurden ausgeschlossen. Die Studienqualität wurde mithilfe einer erweiterten Version des SYRCLE Risk of Bias Tools für Tierstudien ermittelt. Studienmerkmale und Ergebnismessungen waren sehr heterogen und es gab ein unklares oder hohes Verzerrungsrisiko in allen Studien. Wir präsentieren daher eine beschreibende Synthese identifizierter Evidenz. Keine der Studien berichtete eine Fallzahlplanung. Bei über 60 unterschiedlichen Ergebnissen fanden wir Evidenz für Beschwerden aufgrund von Ohrmarken in Form einer Erhöhung von Atemvolumen, Vokalisierung und Blutdruck. Für Zehenmarken wurden erhöhte Vokalisierung und verminderte motorische Aktivität bei Jungtieren ermittelt, ebenso wie Langzeiteffekte in Form reduzierter Greifkraft und Schwimmfähigkeit bei adulten Tieren. Als Fazit lässt sich feststellen, dass zu wenig Evidenz verfügbar ist, um eine zuverlässige Bewertung von Beschwerden aufgrund von Zehen- oder Ohrmarken zu geben und dass die Qualität verfügbarer Evidenz zweifelhaft ist. Es bedarf adäquater gepowerter, aussagekräftiger Studien, die zuverlässige, relevante Ergebnismessungen berichten, um die Auswirkungen dieser Identifizierungsmethoden akkurat bewerten zu können. Solange keine verlässlichere Evidenz vorliegt, kann ein Einfluss von Zehen- oder Ohrmarken auf Tierwohl und Studienergebnisse weder ausgeschlossen...
Resumen

El corte de dedos y orejas (corte o perforación en orejas) suelen ser métodos habituales para la identificación de roedores de laboratorio. Estos métodos potencialmente causan una gran molestia, lo cual puede reducir el bienestar del animal y alterar los resultados de los experimentos. No obstante, no existe ningún resumen sistemático de la evidencia de este tema actualmente. Llevamos a cabo una revisión sistemática de la evidencia de molestias debido al corte de dedos u orejas en roedores. La metodología de revisión fue pre-especificada en un protocolo de revisión registrado. La pregunta PICO fue: En roedores, ¿cuál es el efecto de recortar dedos u orejas en comparación a no recortar o a un recorte simulado en lo referente a los resultados relacionados con el bienestar? A través de una búsqueda sistemática en PubMed, EMBASE, web de ciencias y literatura no convencional, identificamos siete estudios sobre el efecto del recorte de orejas en el bienestar animal y cinco estudios sobre el corte de dedos. Se incluyeron estudios en el documentos si contenían datos originales de un experimento in vivo con roedores, evaluando el efecto del corte de dedos u orejas en un resultado relacionado con el bienestar. Se excluyeron estudios de caso y estudios que aplicaran cointervenciones no adecuadas. Se estimó la calidad del estudio utilizando una versión ampliada del riesgo SYRCLE de herramientas de parcialidad para estudios con animales. Las características del estudio y las medidas de los resultados fueron altamente heterogéneas, y existía un riesgo alto o impreciso de parcialidad en todos los estudios. Por tanto, presentamos una síntesis narrativa de las evidencias identificadas. Ninguno de los estudios indicó un cálculo de un tamaño de muestra. De entre más de 60 distintos resultados, encontramos pruebas de molestias al cortar orejas que se expresaban con un aumento del volumen respiratorio, la vocalización y la presión sanguínea. Para el corte de dedos, se encontró una mayor vocalización y una disminución de la actividad motriz en crías, además de efectos a largo plazo con una reducción de la fuerza de agarre y la capacidad de nadar en adultos. En conclusión, hay muy pocas pruebas para evaluar de forma fiable cualquier molestia al cortar dedos u orejas, y la calidad de las pruebas disponibles es dudosa. Se requieren estudios de alta calidad y con un poder adecuado para informar sobre medidas de resultados relevantes y fiables con el fin de evaluar el impacto de estas técnicas de identificación. Hasta que no haya disponible más pruebas fiables, las consecuencias para el bienestar animal al cortar dedos y orejas y para los resultados de estudios no pueden excluirse ni confirmarse. Este estudio fue financiado por la Organización para el Desarrollo e Investigación Sanitaria de los Países Bajos (n° 114025500).