Comparative Medicine
Copyright 2024
by the American Association for Laboratory Animal Science

Vol 74, No 1
February 2024
Pages 12–18

Review

Consistency in Reporting of Loss of Righting Reflex for Assessment of General Anesthesia in Rats and Mice: A Systematic Review

Michael Z Teng, BSc,[1,+] Dexter Merenick, BSc,[1,+] Anisha Jessel, BSc,[1] Heather Ganshorn, BA, MLIS,[2] and Daniel S J Pang, BVSc, PhD[1,3,*]

General anesthesia induces a reversible loss of consciousness (LOC), a state that is characterized by the inability to feel pain. Identifying LOC in animals poses unique challenges, because the method most commonly used in humans, responding to questions, cannot be used in animals. For over a century, loss of righting reflex (LORR) has been used to assess LOC in animals. This is the only animal method that correlates directly with LOC in humans and has become the standard proxy measure used in research. However, the reporting of how LORR is assessed varies extensively. This systematic literature review examined the consistency and completeness of LORR methods used in rats and mice. The terms 'righting reflex,' 'anesthesia,' 'conscious,' 'rats,' 'mice,' and their derivatives were used to search 5 electronic databases. The abstracts of the 985 articles identified were screened for indications that the study assessed LORR in mice or rats. Full texts of selected articles were reviewed for LORR methodological completeness, with reported methods categorized by 1) animal placement method, 2) behavioral presence of righting reflex, 3) duration of LORR testing, 4) behavioral LORR, and 5) animal position for testing LORR. Only 22 papers reported on all 5 methodological categories. Of the 22 papers, 21 used unique LORR methodologies, with descriptions of LORR methods differing in at least one category as compared with all other studies. This variability indicates that even papers that included all 5 categories still had substantial differences in their methodological descriptions. These findings reveal substantial inconsistencies in LORR methodology and reporting in the biomedical literature likely compromising study replicability and data interpretation.

Abbreviations and Acronyms: LOC, loss of consciousness; LORR, loss of righting reflex; LTI, loss of righting reflex testing interval; LTT, loss of righting reflex testing threshold; RORR, return of righting reflex

DOI: 10.30802/AALAS-CM-23-000063

Introduction

By definition, general anesthetics cause reversible states of amnesia, unconsciousness, and immobilization.[1,5,7,15] These states are also recognized as major indicators of anesthesia.[15] The ability to induce a safe and reversible state of unconsciousness is associated with loss of sensation and pain. A variety of accepted methods have been used to measure loss of consciousness (LOC) in humans, with the loss of response to a verbal command being a common primary clinical end point.[13] Unfortunately, this measure cannot be easily applied to other species, such as mice or rats, which are estimated to represent approximately 95% of all mammals used in research.[3,8] Rats and mice are also the 2 species most commonly used in pain research.[17]

Numerous approaches have been used to identify LOC in rats and mice, such as immobility to a noxious stimulus, tail-flick in response to a temperature stimulus, the loss of righting reflex (LORR), and learning or memory end points.[24] The predominant measure used in research for over a century has been LORR, which has a strong positive correlation with LOC in humans.[5] When an animal is placed on its back (dorsal recumbency), the righting reflex activates vestibular organs, eliciting a return to a normal position with respect to gravity (ventral/sternal recumbency).[23] Although LORR in animals has been frequently used for research, the literature appears to be inconsistent with regard to how LORR is performed and assessed across studies. If true, these inconsistencies increase the likelihood of discrepancies in results between studies, limiting the ability to compare results and studies.

A landmark study highlighted concerning oversights the reporting of animal-based research using animals is reported, with a substantial proportion of studies failing to report even the hypothesis or objective of the study;[12] a similar pattern occurs in veterinary clinical research.[2,19-21]

A response to incomplete reporting has been the development of reporting guidelines (for example, ARRIVE [Animals in Research: Reporting In Vivo Experiments]).[12] However, the existence of guidelines has not guaranteed improved reporting.[2,19-21] Incomplete reporting is associated with increased

Submitted: 30 Sep 2023. Revision requested: 20 Nov 2023. Accepted: 23 Jan 2024.
[1]Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada; [2]Libraries and Cultural Resources, University of Calgary, Calgary, Alberta, Canada; and [3]Faculty of Veterinary Medicine, University of Montreal, Montreal, Quebec, Canada
[+]Equally contributed to the article
**Corresponding author. Email: dsjpang@ucalgary.ca*
This article contains supplemental materials online.

effect sizes and irreproducible results and associated ethical and financial costs, conservatively estimated to be in the region of $28 billion per annum in the United States alone.[6,16] These consequences of incomplete reporting are compounded in that most research focuses on novel rather than replication studies, particularly because little funding is available for replication studies,[4,10,18] further emphasizing the importance of carefully conducted and reported studies. The goal of this systematic review was to examine, summarize, and compare the descriptions of LORR that appear in peer-reviewed publications.

Materials and Methods

Experimental design. We performed a systematic search for studies reporting the use of LORR. The following databases were searched from their inception through 6 June 2022 (Figures 1 and 2). The search strategies that were used are provided in Supplemental Figures S1 and S2.

Abstract screening. Three reviewers performed an initial independent screening of identified abstracts. Selected abstracts were then accessed in full text and reviewed, and the data were extracted independently by 2 reviewers. Differences between reviewers were resolved through discussion to reach a consensus, with the assistance of a third party.

Full text screening. Criteria for inclusion of full text screening required a description of an experiment (or experiments) that included rats, mice, or both and indicated assessment of LOC and LORR as well as assessment of the presence and recovery of righting reflex (RORR). Studies that did not contain this information were excluded. Additional criteria for excluding a paper during full text review were lack of peer review, full text not available, not published in English, or a retracted study.

Data extraction. Studies that mentioned the righting reflex but did not provide a more complete description of the assessment method were included to provide information on completeness of reporting. For example, we interpreted the statement, "A positive RORR was determined to be where the animals righted themselves on two additional toppling stimuli immediately after the first observed RORR" as testing LORR despite the absence of an explicit description in the study, as the righting reflex must be lost to identify its return.[11] Some studies that did not directly mention LORR or RORR in their descriptions were nonetheless clearly were referring to them. For example, one study stated "the ability of animals to right themselves after being turned over was used as a measure of anesthetic effect."[22] We inferred from this statement that the inability to perform such an action would be considered LORR, and that the ability to perform such an action indicates that the righting reflex was either present or had been recovered, even without direct mention of either LORR or RORR.

Data table. During data extraction, LORR information was tabulated under the following 5 categories: 1) animal placement method, 2) behavioral presence of righting reflex, 3) duration of LORR testing, 4) behavioral LORR, and 5) animal position for testing LORR. The 5 categories were tabulated according to species and the general anesthetic used. The category 'duration of LORR testing' included 2 methods for describing time. The first method tested LORR over a fixed duration (for example, LORR is considered to have occurred when the rodent remained on its back for 15 s); this will be referred to as the LORR testing threshold (LTT). The second method assessed LORR at distinct intervals (for example, the rat was placed on its back every 15 s for 1 min). This will be referred to as the LORR testing interval (LTI). All experiments were expected to provide a LTT for confirmation of LORR. Depending on the experimental design, not all experiments needed an LTI. If LORR was only tested once, there would be no LTI.

For table data entry, if no specific information for a category was reported from a study, it was described as 'unspecified'; if information was reported, it was referred to as 'specified'. After the table was completed, the number of specified categories from each study were given numeric values and converted into a percentage. For instance, a study that had specified descriptions for 4 of the 5 categories would have a reporting rate of 80% (4 of 5). Studies that tested more than one strain or anesthesia method were calculated differently. For example, a study might test both rats and mice and report 3 out of the 5 categories for mice, and 4 out of 5 categories for rats. Overall, the 2 experiments from this one study would have an overall reporting rate of 70% (4 of 5 + 3 of 5 = 7 of 10).

Summary table. When creating the summary table, synonymous terms encountered during data entry (for example, 'dorsal recumbency,' 'dorsal,' 'supine,' 'back') were grouped under a single descriptor (for example, 'dorsal recumbency'). In cases where descriptors were nonspecific (vague), such as body position during LORR testing, for example, these studies were categorized under the single term of 'unspecified' for the

Database	Platform/Vendor
MEDLINE	Ovid
Embase	Ovid
CAB Abstracts	Ebsco
BIOSIS Previews	Web of Science (Clarivate)
Web of Science Core Collection – Science Citation Index & Emerging Sources Citation Index	Web of Science (Clarivate)

Figure 1. Systematic search of databases.

Study Characteristic:	Description:
Population	Rats and Mice
Intervention	Loss of righting reflex (LORR)
Outcome	Assessing loss of consciousness (LOC) under induction of general anesthesia

Figure 2. Concepts used to create the search and inclusion criteria. Complete search strategies for each database can be found in Figures S1 and S2.

purpose of analysis, recognizing that 'unspecified' encompassed a variety of different nonspecific descriptors.

Data analysis. Data from studies published before 2010 were compared with those published in 2015 and thereafter by using a Kolmogorov-Smirnov test for unpaired data. Results were considered significant if $P < 0.05$. These years were selected to compare reporting before and after publication of the ARRIVE guideline in 2015. These time periods were also compared in an earlier retrospective, observational cohort study that examined the impact of the ARRIVE guidelines on research reporting.[13]

Risk of bias assessment. The systematic review design was specific to the reporting of LORR methodology and did not assess the effect of treatment. Therefore, a risk of bias assessment was not performed.

Registration and protocol. This systematic review did not have a protocol registered. Data supporting the presented results are available in a data repository: https://doi.org/10.7910/DVN/ETTXVE

Results

We initially imported 2,088 studies for screening and removed 1,103 duplicates. The remaining 985 were used for abstract screening, with 42 studies identified as unrelated to the topic of interest. This resulted in 943 studies that were assessed as full texts. This assessment excluded 260 studies for the following reasons: 118 for lack of peer review, 82 due to unavailable full texts, 35 because they were not being published in English, 11 because they did not test LORR, 8 that they did not use rats or mice, 5 for not being an experimental or observational study, and one due to being a retracted study (Figure 3).

The 683 included papers provide data from 750 experiments. Sixteen studies tested both rats and mice using a single anesthesia method, 47 tested one species with using 2 anesthesia methods, and 2 studies tested 3 anesthesia methods on a single species. Because of the volume of data, results are presented in a summary table. Specific details on the study, the supporting quotes for LORR methods, the species, and the anesthesia method tested are provided in Supplemental Tables S1 to S6.

Table 1. Key reporting details are displayed in percentages for the 5 LORR methodological categories

LORR methodological category	Percentage of unreported LORR methods ('unspecified')	Percentage of reported LORR methods ('specified')
Animal placement method	83.3	16.7
Behavioral presence of righting reflex	59.2	40.8
Duration of LORR testing	58.5	41.5
Behavioral loss of righting reflex	54.8	45.2
Animal position for testing LORR	44.5	55.5

Summary table analysis. Our analysis found that 60% of the 5 categories were 'unspecified' and therefore essentially unreported. The separate methodological categories ranged between 45% and 83%, but 'unspecified' was always the most common descriptor (Table 1).

The most frequent description used for the category 'animal placement method' was 'the container was rotated at an unspecified angle' (Table 2); this description applied to 29% of all specified LORR methods (Tables 1 and 2). The most frequent description used for the category 'behavioral presence of the righting reflex' was 'animal returned to sternal recumbency/normal position' (Table 3); this description applied to 32% of all specified LORR methods (Tables 1 and 3). The most frequent description used for the category 'duration of LORR testing ' was '30 seconds' (Table 4); this description applied to 23% of all specified LORR methods (Tables 1 and 4).

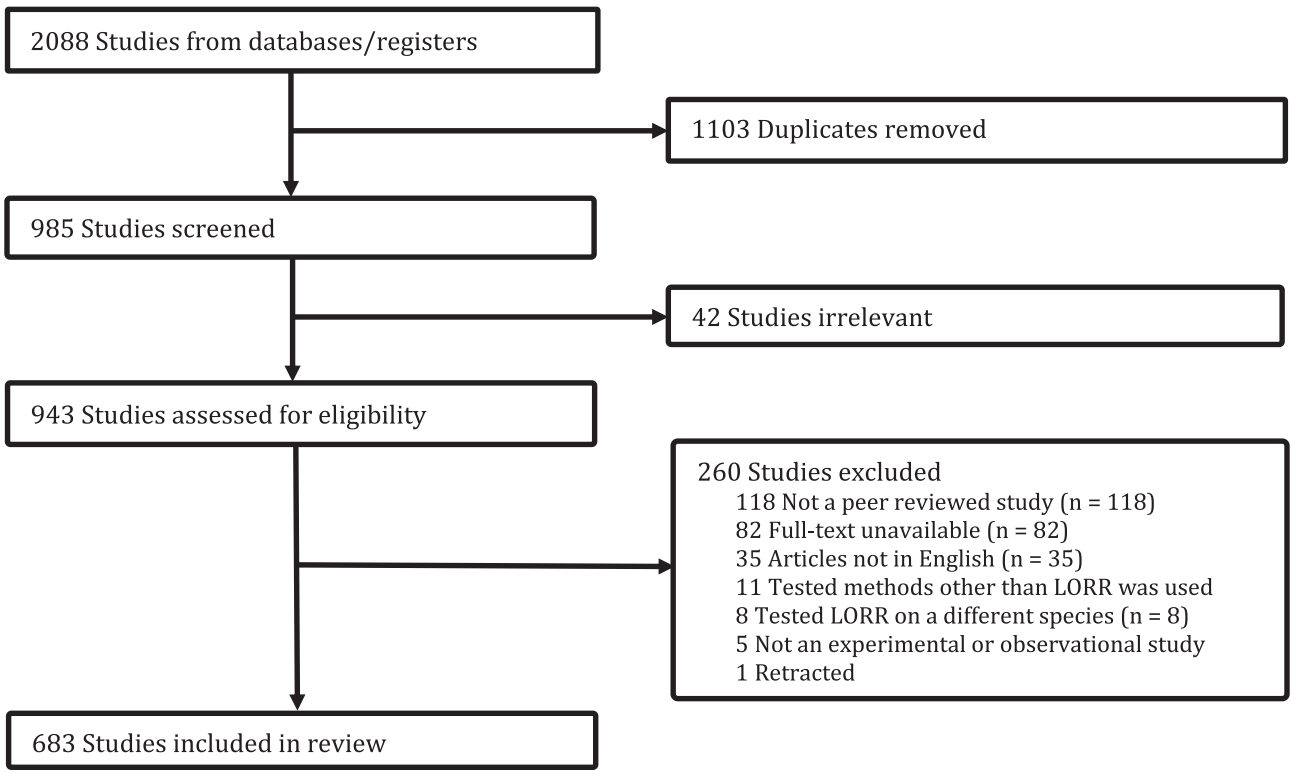

Figure 3. Flow chart of the search strategy used to identify articles for inclusion and exclusion.

Table 2. Summary table for the category 'animal placement method'[a]

Animal placement method (n = 750)	Percentage of experiments with the same description out of all experiments
Positioned by hand	0.9 (n = 7)
The container was rotated 15°	0.1 (n = 1)
The container was rotated 30°	0.8 (n = 6)
The container was rotated 45°	1.7 (n = 13)
The container was rotated 90°	0.9 (n = 7)
The container was rotated 180°	2.1 (n = 16)
The container was rotated at an unspecified angle	4.9 (n = 37)
The container was rotated twice in rapid succession	0.1 (n = 1)
The container was rotated automatically at 1 rpm	0.1 (n = 1)
The container was rotated automatically at 2 rpm	0.1 (n = 1)
The container was rotated automatically at 3 rpm	0.5 (n = 4)
The container was rotated automatically at 4 rpm	2.4 (n = 18)
The container was rotated automatically at 5 rpm	0.3 (n = 2)
The container was rotated automatically at 20 rpm	0.1 (n = 1)
The container was rotated automatically at an unspecified speed	1.2 (n = 9)
Specific unique methodology	0.9 (n = 7)
Unspecified	83.3 (n = 625)

Note that whereas there are only 750 total experiments, one experiment can have multiple descriptions (756 total descriptions).
[a]A list of paper references for each experiment, as well as the full summary table and data table, is available in Tables S1 and S2

The most frequent description used for the category 'behavioral loss of righting reflex' was 'animal failed to return to sternal recumbency/normal position' (Table 5); this description applied to 22% of all specified LORR methods (Tables 1 and 5). The most frequent description for the category 'animal position for testing LORR ' was 'dorsal recumbency' (Table 6); this description applied to 78% of all specified LORR methods (Tables 1 and 6).

Overall data analysis. Of the 5 categories, the median reporting percentage was 40% (range, 0% to 100%) (Table 1). The median reporting percentage for studies published before 2010 (n = 375) was 40% (range, 0% to 100%), and the median percentage during and after 2015 (n = 209) was also 40% (range, 0% to 100%), with significant difference in reporting between the 2 time periods (P = 0.80).

Of the 683 studies examined in this systematic review, 200 had their methods classified as unspecified for all categories examined, whereas only 22 had reported on all categories. These 22 studies included 21 unique LORR methodologies, meaning that the methods described for at least one category differed from every other study in the group (Figure 4).

Discussion

This systematic review showed that in the majority of studies, LORR methodology was not described completely. The studies that did provide details used that many different methods for LORR testing across the 5 categories assessed (animal placement method, behavioral presence of righting reflex, duration of LORR testing, behavioral LORR, and animal position for testing LORR). Incomplete and inconsistent reporting limit the ability to compare results among studies and prevent accurate study replication. This problem is common in many different areas of scientific research.[2,14,19-21] Although reporting guidelines (such as ARRIVE) have been created to address this issue,[11] the introduction of guidelines has not resulted in significant improvement in reporting.[2,14,17]

The completeness of reporting in each methodological category was disturbingly low. No single specified description for any category constituted an absolute majority (>50%) of experiments, with the highest being the specified description for animal position for testing LORR at 44% (dorsal recumbency). This was the only category in which 'unspecified' was not the absolute majority in descriptions, yet 'unspecified' was the most common description in all experiments (45%). Furthermore, nearly a third of all studies were not specific about any of the 5 categories we examined, providing specifics only on the species and method of general anesthesia used.

Table 3. Summary table for the category 'behavioral presence of righting reflex'[a]

Behavioral presence of righting reflex (n = 750)	Percentage of experiments with the same description out of all experiments
Animal returned to sternal recumbency/normal position	13.2 (n = 99)
Animal has ability to right themselves	2.1 (n = 16)
Ability to perform 2 consecutive righting reflexes	2.4 (n = 18)
Ability to perform 3 consecutive righting reflexes	8.1 (n = 61)
Righting themselves so any paw touches the surface they were laid on	0.1 (n = 1)
Righting themselves so both forepaws touch the surface they were laid on	0.1 (n = 1)
Righting themselves so at least 3 paws touch the surface they were laid on	0.3 (n = 2)
Righting themselves so all 4 paws touch the surface they were laid on	7.5 (n = 56)
Animal has ability to remain on paws through one rotation	0.3 (n = 2)
Return to ambulation	1.3 (n = 10)
When the animal made a purposeful attempt to right themselves	0.4 (n = 3)
Specific scoring system for righting reflex/depth of anesthesia	4.1 (n = 31)
Specific unique methodology	2.5 (n = 19)
Unspecified	59.2 (n = 444)

Note that whereas there are only 750 total experiments, one experiment can have multiple descriptions (763 total descriptions).
[a]A list of paper references for each experiment, as well as the full summary table and data table, is available in Tables S1 and S3.

Table 4. Summary table for the category 'duration of loss of righting reflex testing'[a]

Duration of LORR testing (n = 750)	Percentage of experiments with the same description out of all experiments
LTT LORR was considered to have occurred after:	
2 s	0.4% (n = 3)
3 s	0.1% (n = 1)
5 s	2.8% (n = 21)
8 s	0.1% (n = 1)
10 s	6.7% (n = 50)
15 s	3.6% (n = 27)
20 s	0.7% (n = 5)
30 s	9.6% (n = 72)
1 min	5.2% (n = 39)
1 min 15 s	1.3% (n = 10)
2 min	0.7% (n = 5)
5 min	0.5% (n = 4)
15 min	0.1% (n = 1)
20 min	0.4% (n = 3)
40 min	0.1% (n = 1)
LTI, LORR was tested every:	
10 s	0.7% (n = 5)
10–15 s	0.3% (n = 2)
15 s	2.5% (n = 19)
15–30 s	0.5% (n = 4)
20 s	0.3% (n = 2)
30 s	1.7% (n = 13)
1 min	1.2% (n = 9)
2 min	2.1% (n = 16)
3 min	1.1% (n = 8)
4 min	0.3% (n = 2)
5 min	1.7% (n = 13)
10 min	0.3% (n = 2)
12 min	0.1% (n = 1)
15 min	0.4% (n = 3)
30 min	0.1% (n = 1)
Both LTT and LTI	
Specific unique methodology	1.7% (n = 12)
Unspecified	58.5% (n = 439)

Note that whereas there are only 750 total experiments, one experiment can have multiple descriptions (794 total descriptions).
[a]A list of paper references for each experiment, as well as the full summary table and data table, is available in Tables S1 and S4.

Table 5. Summary table for the category 'behavioral loss of righting reflex'[a]

Behavioral loss of righting reflex (n = 750)	Percentage of experiments with the same description out of all experiments
Failed to return to sternal recumbency/normal	10.1% (n = 76)
Failed to return to sternal recumbency/ normal position or lateral recumbency	0.3% (n = 2)
Animal did not right themselves	7.7% (n = 58)
Animal remained in position for testing of LORR	4.9% (n = 37)
Animal did not respond when placed in position for testing of LORR	0.3% (n = 2)
Animal is immobile	0.5% (n = 4)
Animal had at least 3 paws up in the air while in dorsal recumbency	2.3% (n = 17)
Animal failed to right themselves onto all 4 paws	6.3% (n = 47)
Animal failed to perform 2 consecutive righting reflexes	0.4 (n = 3)
Animal failed to perform 3 consecutive righting reflexes	3.1% (n = 23)
Animal rolled onto side and all 4 paws left the surface of the rotating tube	0.4% (n = 3)
Animal rolled on their backs and did not attempt to right themselves	1.1% (n = 8)
Animal rolled on their backs on at least 2 out of 5 complete rotations	2.1% (n = 16)
Specific scoring system for righting reflex/depth of anesthesia	2.4% (n = 18)
Specific unique methodology	5.3% (n = 40)
Unspecified	54.8% (n = 411)

Note that while there are only 750 total experiments, one experiment can have multiple descriptions (765 total descriptions).
[a]A list of paper references for each experiment, as well as the full summary table and data table, is available in Tables S1 and S5.

Table 6. Summary table for the category 'animal position for testing LORR'[a]

Animal position for testing LORR (n = 750)	Percentage of experiments with the same description out of all experiments
Dorsal recumbency	43.5% (n = 326)
Lateral recumbency	4.5% (n = 34)
Ventral recumbency	1.1% (n = 8)
Dorsal or lateral recumbency	2.5% (n = 19)
Constantly rolling	3.5% (n = 26)
Specific unique methodology	0.4% (n = 3)
Unspecified	44.5% (n = 334)

Note that while there are only 750 total experiments, one experiment can have multiple descriptions (750 total descriptions).
[a]A list of paper references for each experiment, as well as the full summary table and data table, is available in Tables S1 and S6.

Consistency in descriptions was also low in the other categories. Even in studies with specific descriptions, many different descriptions were used (Tables 2–6). Twenty-two studies reported on all LORR categories. Of these, 2 used the same methodology, but the remaining 20 had at least one different description. This equates to 21 total unique LORR methodologies for the 22 studies, emphasizing the inconsistencies between descriptions of LORR methodology even in studies that reported completely on their LORR methods.

The category with the highest reporting rate was animal position for testing LORR, which had a reporting rate that was over twice as great as the category with the lowest reporting rate (animal placement method; Table 1). This result is surprising because one might expect that the testing position and placement method would both be described. Descriptions of animal placement methods may be underreported due to the common practice of placing animals by hand into the required test position. However, without a specified description, this remains an assumption. Only 7 of 750 experiments explicitly state that researchers used their hands to place the animal into the test position.

Complete reporting of study methods is fundamental to enabling experimental replicability and informing future

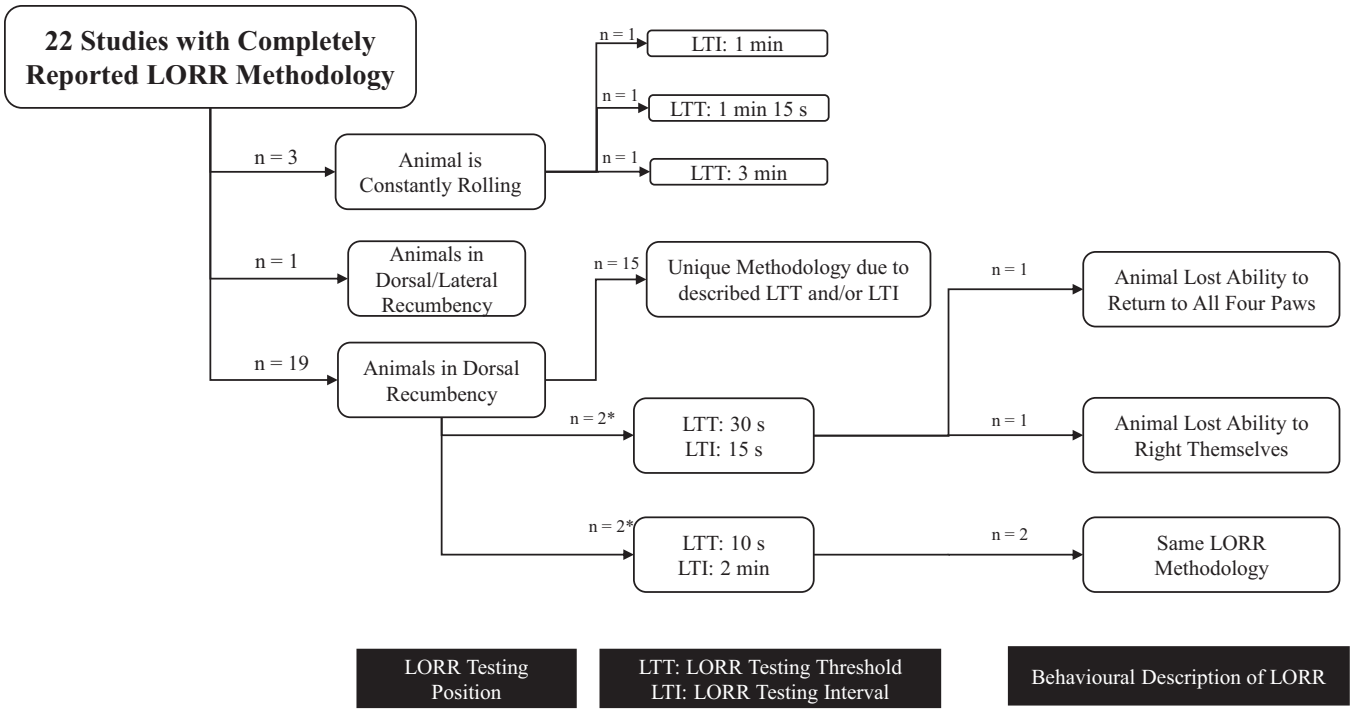

Figure 4. Tree diagram consisting of information from the 22 studies that fully reported each of the 5 categories for describing loss of righting reflex (LORR): 1) animal placement method, 2) behavioral presence of righting reflex, 3) duration of LORR testing, 4) behavioral LORR, and 5) animal position for testing LORR. To see the 22 studies used in this diagram, see Table S2. n = number of studies. *Studies that had the same descriptions for categories 1 and 2.

investigations in the field. The first version of the ARRIVE guidelines was published in 2010 to address the issue of incomplete reporting in animal research.[12] However, the existence of reporting guidelines, accompanied by widespread publicity, has improved reporting to a very limited extent.[2,14,19-21] Unlike a previous study,[14] our analysis revealed no difference in reporting completeness among studies published before or after the creation of the ARRIVE guidelines. Our analysis was simply intended to determine whether availability of the guidelines had resulted in a significant increase in the completeness of methodology reporting. Therefore, our analysis, which used a time frame intended to allow such comparison, indicates that, at least in terms of LORR reporting, the available guidelines have had little impact.

The major issue uncovered in our analysis was the incompleteness of LORR methodological reporting, given that even small differences in LORR methodology can lead to substantial differences in results. Two pre-2015 studies that provided substantial detail of their methodology both assessed LORR duration after administration of an intraperitoneal injections of 50 g/kg of pentobarbital in male Sprague-Dawley rats.[9,25] Despite consistency of many key factors (sex, species, strain, drug, dose, injection route, initial animal position, similar age, and weight), one reported a LORR duration of 148 ± 3 min (mean ± SD) and the other reported a LORR duration of 103 ± 3 min (mean ± SD).[9,25] This 45-min difference in LORR durations under nearly identical conditions could be partially attributed to differences in how recovery from LORR was determined. One paper stated, "timing was terminated when the animals could right themselves three times in 30 sec."[25] The other stated, "return of righting reflex was recorded when the animal had righted itself three times within 1 min."[9] Because one study allowed twice the time for righting as did the other, the shorter LORR duration results in one study might be expected. This example highlights how slight changes in LORR methodology can greatly influence the results reported.

Given the variability of LORR methods revealed by our analysis, recommendation of a specific LORR methodology is compromised by the lack of research that directly compares different methodologies. For instance, while manual positioning an animal during LORR assessment might provide a source of stimulation, the extent to which this might affect testing as compared with an automated rotating chamber is unknown. However, studies might be more easily compared if the most frequently specified descriptions for LORR testing were used. Specifically, this would mean that testing consists of placing an animal in dorsal recumbency using a container. If the animal rights itself to sternal recumbency within 30 s, the righting reflex is intact. If the animal fails to return to sternal recumbency within 30 s, the righting reflex is lost.

This study was conducted under some limitations to avoid making biased assumptions about experimental reporting. We were strict in identifying items considered 'unspecified.' For example, if a study did not describe LORR but did describe the presence of a righting reflex, LORR was still classified as unspecified. This approach may have underestimated the completeness of LORR methodological reporting, but we wanted to avoid false assumptions and instead rely on what was specifically reported. This approach also allowed identification of studies that did provide a complete description of the methods.

In conclusion, this systematic review identified longstanding persistent problems in the reporting of LORR methodology. These results suggest considerable limitations to study replication and comparison between studies due to the incompleteness of reporting and inconsistency in methods for LORR.

Supplemental Materials

Figure S1. Search strategies for databases: Ovid MEDLINE and Epub ahead of print, in-process, in-data-review, and other nonindexed citations and daily, 1946 to 9 Jun 2022

Figure S2. Search strategies for database: Embase, 1974 to 9 Jun 2022

Table S1. Data table with full reference list and extracted data

Table S2. Summary table for the category 'animal placement method' with every study referenced and separated by species and method of general anesthesia

Table S3. Summary table for the category 'behavioral presence of righting reflex' with every study referenced and separated by species and method of general anesthesia

Table S4. Summary table for the category 'duration of LORR testing' with every study referenced and separated by species and method of general anesthesia

Table S5. Summary table for the category 'behavioral loss of righting reflex' with every study referenced and separated by species and method of general anesthesia

Table S6. Summary table for the category 'animal position for testing LORR' with every study referenced and separated by species and method of general anesthesia

Conflict of Interest

The author(s) have no conflict(s) of interest to declare.

Funding

This work was internally funded.

References

1. **Antognini JF, Carstens E.** 2002. In vivo characterization of clinical anaesthesia and its components. Br J Anaesth **891**:156–166. https://doi.org/10.1093/bja/aef156.
2. **Bergen P, Munro BA, Pang DS.** 2023. Quality of reporting of prospective in vivo and ex vivo studies published in the J Vet Emerg Crit Car over a 10-year period 2009–2019. J Vet Emerg Crit Car **33**:435–441. https://doi.org/10.1111/vec.13312.
3. **Carbone L.** 2021. Estimating mouse and rat use in American laboratories by extrapolation from Animal Welfare Act–regulated species. Sci Rep-UK **111**:493. https://doi.org/10.1038/s41598-020-79961-0.
4. **Flier JS.** 2022. The problem of irreproducible bioscience research. Perspect Biol Med **65**:373–395. https://doi.org/10.1353/pbm.2022.0032.
5. **Franks NP.** 2008. General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. Nat Rev Neurosci **95**:370–386. https://doi.org/10.1038/nrn2372.
6. **Freedman LP, Cockburn IM, Simcoe TS.** 2015. The economics of reproducibility in preclinical research. PLoS Biol **136**:e1002165. https://doi.org/10.1371/journal.pbio.1002165.
7. **Grasshoff C, Rudolph U, Antkowiak B.** 2005. Molecular and systemic mechanisms of general anaesthesia: the 'multi-site and multiple mechanisms' concept. Curr Opin Anaesthesiol **184**:386–391. https://doi.org/10.1097/01.aco.0000174961.90135.dc.
8. **Hickman DL, Johnson J, Vemulapalli TH, Crisler JR, Shepherd R.** 2017. Commonly used animal models, p 117–175. In: Suckow MA, Stewart KL, editors. Principles of animal research for graduate and undergraduate students. Cambridge (MA): Academic Press.
9. **Horita A, Carino MA, Ukai Y.** 1994. The analeptic effect of methamphetamine in pentobarbital-narcotized rats is mediated via a dopaminergic-cholinergic mechanism. J Pharmacol Exp Ther **268**:311–318.
10. **Ioannidis JP.** 2012. Why science is not necessarily self-correcting. Perspect Psychol Sci **7**:645–654. https://doi.org/10.1177/1745691612464056.
11. **Jacobson GM, Voss LJ, Melin SM, Cursons RTM, Sleigh JW.** 2011. The role of connexin36 gap junctions in modulating the hypnotic effects of isoflurane and propofol in mice. Anaesthesia **66**:361–367. https://doi.org/10.1111/j.1365-2044.2011.06658.x.
12. **Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG.** 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. J Pharmacol Pharmacother **12**:94–99. https://doi.org/10.4103/0976-500X.72351.
13. **Kim DW, Kil HY, White PF.** 2002. Relationship between clinical endpoints for induction of anesthesia and bispectral index and effect-site concentration values. J Clin Anesth **144**:241–245. https://doi.org/10.1016/S0952-8180(02)00348-3.
14. **Leung V, Rousseau-Blass F, Beauchamp G, Pang DS.** 2018. ARRIVE has not ARRIVEd: support for the ARRIVE Animal Research: reporting of in vivos guidelines does not improve the reporting quality of papers in animal welfare, analgesia or anesthesia. PLoS One **135**:e0197882. https://doi.org/10.1371/journal.pone.0197882.
15. **Mashour GA, Forman SA, Campagna JA.** 2005. Mechanisms of general anesthesia: from molecules to mind. Best Pract Res Cl Anaesth **193**:349–364. https://doi.org/10.1016/j.bpa.2005.01.004.
16. **McCann SK, Cramond F, Macleod MR, Sena ES.** 2016. Systematic review and meta-analysis of the efficacy of interleukin-1 receptor antagonist in animal models of stroke: an update. Transl Stroke Res **7**:395–406. https://doi.org/10.1007/s12975-016-0489-z.
17. **Mogil JS.** 2009. Animal models of pain: progress and challenges. Nat Rev Neurosci **104**:283–294. https://doi.org/10.1038/nrn2606.
18. **Munafò MR, Nosek BA, Bishop DV, Button KS, Chambers CD, Percie du Sert N, Simonsohn U, Wagenmakers E-J, Ware JJ, Ioannidis J.** 2017. A manifesto for reproducible science. Nat Hum Behav **1**:1–9. https://doi.org/10.1038/s41562-016-0021.
19. **Munro BA, Bergen P, Pang DS.** 2022. Randomization, blinding, data handling and sample size estimation in papers published in Vet Anaesth Analg in 2009 and 2019. Vet Anaesth Analg **491**:18–25. https://doi.org/10.1016/j.vaa.2021.09.004.
20. **Rufiange M, Rousseau-Blass F, Pang DS.** 2019. Incomplete reporting of experimental studies and items associated with risk of bias in veterinary research. Vet Rec Open **61**:e000322. https://doi.org/10.1136/vetreco-2018-000322.
21. **Sargeant JM, Elgie R, Valcour J, Saint-Onge J, Thompson A, Marcynuk P, Snedeker K.** 2009. Methodological quality and completeness of reporting in clinical trials conducted in livestock species. Prev Vet Med **912-4**:107–115. https://doi.org/10.1016/j.prevetmed.2009.06.002.
22. **Shih CL, Chen HH, Chiu TH.** 2001. Acute exposure to trichloroethylene differentially alters the susceptibility to chemoconvulsants in mice. Toxicology **162**:35–42. https://doi.org/10.1016/S0300-483X(01)00358-4.
23. **Simpson JA, Fitch W.** 1988. Reflex activity, p 155–170. In: Applied neurophysiology: with particular reference to anaesthesia. Cambridge (Great Britain): Butterworth–Heinemann.
24. **Zalucki O, Van Swinderen B.** 2016. What is unconsciousness in a fly or a worm? A review of general anesthesia in different animal models. Conscious Cogn **44**:72–88. https://doi.org/10.1016/j.concog.2016.06.017.
25. **Zucker J, Calkins D, Zabawska J, Lai H, Horita A.** 1987. Effects of intraseptal drug administration on pentobarbital-induced narcosis and hippocampal choline uptake. Pharmacol Biochem Behav **28**:433–436. https://doi.org/10.1016/0091-3057(87)90501-6.