

# Systematic review and meta-analysis of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators tested for antinociceptive effects in animal models of injury-related or pathological persistent pain

Nadia Soliman<sup>a,\*</sup>, Simon Haroutounian<sup>b</sup>, Andrea G. Hohmann<sup>c</sup>, Elliot Krane<sup>d</sup>, Jing Liao<sup>e</sup>, Malcolm Macleod<sup>e</sup>, Daniel Segelcke<sup>f</sup>, Christopher Sena<sup>e</sup>, James Thomas<sup>g</sup>, Jan Vollert<sup>a</sup>, Kimberley Wever<sup>h</sup>, Harutyun Alaverdyan<sup>b</sup>, Ahmed Barakat<sup>i,j</sup>, Tyler Barthlow<sup>j</sup>, Amber L. Harris Bozer<sup>k</sup>, Alexander Davidson<sup>l</sup>, Marta Diaz-delCastillo<sup>j</sup>, Antonina Dolgorukova<sup>m</sup>, Mehnaz I. Ferdousi<sup>n</sup>, Catherine Healy<sup>n</sup>, Simon Hong<sup>o</sup>, Mary Hopkins<sup>n</sup>, Arul James<sup>p</sup>, Hayley B. Leake<sup>q,r</sup>, Nathalie M. Malewicz<sup>s</sup>, Michael Mansfield<sup>t</sup>, Amelia K. Mardon<sup>q</sup>, Darragh Mattimoe<sup>n</sup>, Daniel P. McLoone<sup>n</sup>, Gith Noes-Holt<sup>u</sup>, Esther M. Pogatzki-Zahn<sup>f</sup>, Emer Power<sup>n</sup>, Bruno Pradier<sup>f</sup>, Eleny Romanos-Sirakis<sup>v,w</sup>, Astra Segelcke<sup>x</sup>, Rafael Vinagre<sup>y</sup>, Julio A. Yanes<sup>z</sup>, Jingwen Zhang<sup>aa</sup>, Xue Ying Zhang<sup>a</sup>, David P. Finn<sup>n</sup>, Andrew S.C. Rice<sup>a</sup>

## Abstract

We report a systematic review and meta-analysis of studies that assessed the antinociceptive efficacy of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators on pain-associated behavioural outcomes in animal models of pathological or injury-related persistent pain. In April 2019, we systematically searched 3 online databases and used crowd science and machine learning to identify studies for inclusion. We calculated a standardised mean difference effect size for each comparison and performed a random-effects meta-analysis. We assessed the impact of study design characteristics and reporting of mitigations to reduce the risk of bias. We meta-analysed 374 studies in which 171 interventions were assessed for antinociceptive efficacy in rodent models of pathological or injury-related pain. Most experiments were conducted in male animals (86%). Antinociceptive efficacy was most frequently measured by attenuation of hypersensitivity to evoked limb withdrawal. Selective cannabinoid type 1, cannabinoid type 2, nonselective cannabinoid receptor agonists (including delta-9-tetrahydrocannabinol) and peroxisome proliferator-activated receptor- $\alpha$  agonists (predominantly palmitoylethanolamide) significantly attenuated pain-associated behaviours in a broad range of inflammatory and neuropathic pain models. Fatty acid amide hydrolase inhibitors, monoacylglycerol lipase inhibitors, and cannabidiol significantly attenuated pain-associated behaviours in neuropathic pain models but yielded mixed results in inflammatory pain models. The reporting of criteria to reduce the risk of bias was low; therefore, the studies have an unclear risk of bias. The value of future studies could be enhanced by improving the reporting of methodological criteria, the clinical relevance of the models, and behavioural assessments. Notwithstanding, the evidence supports the hypothesis of cannabinoid-induced analgesia.

**Keywords:** Cannabinoids, Cannabis-based medicine, Endocannabinoid system modulator, Animal models, Pain, Systematic review and meta-analysis, Preclinical

## 1. Introduction

Cannabinoids, cannabis-based medicines, and endocannabinoid system modulators as potential therapeutics for pain management are of increasing research interest. The endocannabinoid system, composed of the cannabinoid type 1 (CB<sub>1</sub>) and cannabinoid type 2 (CB<sub>2</sub>) receptors, their endogenous ligands, and the enzymes that metabolise the endogenous ligands, is implicated in pain modulation *in vivo*. Hence, cannabinoids, cannabis-based medicines, and endocannabinoid system modulators as possible therapeutics for pain management have been studied extensively in animal models (reviewed most recently in our companion narrative review Finn et al.<sup>19</sup> and Refs. 48 and 56). **Table 1** provides examples and current terminology and definitions of this diverse range of potential therapeutics.

There are, however, several unanswered questions remaining. For example, there is uncertainty regarding the clinical evidence for the analgesic efficacy of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators, and it is not clear whether the current clinical evidence, based upon efficacy and safety considerations, justifies their use for pain management.<sup>21,46</sup>

These findings strengthen the rationale for assessing the full evidence base. Improving our understanding of the preclinical literature will better inform future clinical research. Animal models of injury-related and pathological persistent pain are used to investigate the underlying pathophysiology as well as to assess the efficacy and adverse effect profile of potential analgesics. Such studies provide justification and indications for clinical trials. The failure to translate findings from preclinical

**Table 1****Terminology and Definitions (Adapted from Soliman et al. (2019) after modification from Hauser et al. (2018)).**

Term	Definition	Examples/typical products
(Herbal) cannabis	The whole plant or parts or material from the plant (eg, flowers, buds, resin, and leaves)	<i>Cannabis sativa</i> , hashish
Medical or medicinal cannabis	The term “medical/medicinal cannabis” (or “medical/medicinal marijuana”) is used for cannabis plants, plant material, or full plant extracts used for medical purposes.	Bedrocan, Bedrobinol, Tilray 10THC/10CBD
Cannabis-based (or cannabis-derived) medicines	Medicinal cannabis extracts or products with regulatory approval for marketing as a therapeutic with defined and standardized THC and/or CBD content.	Nabiximols (Sativex), dronabinol, marinol, Epidiolex
Cannabinoids	Cannabinoids are biologically active constituents of cannabis, or synthetic compounds, usually having affinity for and activity at cannabinoid receptors.	THC, CBD, CP55,940, WIN55,212-2, HU210, nabilone
Phytocannabinoid	A cannabinoid found in cannabis plants or purified/extracted from plant material	THC, CBD
Endocannabinoid	An endogenous ligand found in the body of humans and other animals and which has affinity for, and activity at, cannabinoid receptors	Anandamide, 2-AG
Cannabinoid receptor antagonists and negative allosteric modulators	Directly block cannabinoid receptors or reduce signalling indirectly via impeding action of endogenous ligand through actions at a distinct site	Cannabinoid receptor antagonists (rimonabant [SR141716A], AM251, SR144528, AM630), negative allosteric modulators (PSNCBAM-1), DAGL inhibitors (RHC80267)
Modulators that increase or enhance endocannabinoid system activity	In addition to individual phytocannabinoids, cannabis-derived or cannabis-based medicines, and cannabis extracts, other pharmacological approaches under development for manipulation of the endocannabinoid system include selective synthetic cannabinoid receptor agonists, inhibitors of the catabolism (eg, FAAH inhibitors), transport (eg, FABP inhibitors) or reuptake of endocannabinoids, or positive allosteric modulators of cannabinoid receptor signalling.	FAAH inhibitors (PF-04457845, URB597, URB937), anandamide transport inhibitors (AM404, VDM11), MGL inhibitors (URB602, JZL184, MJN110), positive allosteric modulators of the CB <sub>1</sub> receptor (ZCZ011, GAT211)

CB1, cannabinoid type 1; CBD, cannabidiol; FAAH, fatty acid amide hydrolase; FABP, fatty acid-binding protein; THC,  $\Delta^9$ -tetrahydrocannabinol; 2-AG, 2-arachidonoyl glycerol; MGL, monoacylglycerol lipase.

research to clinical treatment has raised questions about the predictive validity and utility of animal models in drug development.<sup>30</sup> Limitations in experimental design,<sup>2,36</sup> conduct,<sup>10,37</sup> analysis, and reporting<sup>7,43</sup> may be compounding

the challenges of translational pain research and hindering the development of effective therapies.

As part of the International Association for the Study of Pain (IASP) Presidential Taskforce on Cannabis and Cannabinoid Analgesia, we

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

<sup>a</sup> Pain Research, Department of Surgery and Cancer, Imperial College London, London, United Kingdom, <sup>b</sup> Department of Anesthesiology, Washington University Pain Center, Washington University School of Medicine, St. Louis, Missouri, United States, <sup>c</sup> Department of Psychological and Brain Sciences, Program in Neuroscience and Gill Center for Biomolecular Science, Bloomington, IN, United States, <sup>d</sup> Departments of Anesthesiology, Perioperative, and Pain Medicine, & Pediatrics, Stanford University School of Medicine, Stanford, CA, United States, <sup>e</sup> CAMARADES, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom, <sup>f</sup> Department of Anesthesiology, Intensive Care and Pain Medicine University Hospital Muenster, Muenster, Germany, <sup>g</sup> EPPI-Centre, University College London, London, United Kingdom, <sup>h</sup> SYRCLE at Central Animal Laboratory, Radboud University Medical Center, Nijmegen, the Netherlands, <sup>i</sup> Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, <sup>j</sup> Department of Pharmacology and Toxicology, Faculty of Pharmacy, Assiut University, Assiut, Egypt, <sup>k</sup> Independent Researcher, <sup>l</sup> Department of Psychological Sciences, Tarleton State University, Stephenville, TX, United States, <sup>m</sup> Countess of Chester Hospital, Chester, United Kingdom, <sup>n</sup> Valdmann Institute of Pharmacology, Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia, <sup>o</sup> Pharmacology and Therapeutics, School of Medicine, Galway Neuroscience Centre and Centre for Pain Research, Human Biology Building, National University of Ireland Galway, Galway, Ireland, <sup>p</sup> Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, United States, <sup>q</sup> Leicester General Hospital, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom, <sup>r</sup> IIMPACT in Health, University of South Australia, Adelaide, South Australia, Australia, <sup>s</sup> Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, New South Wales, Australia, <sup>t</sup> Department of Anaesthesiology, Intensive Care Medicine and Pain Management, Medical Faculty of Ruhr-University Bochum, BG University Hospital Bergmannsheil gGmbH, Bochum, Germany, <sup>u</sup> Department of Allied Health Sciences, Institute of Health and Social Care, Pain Research Cluster, Ageing, Acute and Long Term Conditions Research Group, London South Bank University, London, United Kingdom, <sup>v</sup> Molecular Neuropharmacology and Genetics Laboratory, Department of Neuroscience, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, <sup>w</sup> Staten Island University Hospital Northwell Health, Staten Island, NY, United States, <sup>x</sup> Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States, <sup>y</sup> Visiting Scholar, Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA, United States, <sup>z</sup> Department of Psychological Sciences, Auburn University, Auburn, AL, United States, <sup>aa</sup> King's College London GKT School of Medical Education, King's College London, London, United Kingdom

\*Corresponding author. Address: Pain Research, Department of Surgery and Cancer, Imperial College London, London, SW10 9NH United Kingdom. Tel.: 0044 7793708361. E-mail address: n.soliman16@imperial.ac.uk (N. Soliman).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 162 (2021) S26–S44

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Association for the Study of Pain. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

<http://dx.doi.org/10.1097/j.pain.0000000000002269>

performed a preclinical systematic review and meta-analysis of the available evidence on cannabinoids, cannabis-based medicines, and endocannabinoid system modulators tested for antinociceptive effects in animal models of injury-related or pathological persistent pain. We have made the full data set available on the Open Science Framework for further investigation (<https://osf.io/2qde5/>).

## 2. Aims and objectives

The review was conducted using the CAMARADES Systematic Review Facility online platform (SyRF; [www.syrf.org.uk](http://www.syrf.org.uk)). A crowd was recruited to assist with the study selection, annotation, and data extraction stages of the review. In addition, machine learning was used to perform error analysis to ensure that all relevant studies were identified for inclusion. We aimed to (1) estimate the efficacy of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators in animal models of injury-related or pathological persistent pain, (2) assess the impact of the studies' internal and external validity on the reported behavioural outcome measures, and (3) identify the presence of publication bias and determine its magnitude. By exploring the reported quality and design characteristics of preclinical studies testing the efficacy of these drugs, we aimed to provide evidence and useful information for preclinical researchers wishing to increase scientific validity, improve the design of experiments, and refine the in vivo modelling of injury-related or pathological persistent pain.

## 3. Methods

The methods for the review were prespecified in the study protocol, registered on PROSPERO (CRD42019124804; [https://www.crd.york.ac.uk/prospero/display\\_record.php?ID=CRD42019124804](https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019124804)) and published.<sup>47</sup> We do not have any deviations from the protocol to report.

### 3.1. Crowd recruitment, training, and contribution

Ethical approval to use a crowd was obtained in March 2019 from Imperial College London's Head of Surgery and Cancer and Science, Engineering and Technology Research Ethics Committee. Crowd members were recruited by advertising through the IASP network, collaborators, and students using direct communication, newsletters, and social media. Volunteers were required to pass training modules (developed by NS hosted on the platform Learn to SyRF; [learn.syrf.org.uk](http://learn.syrf.org.uk)) for both screening and data extraction. To pass the screening training, participants had to correctly make the include or exclude decisions for 10 consecutive publications. For data extraction, they were presented with studies to extract data from, and were required to score greater than 80% (compared to "gold standard") for 5 successive studies. The pass thresholds were determined based upon agreement between expert reviewers in prior reviews. In accordance with ICMJE criteria, crowd members could join the authorship team upon completion of screening over 350 studies and extracting data from 35 studies (or 40 studies if they joined later in the review process). Those who did not meet these thresholds are acknowledged for their contribution.

### 3.2. Search

In April 2019, 3 online databases (PubMed, Web of Science, and Ovid Embase) were systematically searched with no language restrictions to identify publications reporting testing of a cannabinoid, cannabis-based medicine, or endocannabinoid system modulator for antinociception in an in vivo model of

persistent pain. The general search terms are given below; full search strategy can be found on the Open Science Framework Cannabinoid Preclinical SR search strategy ([https://osf.io/2qde5/?view\\_only=45aa94bb8ed64d4aabc21e79af5a0e72](https://osf.io/2qde5/?view_only=45aa94bb8ed64d4aabc21e79af5a0e72)):

Cannabinoids OR cannabis OR marijuana OR marihuana OR hemp OR hashish OR cannabinoid OR cannabinoids OR cannabidiol OR tetrahydrocannabinol OR "endocannabinoid modulator" OR FAAH OR monoacylglycerol lipase (MGL) OR MAGL OR ABHD6 OR ABHD12 OR "fatty acid binding protein" OR NAAA OR endocannabinoid OR endocannabinoids OR endocannabinoid OR FAAH inhibitor OR FAAH inhibition OR MAGL inhibitor OR MAGL inhibition OR MGL inhibitor OR MGL inhibition OR anandamide transport inhibitor OR anandamide transport inhibition OR "ABHD6 inhibitor" OR "ABHD6 inhibition" OR "ABHD12 inhibitor" OR "ABHD12 inhibition" OR NAAA inhibitor OR NAAA inhibition OR "Fatty acid Binding Protein inhibitor" OR "fatty acid binding protein inhibition" OR FABP inhibition OR FABP inhibitor OR allosteric modulator OR "endocannabinoid modulators" OR "endo-cannabinoid modulators" OR "endo-cannabinoid modulator" OR FAAH inhibitors OR MAGL inhibitors OR MGL inhibitors OR anandamide transport inhibitors OR "ABHD6 inhibitors" OR "ABHD12 inhibitors" OR NAAA inhibitors OR "Fatty acid Binding Protein inhibitors" OR FABP inhibitors OR allosteric modulators OR PEA OR palmitoylethanolamide AND Pain OR Hyperalgesia OR pain OR analgesia OR analgesic OR analgesics OR allodynia OR neuralgia OR hypersensitivity OR hyperalgesia OR hyperalgesic OR antinociception OR antinociception OR hypoalgesia OR hypoalgesic OR antihyperalgesia OR antihyperalgesia OR antihyperalgesic OR antihyperalgesic OR anti-allodynic OR antiallodynic OR anti-allodynia OR antiallodynia AND Animal search filters.

Search results were limited to animal studies using specific data-base search filters.<sup>11,27</sup> The search results were amalgamated into an Endnote (X7) library and duplicates removed.

### 3.3. Inclusion and exclusion criteria

Population: any injury-related or pathological persistent pain model. Persistent pain was described as typically studied over a period of hours, days, weeks, or months, and therefore for inclusion, a minimum experiment length of 1 hour. Intervention: any cannabinoid, cannabis-based medicine, or endocannabinoid system modulator administered to assess antinociceptive effect. Comparison: a separate cohort of animals in which the model was induced and was given a vehicle control treatment. Outcome: any pain-associated behavioural outcome measures.

For the meta-analysis, studies were required to report the number of animals per group, the mean, and a measure of variance (either the SEM or SD). Studies assessing the drug intervention in a model of acute pain were excluded (pathological or injury-related models persisting for less than 1 hour or naive/healthy animals used in pain-associated behavioural assessments). Similarly, studies that did not have an appropriate control were excluded. For example, the same animal could not be used for both, eg, contralateral is not suitable control for ipsilateral due to the possibility of contralateral sensory changes that could affect outcome measures.

### 3.4. Study selection

Using SyRF, the articles identified in the search were manually screened based on title and abstract by 2 independent reviewers, with discrepancies reconciled by a third. To ensure that the crowd had correctly identified and included all relevant studies, the human decisions were used to train a machine learning algorithm.<sup>3</sup> Error

analysis was conducted in an iterative manner, presenting the top 200 studies for possible inclusion for screening. These were screened by an expert reviewer (N.S.) until there was not a change in the decisions for either inclusion or exclusion.

### 3.5. Risk of bias

In accordance with the CAMARADES checklist<sup>33</sup> and adaptation of the SYRCLE Risk of Bias tool,<sup>26</sup> the risk of bias of the included studies was assessed by recording the reporting of 5 methodological quality criteria: sample size calculation, randomisation, allocation concealment, blinded assessment of outcome, and reporting of animal exclusions. Criteria were extracted by 2 independent reviewers and discrepancies reconciled by a third. Risk was scored high, low, or unclear based upon the reporting of the method. Reporting of potential conflicts of interest and compliance with animal welfare regulations<sup>32,33</sup> were collected but were not included in the overall risk of bias.

### 3.6. Data extraction

Data were extracted into SyRF. For all included studies, details of publication, model, intervention, outcome assessment (Table 2), and other experiment details were extracted (Table 3). Outcome data presented graphically were extracted using digital ruler software (Universal Desktop Ruler, Adobe ruler, Webplotdigitizer) to determine values. When multiple time points were presented, the time point that showed the greatest difference between the control group and treatment group was extracted. If the type of variance (eg, SEM or SD) was not reported, it was characterised as SEM because this is the more conservative approach, as studies are weighted in part by the inverse of the observed variance. For each study, data were extracted by 2 independent reviewers.

### 3.7. Data reconciliation

Data extracted by 2 independent reviewers were compared, and any discrepancies reconciled by a third independent reviewer. For outcome data, which were predominantly reported in graphs, the standardised mean difference (SMD) effect size of individual comparison was calculated for each reviewer's extracted data; if these differed by >10%, they required reconciliation. When they differed by <10%, a mean of the 2 means and variances was calculated.

### 3.8. Data analysis

The meta-analysis was conducted in accordance with the guidelines described by Vesterinen et al.<sup>51</sup> The data have been analysed as a whole and subgroup analyses were conducted to investigate how effect sizes vary according to study characteristics, eg, species, strain, sex, the type of injury/pathology modelled, and the outcome measure. Data pertaining to the cannabinoid-related intervention were extracted including the time point of admission relative to model induction, either prophylactically (pre-model induction) or therapeutically (post-model induction).

A Hedges' g SMD effect size was calculated for each comparison. Effect sizes were weighted using the inverse variance method to reflect the contribution of each comparison to the total effect estimate. When a single control group served multiple treatment groups, the control group sample size was adjusted by dividing the number of animals in the control group by the number of treatment groups served to avoid artificial inflation of *n*. When more than one pain-associated behavioural outcome was reported for the same cohort of

**Table 2**

**Study-level data extracted from each included publication.**

Meta-data	Risk of bias	Reporting quality	Curated content
First author	Sample size calculation	Compliance with animal welfare regulations	Locomotor assessment
Year	Randomisation	Statement of potential conflict(s) of interest	Confirmation of drug target
	Allocation concealment		Electrophysiology
	Blinded assessment of outcome		Markers of neuronal activity
	Animal exclusions		Assessment of depression/anxiety-related behaviour

animals, the comparisons were combined to provide a single nested comparison that is a summary effect for each cohort. Cohort-level effect sizes were then pooled using a random-effects model (adjusted using the Hartung-Knapp-Sidik-Jonkman method<sup>24,25</sup>) and the restricted maximum-likelihood model was used to estimate heterogeneity, the variation of outcomes across studies.<sup>50</sup>

Subgroup analyses (stratified meta-analyses for categorical variables) were performed to investigate how study characteristics influence the overall estimates of effects. The subgroup analyses aim to provide empirical evidence to inform experimental design and refine modelling of injury-related or pathological persistent pain and the extent to which predefined study design and study risk of bias characteristics differ in their overall estimates of effect. The study design factors analysed using stratified meta-analysis were animal species, strain and sex, model type, outcome measure type, therapeutic intervention, and intervention type. Where possible, drugs have been classed by mechanism of action as listed in the IUPHAR/BPS Guide to Pharmacology.<sup>1</sup> Several drugs have multiple potential sites of action and these have been classified accordingly. Characteristics of the intervention, eg, dose and route of delivery are important but inextricably linked to the intervention and, therefore, were not analysed independently. We also assessed the impact of the reporting of methodological quality criteria and compared the

**Table 3**

**Experimental-level data extracted from each included publication.**

Animal	Model	Intervention	Outcome measure
Species	Type of model	Drug	Outcome measure type
Strain		Time of administration in relation to model induction	Units
Sex	Method of induction		Direction of effect
		Dose	
Age		Route of administration	Number of treatment groups served by control
Weight			Time of assessment
			For each group;
			Sample size
			Mean outcome
			Variance



pooled effect size of studies that did report a specific criterion with the pooled effect of studies that did not.

The analyses were conducted using R version 3.6.2. and the packages meta (version 4.15.1), metafor (version 2.4.0),<sup>52</sup> and dmetar (version 0.0.9000).<sup>23</sup>

### 3.9. Publication bias

Funnel plots were generated by plotting each study's effect size on the x-axis against its sample size-based precision estimate  $1/\sqrt{N}$  on the y-axis, in accordance with guidance by Zwetsloot et al.<sup>57</sup> The potential for publication bias was assessed by visual inspection of the asymmetry of funnel plots. Trim and fill analysis attempted to correct for funnel plot asymmetry by imputing the theoretically missing studies on the left-hand side of the plot and enabling a recalculation of the overall effect size.<sup>14</sup> In addition, Egger regression allowed for a statistical assessment of the presence of publication bias.<sup>16</sup>

## 4. Results

### 4.1. Crowd recruitment, training, and contribution

The recruitment strategy aimed to target bioscientists but there were not any prerequisite criteria. Volunteers had varied knowledge of the topic and experience of the systematic review process. Four hundred fifty-three people from 44 countries signed up to the project. Of those, 100 went on to make a meaningful contribution to the project, with 28 making a large enough contribution to meet authorship criteria. The crowd took 6 weeks to complete the screening phase and 37 weeks to complete data extraction.

### 4.2. Systematic search, study selection, and error analysis

The systematic search identified 10,816 articles for screening against the inclusion/exclusion criteria (search results are available on the Open Science Framework [https://osf.io/2qde5/?view\\_only=45aa94bb8ed64d4aabc21e79af5a0e72](https://osf.io/2qde5/?view_only=45aa94bb8ed64d4aabc21e79af5a0e72)). A total of 850 studies were initially identified for inclusion. Error analysis of the included studies was performed manually and 228 (26.8% of the included studies) had been wrongly included. The incorrectly included studies all tested an intervention for physiological antinociception in healthy animals, not in a model of pathological or injury-related persistent pain. All human decisions were then used to train the machine learning algorithm. The machine ranked and presented the studies where there was a difference between human and machine decisions. In the first iteration, the machine performed with 95% sensitivity and 89% specificity, which improved with each iteration to an eventual sensitivity of 95% and specificity of 94%. During the error analysis process, a further 1200 studies were presented for screening. A total of 137 decisions changed, leading to an eventual inclusion of 751 studies: 129 wrongly excluded were included and 8 wrongly included were excluded.

A further 278 articles were excluded at full text screening, conducted concurrently to the annotation and data extraction stages, leading to an inclusion of 473 studies (Fig. 1). 99 studies were missing key information for meta-analysis. Data extracted from the 374 studies qualifying for inclusion and meta-analysis are presented here.

### 4.3 Study characteristics

In the 374 studies included in the meta-analysis, 171 interventions were assessed for antinociceptive effect in models of pathological or injury-related persistent pain (see Table, supplemental digital content 1, for included study list and study

characteristics; available at <http://links.lww.com/PAIN/B332>). The drugs are listed by mechanism of action (23 drug classes) in Table 4 (see Table, supplemental digital content 2, for full list; available at <http://links.lww.com/PAIN/B333>). Cannabinoid type 1 and CB<sub>2</sub> receptor agonists were tested most frequently, 281 (19%) and 299 (20%) comparisons, respectively. The most frequently tested drug was the CB<sub>1</sub> receptor agonist, WIN55,212-2 ( $n = 194$  comparisons, 13%). 20 model types were used (Table 5); inflammation, nerve injury, and formalin were used most frequently: 30% ( $n = 467$  comparisons), 27% ( $n = 413$  comparisons), 15% ( $n = 235$  comparisons), respectively. Over 16,000 animals were included in the analysis with a median of 24 animals' pain-associated behavioural outcome data extracted from each included study (a range of 4–162).

All experiments were conducted in rodents. Fifty-six percent ( $n = 865$  comparisons) were conducted in rats and 44% ( $n = 678$  comparisons) were conducted in mice. Male animals were used in 86% ( $n = 1334$  comparisons) and female animals were used in 7% ( $n = 110$  comparisons). Two percent ( $n = 28$ ) used mixed sex groups and 5% ( $n = 74$  comparisons) did not report the sex of the animals used. Evoked limb withdrawal to mechanical and thermal stimuli were the most frequently used pain-associated behavioural outcome measures; 51% ( $n = 791$  comparisons) and 22% ( $n = 343$  comparisons), respectively.

To inform a narrative review<sup>19</sup> and future research, the following was annotated for each study and is summarised in Table 6: Whether the study investigated the effects of drug on non-pain-related motor activity, the pharmacokinetic properties, tissue concentrations, and where applicable confirmed the cannabinoid receptors as the target. In addition, we assessed whether the study investigated potential toxic effects, effects on dependency, and on aspects of animal behaviour potentially reflecting anxiety- and depression-related behaviour. We also assessed whether electrophysiology (eg, wide dynamic range and nociceptive specific cells) or markers for neuronal activity (eg, c-fos, Fos, ERK, and p38 MAPK) were measured in studies of antinociceptive efficacy.

### 4.4. Meta-analysis of the antinociceptive efficacy of treatment with a cannabinoid, cannabis-based medicine, or endocannabinoid system modulator

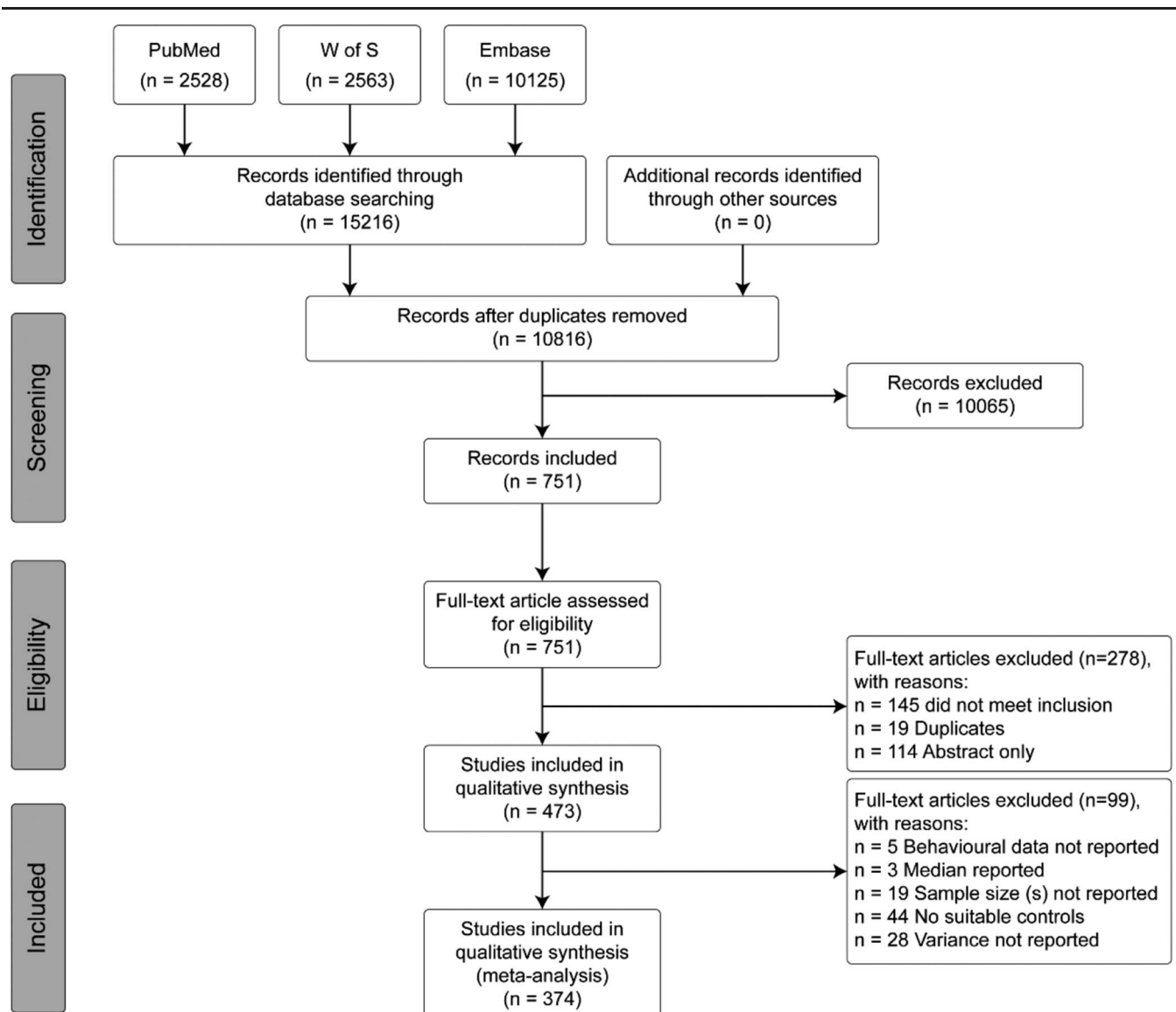
A total of 374 studies, comprising 1544 comparisons, investigated the effects of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators in models of pathological or injury-related persistent pain. Prophylactic and/or therapeutic administration of the drugs led to a significant attenuation of pain-associated behaviour compared with control {SMD = 1.321 (95% confidence interval [CI] 1.232–1.411)}. Heterogeneity was moderate ( $I^2 = 61.58\%$ ) (Fig. 2).

#### 4.4.1. Effects of study characteristics on antinociceptive efficacy

Subgroup analyses demonstrated that species accounted for a significant proportion of heterogeneity ( $Q = 17$ ,  $df\ 2$ ,  $P < 0.005$ ). Therefore, rats and mice have been analysed separately. Further subgroup analyses were conducted to determine whether the antinociceptive effect varies due to study design characteristics.

##### 4.4.1.1. Rats

Rats were used in 276 studies ( $n = 6479$ , 864 comparisons) to assess the potential antinociceptive effect of the treatments. 95



**Figure 1.** A flow diagram of articles identified from the bibliographic search of 3 electronic databases: PubMed, Web of Science (W of S), and Embase, conducted on April 9, 2019. The diagram provides the breakdown of records through deduplication, screening, and eligibility until final inclusion in both qualitative and quantitative analysis. Reported in accordance with the PRISMA guidelines. W of S, Web of Science.

interventions were tested in 11 model types. The treatments led to a significant attenuation of pain-associated behaviour compared with control (SMD = 1.306 [95% CI 1.199-1.412]). Heterogeneity was moderate ( $I^2 = 57.8\%$ ,  $Q = 2044.69$ ,  $df\ 863$ ,  $P < 0.0001$ ).

The drug and drug class accounted for a significant proportion of heterogeneity ( $Q = 1338.17$ ,  $df\ 94$ ,  $P < 0.0001$ ; and  $Q = 36.37$   $df\ 20$ ,  $P < 0.05$ , respectively). Novel compounds not included in the classification were classified based upon the mechanism of action reported by the study authors. Cannabinoid type 1 receptor agonists and CB<sub>2</sub> receptor agonists were most frequently assessed (194 and 159 comparisons, respectively). Most drug classes resulted in a significant antinociceptive effect; NAAA inhibitors produced the largest significant attenuation of pain-associated behaviour compared with control (SMD = 1.59 [95% CI 1.17-2.01]), whereas peroxisome proliferator-activated receptor (PPAR)-gamma antagonist, GPR55 agonist, hemp oil, FABP inhibitors, and CB<sub>2</sub> receptor inverse agonist did not have significant effect; however, these were of single studies or single comparisons. Although their effect is smaller, the CB<sub>1</sub> and CB<sub>2</sub> receptor agonists subgroups are comprised of data

from a high number of animals and comparisons therefore we can have greater confidence in the stability of the effect. The smallest significant effect was elicited by cannabidiol (CBD) (SMD = 1.12 [95% CI 0.84-1.40]). Most drugs were assessed after model induction (702 comparisons). Whether the drug was administered prophylactically, or therapeutically, did not account for a significant amount of heterogeneity ( $Q = 0.30$ ,  $df\ 2$ ,  $P = 0.9$ ) (Fig. 3A).

#### 4.4.1.2. Mice

Mice were used in 153 studies ( $n = 6876$ , 677 comparisons) to assess the potential antinociceptive effect of the treatments. One hundred ten drugs were tested in 15 model types. Overall, the treatments led to a statistically significant attenuation of pain-associated behaviour compared with control (SMD = 1.353 [95% CI 1.199-1.506],  $P < 0.0001$ ). Heterogeneity was moderate ( $I^2 = 66.7\%$ ,  $Q = 2027.72$ ,  $df\ 676$ ,  $P < 0.0001$ ).

The drug and drug class accounted for a significant proportion of heterogeneity ( $Q = 5332.76$ ,  $df\ 109$ ,  $P < 0.0001$ ; and  $Q = 82.26$ ,  $df$

**Table 4**

**Summary of the drug classes assessed for antinociceptive effect in animal models of injury-related or pathological persistent pain.**

Drug class	No. of studies	No. of nested comparisons
CB <sub>2</sub> receptor agonist	75	299
CB <sub>1</sub> receptor agonist	88	281
Nonselective cannabinoid receptor agonist	71	230
FAAH inhibitor	57	217
PPAR-alpha agonist	40	121
THC	16	69
Anandamide transport inhibitor	18	64
CBD	17	63
Monoacylglycerol lipase inhibitor	23	58
FABP inhibitor	3	31
Unknown mechanism of action	6	25
NAAA inhibitor	4	20
CB <sub>1</sub> receptor inverse agonist	7	19
Diacylglycerol lipase inhibitor	3	14
Dual FAAH/MGL inhibitor	4	10
CB <sub>1</sub> receptor PAM	1	5
FAAH inhibitor/TRPV1 agonist	1	5
CB <sub>2</sub> receptor inverse agonist	2	4
ABHD6 inhibitor	1	3
FAAH inhibitor/TRPA1 agonist	1	2
PPAR-gamma antagonist	1	2
GPR55 agonist	1	1
Hemp oil	1	1

ABHD6, abhydrolase domain containing 6; CB<sub>1</sub>, cannabinoid type 1; CB<sub>2</sub>, cannabinoid type 2; CBD, cannabidiol; FABP, fatty acid-binding protein; FAAH, fatty acid amide hydrolase; MGL, monoacylglycerol lipase; NAAA, N-acyl ethanolamine-hydrolysing acid amidase; PPAR, peroxisome proliferator-activated receptor; PAM, positive allosteric modulator; TRPV1, transient receptor potential vanilloid receptor 1; TRPA1, transient receptor potential ankyrin 1; THC, delta-9 tetrahydrocannabinol.

23,  $P < 0.0001$ , respectively). Cannabinoid type 2 receptor agonists and FAAH inhibitors were assessed the most (137 and 103 comparisons, respectively). As in rats, most drug classes produced a significant antinociceptive effect. The largest significant attenuation of pain-associated behaviour compared with control was reported for NAAA inhibitors (SMD = 3.23 [95% CI 1.97–4.50]); however, we can have more confidence in the smaller effect sizes of the CB<sub>2</sub> receptor agonists and FAAH inhibitors. In addition, an FAAH inhibitor/TRPV1 agonist and ABHD6 inhibitors did not significantly attenuate pain-associated behaviours.

Most drugs were assessed after model induction (475 comparisons) and whether they were administered pre- or post-model induction accounted for a significant amount of heterogeneity (SMD = 1.173 [95% CI 0.921–1.434] and SMD = 1.415 [95% CI 1.227–1.602]  $Q = 15.07$ ,  $df 2$ ,  $P = 0.005$ ) (**Fig. 3b**).

#### 4.5. Interpreting effect sizes

Effect sizes are influenced by 2 factors: the mean difference between groups and the variance within the groups. This is of importance for biomedical research as preclinical studies typically lead to larger

**Table 5**

**Summary of the model types used to assess the antinociceptive effect of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators.**

Model type	No. of studies	No. of nested comparisons
Inflammation	434	467
Nerve injury	348	413
Formalin	223	235
Chemotherapy	112	128
Diabetes	63	74
Cancer	57	65
Postoperative	27	52
Visceral inflammation	20	31
Chemical cauterization	1	16
Migraine	9	13
HIV	4	11
Capsaicin	5	9
Heat injury	2	7
Multiple sclerosis	6	7
Musculoskeletal	2	4
Antiretroviral	1	3
Burn injury	1	3
Mustard oil	3	3
Sickle cell disease	2	2
Mild traumatic brain injury	1	1

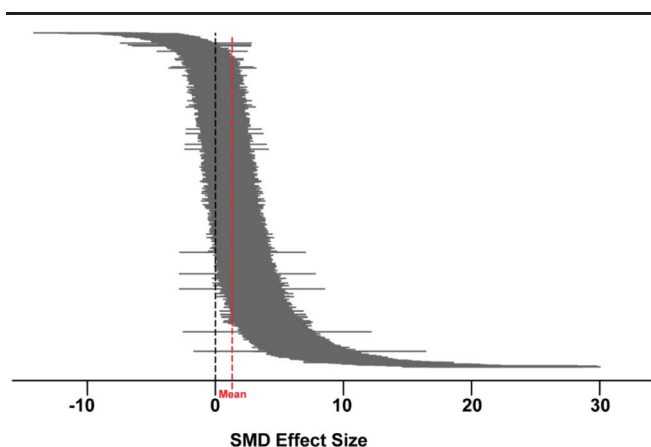
effect sizes than clinical studies due to the homogenous and controlled nature of the experiments limiting the observed variance (ie, the larger effect size is often due to smaller variance and not larger mean differences). To assist interpretation, the overall SMD effect of 1.321 suggests that 90.7% of the treatment group will have a mean larger than the mean of the control group, with an overlap between the 2 groups of 50.9% (**Fig. 4**). This concept is further illustrated for the drug classes assessed for antinociceptive effect in both rats and mice in **Table 7**.

**Table 6**

**The number of studies that conducted further experimentation to gain further understanding of the cannabinoids, cannabis-based medicines, and endocannabinoid system modulators in conjunction with antinociceptive effect.**

	No. of studies	%
Confirmation (where applicable) of the CB <sub>1</sub> /CB <sub>2</sub> target	207	69
Effects on motor activity	124	33
Investigate pharmacokinetics	26	7
Tissue concentrations	25	7
Electrophysiology	20	5
Potential toxic effects	19	5
Measure markers of neuronal activity	19	5
Effect on anxiety/depression	11	3
Effects on dependency	10	3

CB<sub>2</sub>, cannabinoid type 2.



**Figure 2.** A caterpillar plot of the 1544 nested comparisons extracted from the 374 studies included in the meta-analysis. Hedges' *g* standardised mean differences (SMD) were calculated for each comparison. Effect sizes were pooled using the random-effects model and heterogeneity estimated with the restricted maximum-likelihood model (red dashed line indicates overall mean). Overall effect size = 1.321.  $Q = 4101.26$ ,  $df\ 1543$ ,  $P < 0.0001$ ,  $I^2 = 61.58\%$ .

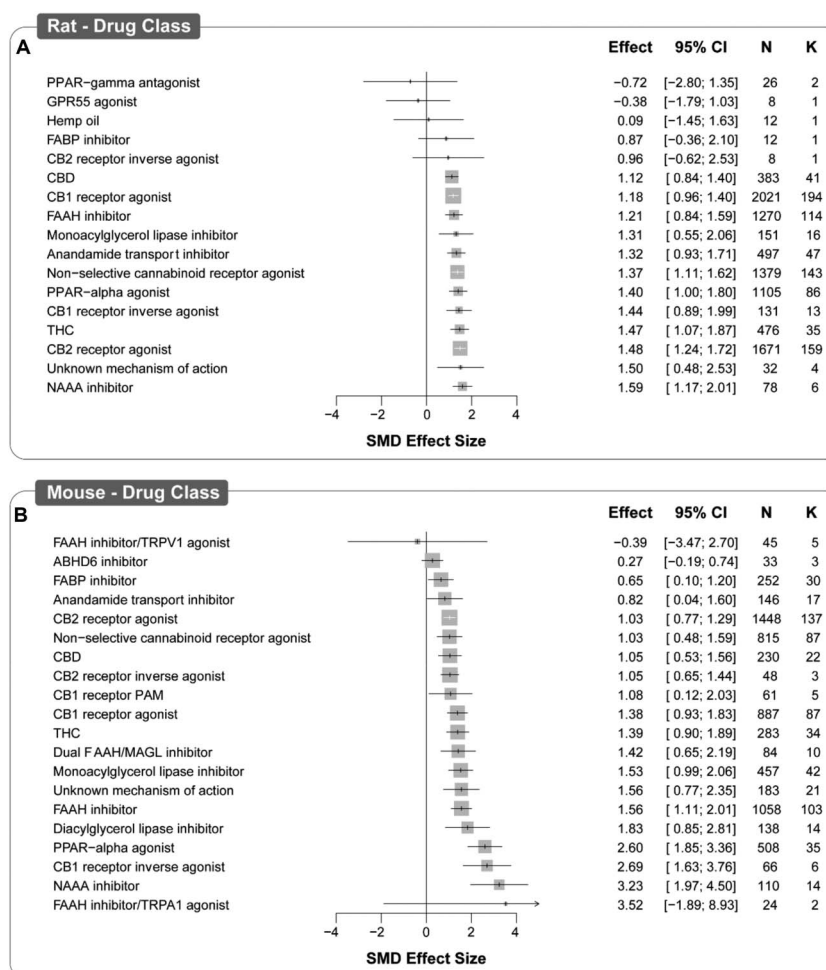
#### 4.6. Animal model characteristics

Stratified meta-analyses were conducted to determine the influence of animal model characteristics on the observed effect sizes; the forest

plots pertaining to rat characteristics are presented in **Figure 5**: model (A), strain (B), sex (C), and outcome measure (D). Similarly, mouse characteristics are presented in **Figure 6**: model (A), strain (B), sex (C), and outcome measure (D). The results are presented thematically below and with reference to both **Figures 5 and 6**.

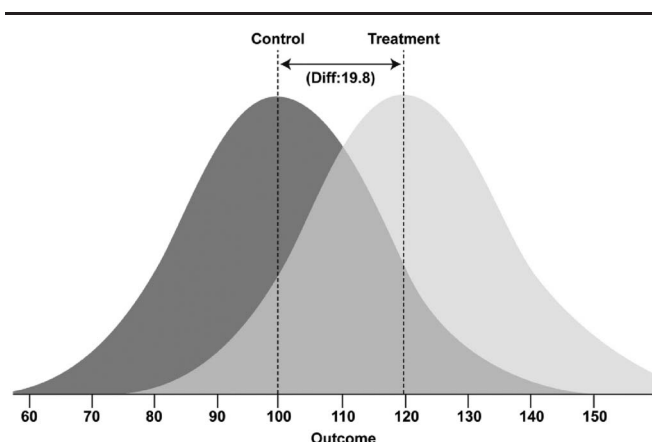
##### 4.6.1. Model type

In both rats and mice, the model type accounted for a significant proportion of heterogeneity ( $Q = 36.56$ ,  $df\ 13$ ,  $P < 0.0025$ ; and  $Q = 173.03$ ,  $df\ 14$ ,  $P < 0.0001$ , respectively). Of the 15 model types, inflammation followed by nerve injury were modelled most frequently in both rats (269 and 259 comparisons, respectively) and mice (195 and 154 comparisons, respectively). In rats, the largest attenuation of pain-associated behaviour compared with control was reported in models of burn injury (SMD = 2.23 [95% CI 0.33-4.14]) and the smallest significant attenuation was reported in models of inflammation (SMD = 1.16 [95% CI 0.93-1.38]). The overall estimate of effect was significant for most rat model types except heat injury, migraine, and capsaicin models (**Fig. 5A**). In mice, the largest significant attenuation of pain-associated behaviour compared with control was reported in the mouse model of HIV protein-associated neuropathy (SMD = 7.932 [95% CI 5.115-10.748]), whereas the smallest significant attenuation was reported in nerve injury models (SMD = 1.04



**Figure 3.** Forest plot of drug classes assessed in rat (A) and mouse (B) models of injury-related or pathological persistent pain. The size of the square represents the weight (%). The weight is the influence that individual subgroup has on the pooled result. N denotes the number of animals that contribute to that subgroup. K denotes the number of comparisons that comprise each subgroup SMD, standardised mean difference..





**Figure 4.** Visualisation of the overlap between control and treatment group distributions of the overall SMD effect size of 1.32.<sup>34</sup> The darker distribution curve represents the control group and the lighter distribution curve represents the treatment group. Animals within each group can fall anywhere within their respective curves, with increasing likelihood towards the peak; imagine each curve a hill of animals with single animals at the tail-ends of the distribution curve. SMD, standardised mean difference.

[95% CI 0.71–1.36]). The estimate of effect was significant in most mouse models, except sickle cell disease, visceral inflammation, multiple sclerosis, and mustard oil (**Fig. 6A**).

#### 4.6.2. Strain

Strain accounted for a significant proportion of heterogeneity; 12 different strains of rats ( $Q = 39.41$ ,  $df\ 11$ ,  $P < 0.0001$ ) and 29 different strains of mice were reported ( $Q = 281.08$ ,  $df\ 28$ ,  $P < 0.0001$ ). In rats, Sprague-Dawley and Wistar strains were reported the most ( $n = 5148$ , 458 comparisons and  $n = 3655$ , 361 comparisons, respectively). All report significant effects except Wistar-Kyoto, Fischer 344, Wistar Han, and Wistar albino, and the obese diabetic ZDF/*cr1-lepr/fa*, as well as those strains that were not reported (**Fig. 5b**). In mice, the C57BL/6J strain was the most frequently reported ( $n = 1794$ , 182 comparisons). The effects of the drugs were significant in over half of the strains but insignificant in 13 strains (**Fig. 6B**).

#### 4.6.3. Sex

In rats, sex accounted for a significant proportion of heterogeneity ( $Q = 21.65$ ,  $df\ 3$ ,  $P < 0.001$ ) (**Fig. 5C**), whereas mouse sex did not ( $Q = 7.39$ ,  $df\ 3$ ,  $P = 0.06$ ) (**Fig. 6C**). Most of the data are from male animals. Female-only animal groups were used in 23 studies (6%,  $n = 1307$ ), mixed sex groups in 8 studies (2%,  $n = 86$ ) compared with 302 studies using male groups. In rats, 91% of the experiments ( $n = 8337$ , 786 comparisons) and 3% ( $n = 650$ , 48 comparisons) of the experiments were conducted using male and female rats, respectively. The sex was not reported for 28 comparisons ( $n = 213$ ), and 4 comparisons ( $n = 60$ ) used mixed sex groups. In mice, 81% of the experiments used male animals ( $n = 5574$ , 678 comparisons) and 9% used female animals ( $n = 657$ , 66 comparisons). The sex was not reported for 43 comparisons ( $n = 438$ ), and 24 comparisons ( $n = 207$ ) used mixed sex groups.

#### 4.6.4. Pain-associated behavioural outcome measures

In both rats (**Fig. 5D**) and mice (**Fig. 6D**), the type of pain-associated outcome measure accounted for a significant proportion of heterogeneity (rat:  $Q = 160.28$ ,  $df\ 5$ ,  $P < 0.0001$ ; and mouse;  $Q = 20.96$ ,  $df\ 5$ ,  $P < 0.001$ , respectively). Evoked limb

withdrawal to mechanical stimulation was most frequently reported (451 and 449 comparisons rats and mice, respectively). Significant effects were observed for all outcome assessment types for rats; however, in mice, effect sizes for complex and nonevoked assessment types were not significant.

#### 4.6.5. Antinociceptive efficacy of drug classes in different model types

In rats, CB<sub>1</sub> and CB<sub>2</sub> receptor agonists were assessed the most, with in 10 and 9 model types, respectively. The antinociceptive effect of the 2 drug classes was significant in most model types (see Table, supplemental digital content 3 for drug class effects in individual model types, available at <http://links.lww.com/PAIN/B334>): nerve injury, chemotherapy-induced peripheral neuropathy, and cancer models. However, CB<sub>1</sub> receptor agonists did not significantly attenuate pain-associated behaviour in inflammation and heat injury models (31 and 7 comparisons, respectively). Within rat inflammation models, the results are mixed; the CB<sub>1</sub> receptor agonists did not significantly attenuate pain-associated behaviours in models of carrageenan (10 comparisons) and osteoarthritis (1 comparison). Cannabinoid type 2 receptor agonists did not significantly attenuate pain-associated behaviour in the rat formalin and migraine models (4 comparisons in each). Fatty acid amide hydrolase inhibitors and peroxisome proliferator-activated receptor- $\alpha$  agonists (exclusively PEA, oleoyl-ethanolamide, and analogues of PEA) were also assessed in a broad range of model types, 6 and 10 model types, respectively. Peroxisome proliferator-activated receptor- $\alpha$  agonists significantly attenuated pain-associated behaviours across all model types, whereas FAAH inhibitors demonstrated antinociceptive effects in neuropathic pain-associated models (eg, nerve injury, CIPN, and diabetes) but a mixed effect was observed in inflammation-associated models. FAAH inhibitors (comprising 26 individual drugs) significantly attenuated pain-associated behaviours in the complete Freund's adjuvant and formalin models but not in carrageenan (11 comparisons) or osteoarthritis (12 comparisons) models (see Table, Supplemental Digital Content 4, for drug class effects in rat inflammation models available at <http://links.lww.com/PAIN/B335>).

In mice, CB<sub>2</sub> receptor agonists and FAAH inhibitors were assessed the most, across 10 and 8 model types, respectively (see Tables, supplemental digital content 5 and 6 for full drug class effects in mouse models, available at <http://links.lww.com/PAIN/B336> and <http://links.lww.com/PAIN/B337>, respectively). Cannabinoid type 2 receptor agonists significantly attenuated pain-associated behaviours in mouse cancer and visceral inflammation models but not in nerve injury (33 comparisons), multiple sclerosis (6 comparisons), and CIPN (2 comparisons) models. Fatty acid amide hydrolase inhibitors significantly attenuated pain-associated behaviours in inflammation (38 comparisons), nerve injury (29 comparisons), formalin (15 comparisons), postoperative (4 comparisons), and diabetes (3 comparisons) models but not in CIPN (8 comparisons), visceral inflammation (5 comparisons), and mustard oil (1 comparison) models. Cannabinoid type 1 and CB<sub>2</sub> receptor agonists were both assessed in 6 model types and significantly attenuated pain-associated behaviours in all model types except nerve injury. Like rats, PPAR- $\alpha$  agonists significantly attenuated pain-associated behaviours in the 4 model types in which they were assessed: nerve injury, inflammation, formalin, and CIPN models.

#### 4.7. Risk of Bias

The overall risk of bias of the 374 included studies is unclear. The reporting of methodological quality criteria was low: 47% (175)

**Table 7**

**Differences and overlap between treatment and control groups for drug classes assessed for antinociceptive efficacy in rats (A) and mice (B), which correspond to the forest plots of Figures 3 and 4.**

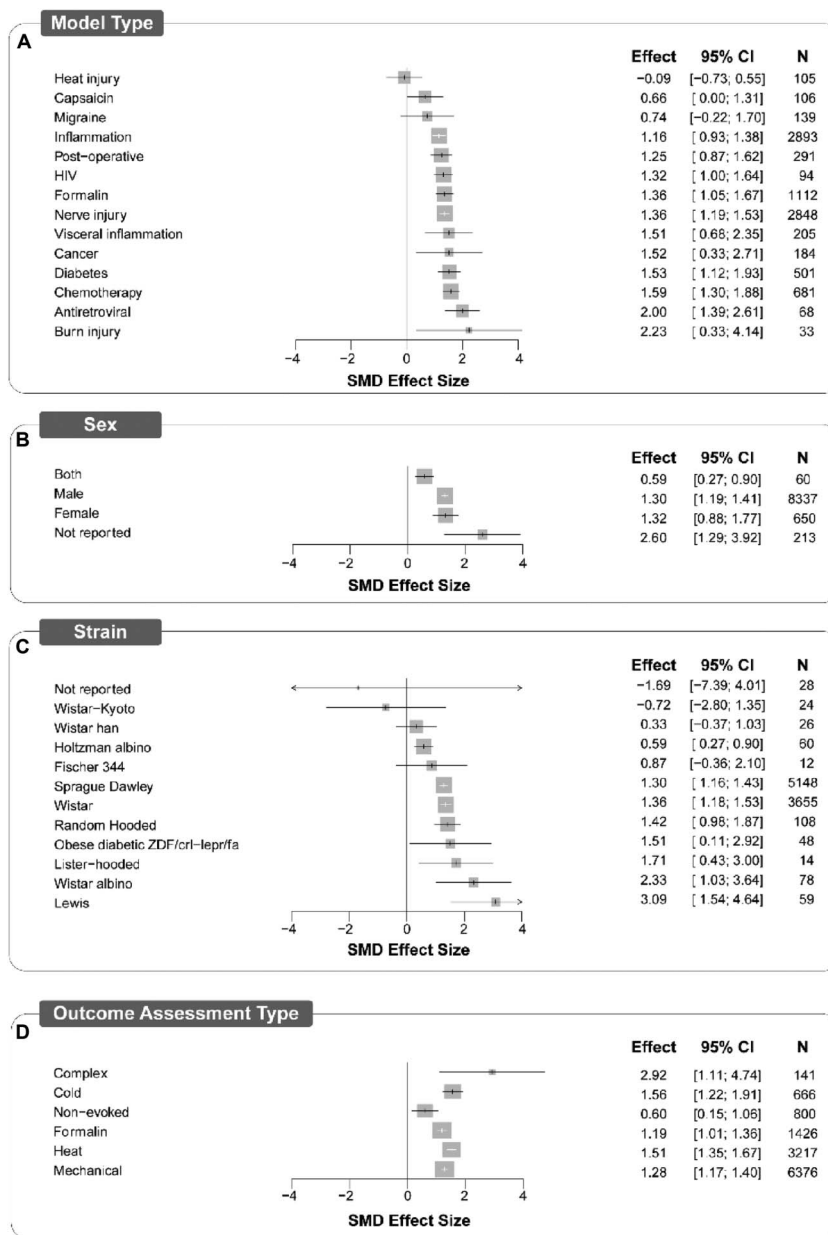
<b>A</b>	<b>Effect Size</b>	<b>% of treatment group with larger mean than control group</b>	<b>% overlap</b>	<b>N</b>
<b>Drug Class</b>				
NAAA inhibitor	1.59	94.4	42.7	78
Unknown mechanism of action	1.50	93.3	45.3	32
CB <sub>2</sub> receptor agonist	1.48	93.1	45.9	1671
THC	1.47	92.9	46.2	476
CB <sub>1</sub> receptor inverse agonist	1.44	92.5	47.2	131
PPAR-alpha agonist	1.40	91.9	48.4	1105
Nonselective cannabinoid receptor agonist	1.37	91.5	49.3	1379
Anandamide transport inhibitor	1.32	90.7	50.9	497
Monoacylglycerol lipase inhibitor	1.31	90.5	51.2	151
FAAH inhibitor	1.21	88.7	54.5	1270
CB <sub>1</sub> receptor agonist	1.18	88.1	55.5	2021
CBD	1.12	86.9	57.5	383
CB <sub>2</sub> receptor inverse agonist	0.96	83.1	63.1	8
FABP inhibitor	0.87	80.8	66.4	12
Hemp oil	0.09	53.6	96.4	12
GPR55 agonist	−0.38	35.2	84.9	8
PPAR-gamma antagonist	−0.72	23.6	71.9	26
<b>B</b>	<b>Effect Size</b>	<b>% of treatment group with larger mean than control group</b>	<b>% overlap</b>	<b>N</b>
<b>Drug Class</b>				
FAAH inhibitor/TRPA1 agonist	3.52	100	7.84	24
NAAA inhibitor	3.23	99.9	10.6	110
CB <sub>1</sub> receptor inverse agonist	2.69	99.6	17.9	66
PPAR-alpha agonist	2.60	99.5	19.4	508
Diacylglycerol lipase inhibitor	1.83	96.6	36	138
FAAH inhibitor	1.56	94.1	43.5	1058
Unknown mechanism of action	1.56	94.1	43.5	183
Monoacylglycerol lipase inhibitor	1.53	93.7	44.4	457
Dual FAAH/MGL inhibitor	1.42	92.2	47.8	84
THC	1.39	91.8	48.7	283
CB <sub>1</sub> receptor agonist	1.38	91.6	49	887
CB <sub>1</sub> receptor PAM	1.08	86	58.9	61
CB <sub>2</sub> receptor inverse agonist	1.05	85.3	60	48
CBD	1.05	85.3	60	230
Nonselective cannabinoid receptor agonist	1.03	84.8	60.7	815
CB <sub>2</sub> receptor agonist	1.03	84.8	60.7	1448
Anandamide transport inhibitor	0.82	79.4	68.2	146
FABP inhibitor	0.65	74.2	74.5	252
ABHD6 inhibitor	0.27	60.6	89.3	33
FAAH inhibitor/TRPV1 agonist	−0.39	34.8	84.5	45

N, number of animals (Magnusson, 2020).

ABHD6, abhydrolase domain containing 6; CB<sub>2</sub>, cannabinoid type 2; CBD, cannabidiol; FAAH, fatty acid amide hydrolase; FABP, fatty acid-binding protein; MGL, monoacylglycerol lipase; NAAA, N-acyl ethanolamine-hydrolysing acid amidase; PAM, positive allosteric modulator; PPAR, peroxisome proliferator-activated receptor; TRPV1, transient receptor potential vanilloid receptor 1; THC, delta-9-tetrahydrocannabinol.

reported blinded assessment of outcome, 32% (118) reported randomisation (to treatment or control group), 14% (54) reported animal exclusions, 13% (49) reported predetermined animal

exclusion criteria, 4% (15) reported allocation concealment, and 3% (12) reported a sample size calculation (**Fig. 7A**). This contrasts with the reporting of conflict of interest, 54% (203), and



**Figure 5.** Forest plots of study design characteristics of experiments in which treatments were assessed for antinociceptive efficacy in rat models of persistent or injury-related persistent pain. Model type (A), sex (B), strain (C), and outcome assessment type (D) account for a significant proportion of heterogeneity. The size of the square represents the weight (%). The weight is the influence that individual subgroup has on the pooled result. N denotes the number of animals that contribute to that subgroup. SMD, standardised mean difference

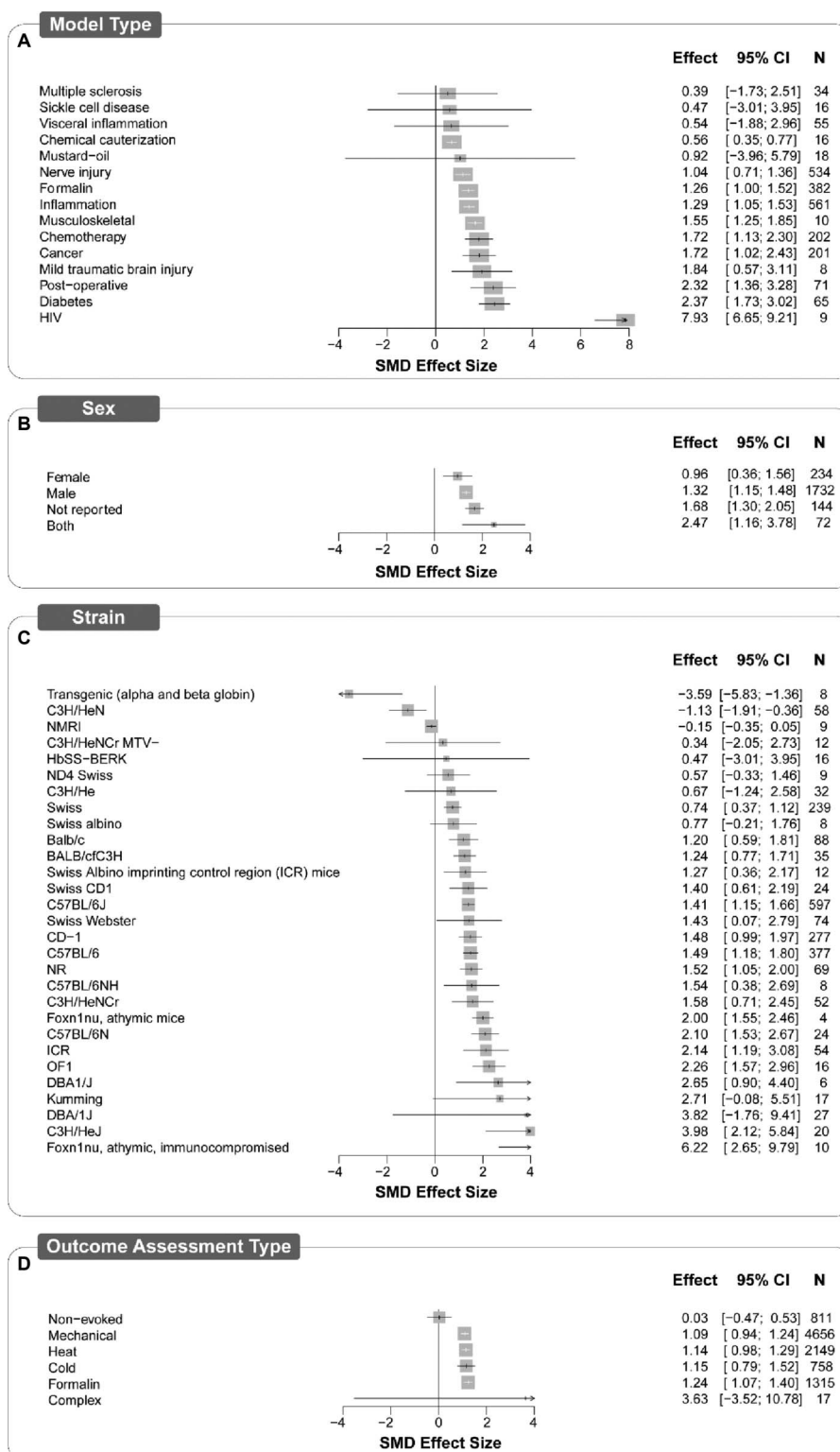
high reporting of compliance with animal welfare regulations, 94% (353). The highest number of criteria reported was 4 out of 6 in 10 studies. The methods for how bias was mitigated were rarely reported; therefore, the studies are at an unclear risk of bias (Fig. 7B; see Table, supplemental digital content 7, a traffic light plot presenting the risk of bias judgement for each study, available at <http://links.lww.com/PAIN/B338>).

Reporting of blinded assessment of outcome, predetermined animal inclusion criteria, and animal exclusions did not account for a significant proportion of heterogeneity. Allocation concealment, randomisation, and sample size calculation accounted for a significant proportion of heterogeneity. In the case of sample size calculation and allocation concealment, the low prevalence of reporting may limit our ability to accurately determine their influence

on the reported outcomes. However, larger effect sizes were reported in the studies that did not report allocation concealment and sample size calculations,  $SMD = 1.345$  vs  $SMD = 1.055$  ( $Q = 5.299$ ,  $df 1$ ,  $P = 0.021$ ) and  $SMD = 0.221$  vs  $SMD = 1.349$  ( $Q = 18.104$ ,  $df 1$ ,  $P < 0.0001$ ), respectively. It was the converse for randomisation, where larger effect sizes were reported in studies that did report randomisation ( $SMD = 1.471$  vs  $SMD = 1.245$ ,  $Q = 5.792$ ,  $df 1$ ,  $P = 0.016$ ) (Fig. 8).

#### 4.8. Publication bias

Analysis of the data from the 374 included studies has an overall effect size (SMD) of 1.321 (95% CI 1.232–1.411). Visual inspection of the funnel plot shows normal distribution. Egger's



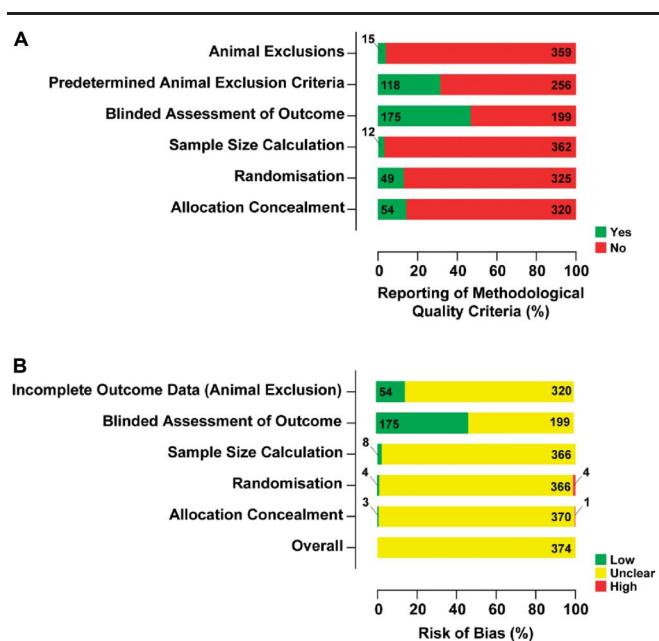
**Figure 6.** Forest plots of study design characteristics of experiments in which treatments were assessed for antinociceptive efficacy in mouse models of persistent or injury-related persistent pain. Model type (A), sex (B), strain (C), and outcome assessment type (D) all account for a significant proportion of heterogeneity. The size of the square represents the weight (%). The weight is the influence that individual subgroup has on the pooled result. N denotes the number of animals that contribute to that subgroup. SMD, standardised mean difference

regression was not consistent with effects of small studies ( $P = 0.112$ ) and did not indicate the presence of funnel plot asymmetry. Trim and fill analysis did not impute any theoretically missing studies (Fig. 9).

## 6. Discussion

We report a systematic review of preclinical studies in which cannabinoids, cannabis-based medicines, and endocannabinoid system modulators were assessed for behavioural signs of





**Figure 7.** The reporting of methodological quality criteria (A) and a summary bar plot showing the proportion of studies with a given risk of bias for each methodological quality criteria (B) for the 374 included studies. Reporting of conflicts of interest statements and compliance with animal welfare regulations were also collected but are not included in the overall risk of bias.

antinociceptive efficacy in animal models of injury-related or pathological persistent pain.

We identified 374 studies including over 16,000 rodents in which 171 different cannabinoids, cannabis-based medicines, or endocannabinoid system modulators were assessed for antinociceptive efficacy in 20 different animal model types of injury-related or pathological persistent pain. All the included studies investigated effects in rodents only, suggesting a scarcity of studies investigating antinociceptive effects in larger animals, eg, canines and primates.

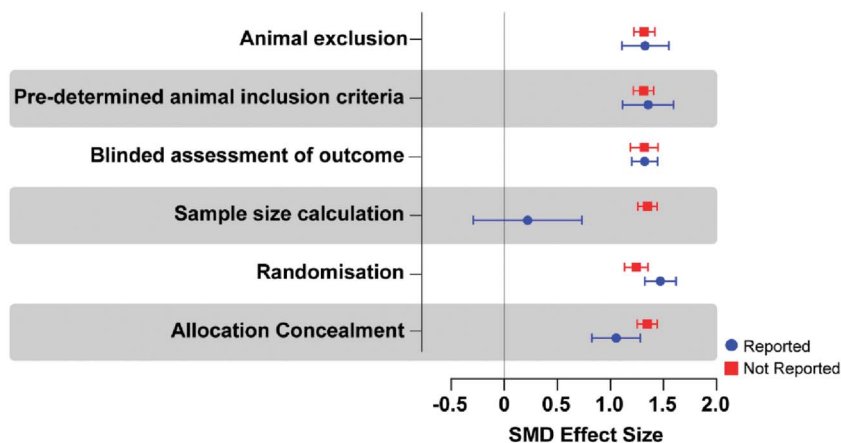
Most experiments were conducted in male Sprague-Dawley rats, with inflammation and nerve injury most frequently modelled. Antinociceptive efficacy was measured predominantly by attenuation of hypersensitivity in evoked limb withdrawal assessments. The interventions led to a statistically significant attenuation of pain-associated behaviour compared with control; the overall

SMD was 1.321 (95% CI 1.232–1.411). However, this groups together a very broad range of drugs, drug classes (some with opposing mechanisms of action), models, and pain-associated behavioural outcome measures. Thus, more useful insight of the antinociceptive efficacy of these drugs has been gained from the subgroup analyses.

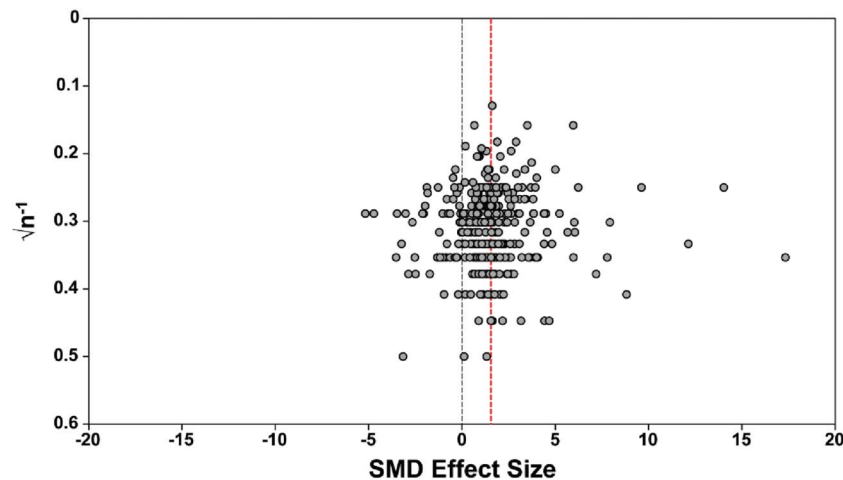
### 6.1. Antinociceptive efficacy of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators

Selective CB<sub>1</sub>, CB<sub>2</sub>, nonselective cannabinoid receptor agonists (including delta-9-tetrahydrocannabinol [THC]), and PPAR- $\alpha$  agonists (predominantly palmitoylethanolamide; PEA) significantly attenuated pain-associated behaviours in a broad range of inflammatory and neuropathic pain models. Fatty acid amide hydrolase inhibitors, MGL inhibitors, and CBD significantly attenuated pain-associated behaviours in neuropathic pain models but yielded mixed results in inflammatory pain models. The differences of antinociceptive efficacy may be inherent to the interventions but are also likely to be influenced by other study design characteristics. Careful consideration should be given to the choice of species, strain, and sex in relation to the clinical condition being modelled, coupled with the need to assess efficacy using multiple species, strains, models, and pain-associated behavioural outcomes; the conclusions will be dependent on these variables.<sup>36</sup> Increasing the biological variation will improve the generalisability of the results.

There are many different strains of the cannabis plant, each containing different amounts of phytocannabinoids.<sup>6</sup> Of note is the psychoactive component, THC, whose pharmacological effects are attributed to activity at both the cannabinoid receptors. Delta-9-tetrahydrocannabinol significantly attenuated pain-associated behaviours in a broad range of models; however, we have not been able to analyse the broader effects of the drug. Similarly, CBD significantly attenuated pain-associated behaviours in a broad range of models. Most drugs assessed were small-molecule CB<sub>1</sub> and CB<sub>2</sub> receptor agonists and nonselective agonists. The CB<sub>1</sub> and CB<sub>2</sub> receptor agonists consistently significantly attenuated pain-associated behaviours in a broad range of rat models; however, their antinociceptive effect was less consistent in the mouse models. Unlike in rats, in mouse nerve injury models, CB<sub>1</sub> and CB<sub>2</sub> receptor agonists did not significantly attenuate pain-related behaviour. We did not assess side effects (eg, motor impairment, hypothermia, or anxiolysis) that could



**Figure 8.** Standardised mean difference effect sizes (and variance) associated with the reporting of methodological quality criteria.



**Figure 9.** Assessment of publication bias. Visual inspection of the funnel plot does not suggest asymmetry. The dashed red line denotes the overall summary effect size. SMD, standardised mean difference.

influence the interpretation of pain-associated behaviour, although these assessments were not commonly reported.

Endocannabinoid system modulators, eg, FAAH and MGL inhibitors, are considered promising targets for analgesic drug development, thereby preventing the harms associated with cannabis and orthosteric cannabinoid agonists (reviewed by Guindon and Hohmann<sup>22</sup>). The evidence for efficacy of FAAH inhibitors is mixed. In rats and mice, FAAH inhibitors significantly attenuated pain-associated behaviours in nerve injury, formalin, and diabetes models. However, they did not significantly attenuate pain-associated behaviours of rat models of inflammation and both rat and mouse models of visceral inflammation. The FAAH inhibitors significantly attenuated pain-associated behaviours in CFA but not in osteoarthritis models. The evidence indicates that FAAH inhibitors are least effective for osteoarthritis but may be a viable candidate for treatment of neuropathic pain conditions. Concomitantly, the FAAH inhibitor, PF-04457845, failed to demonstrate analgesic efficacy in a randomised placebo-controlled clinical trial of osteoarthritis patients.<sup>29</sup> Our findings support the potential utility of a prospective preclinical systematic review and meta-analysis to review the animal efficacy data before clinical trial.

The drugs were grouped by class; however, the drugs themselves can activate different signalling pathways, and the signalling pathways necessary for therapeutic effect are not fully understood. This is particularly pertinent to the antinociceptive potential of inverse agonists; the data do not represent all inverse agonist studies merely when they were assessed for antinociceptive efficacy. The authors of the included studies have postulated that inverse agonists may yield an antinociceptive effect due to the subsequent reduction of proinflammatory and pronociceptive mediators.

## 6.2. External validity

### 6.2.1. Misalignment between animal models and the clinical population

The models used in preclinical pain research generally are not often well matched to the clinical population.<sup>13,44</sup> Sex only accounted for a significant proportion of heterogeneity in rats. The studies have been conducted predominantly using male

animals. It is likely that the paucity of female animals limits our ability to determine the influence of sex in the reported outcomes and hence may reduce the generalisability of the findings. Fisher et al.'s<sup>21</sup> systematic review of randomized controlled trials reported that female patients ( $n = 3691$ ) outnumbered male patients ( $n = 3613$ ) in the 34 randomized controlled trials that reported sex; drugs were assessed in patients with neuropathic pain ( $n = 13$ ), cancer ( $n = 6$ ), acute pain after surgery ( $n = 4$ ), and multiple sclerosis ( $n = 10$ ), and chronic prostatitis, carpal tunnel syndrome, and back pain in one trial each. Furthermore, most rodent studies relied upon stimulus-evoked behavioural outcome measures, which is a measure of hypersensitivity. In conditions where evoked pain is a clinically relevant aspect (eg, postsurgical and musculoskeletal conditions), hypersensitivity measures may be relevant<sup>41</sup> and within these models, the treatment effects are significant indicating potential benefits for these conditions. However, in neuropathic pain conditions, spontaneous pain is more often assessed,<sup>15</sup> and in many common neuropathic pain conditions, such as diabetic neuropathy, sensory loss is more common than evoked hypersensitivity.<sup>4,35</sup> So, although treatment effects are also significant in the neuropathic pain-associated models, translation to the clinic might be limited.

### 6.2.2. Misalignment between the cannabinoids, cannabis-based medicines, and endocannabinoid system modulators assessed in preclinical trials and clinical trials

A substantially more diverse range of potential therapeutics have been assessed for antinociceptive efficacy in animal models (171) than reported in the recent review of randomised controlled trials in which 11 interventions were assessed in patients.<sup>21</sup> In the latter review, evidence of benefit was found for cannabis <7 days and nabiximols >7 days. However, the studies had an unclear or high risk of bias with the evidence scored as low or very-low quality and the authors conclude that "the evidence neither supports nor refutes claims of efficacy." The following interventions were assessed in clinical trials in the clinical meta-analysis but do not feature in the preclinical meta-analysis: nabiximols ( $n = 17$ ), dronabinol ( $n = 2$ ), cannabinoid receptor agonist ( $n = 2$ ; AZD1940 and GW842166), and THC congener ( $n = 1$ ; benzopyran peridine).<sup>21</sup> It is possible that these drugs have not

been tested for efficacy in animal models and/or the studies in which their effects are described are not reported with sufficient detail to be included in the meta-analysis.

The observed misalignment between preclinical and clinical trials suggests that the animal studies are not being optimally used to inform or predict direction for clinical trials or efficacy in the clinic. The justifications for clinical trials were likely borne from patients and their use of cannabis to alleviate pain. Animal studies are being conducted concurrently to clinical studies providing opportunity for both translation and back-translation. Animal to human translation will always be unpredictable, especially considering the challenges of pain research. However, animal studies in the field of cannabinoid research may have greater utility than is currently being recognised. They can be especially useful for providing mechanistic insights into the pharmacology of cannabinoids, cannabis based-medicines, and endocannabinoid system modulators, to facilitate development of human therapies.

### 6.3. Internal validity

All included studies had an unclear risk of bias. The propensity to report methodological quality measures to reduce the risk of bias was low, and this finding is commensurate with other pain preclinical systematic reviews (eg, Currie et al.<sup>10</sup>). It likely reflects the fact that the reporting of many of these quality measures has not been required by journals nor convention in the preclinical field until recently. The included studies rarely reported the performance of power calculations to determine sample size (3%) or animal exclusions (14%); however, randomisation and blinding were more frequently reported (32% and 47%, respectively). Our analyses did not show a consistent relationship between the reporting of methodological quality and smaller effect sizes (akin to Federico et al.<sup>18</sup>); however, larger effect sizes were reported in studies that did not report allocation concealment and sample size calculations, both accounting for a significant proportion of heterogeneity. The methods used to mitigate bias were rarely reported, and it was therefore not possible to accurately assess the risk of bias, leading to uncertainty in the validity of the outcomes.

Sample size should be determined using a power analysis, and experiments are required to use sufficient animals to be adequately powered. In relation, many experiments compared multiple treatment groups with one control group. This reduced the sensitivity because the control group was divided across the multiple treatment groups (the mean number of animals in the control group was reduced from  $n = 8$  to  $n = 3$ ). Bate and Karp<sup>5</sup> provide a strategy for reducing the risk of false positives by increasing the number of animals in the control group, although this is also contingent on the other factors that comprise a sample size calculation (eg, effect, variability, significance level, analysis, and experimental constraints). Using more animals is commensurate with the reduction, refinement, and replacement (3R) principles because it ensures that the animal sacrifice is weighted against the highest possible gain of knowledge, although existing animal care and use committees may view this as a conflict.

### 6.4. Publication bias

Unlike a recent preclinical systematic review of pregabalin in which a 27% overestimate of effect was theorised,<sup>18</sup> our analysis does not suggest presence of publication bias, the phenomenon wherein neutral or negative studies are not published.

### 6.5. Future research

The number of studies that concurrently conducted and reported pharmacokinetic investigation (24 studies) including drug concentrations in tissues after administration (25 studies) was low at 7% of studies. There is a need to conduct pharmacokinetic studies alongside pharmacodynamic studies to determine the relationship between plasma/tissue concentrations of treatments and reflexive or complex nociceptive behavioural measures. Few studies assessed the effects of the drugs on sedation with 124 studies (33%) assessing impact of drugs on motor activity, which is a particular concern for direct-acting CB<sub>1</sub> receptor agonists (but would not be expected to be a confound for CB<sub>2</sub> agonists, FAAH inhibitors, anandamide transport inhibitors, or CB<sub>1</sub> allosteric modulators). Similarly, few studies assess the anxiolytic or depressive-like effects (11 studies, 3%) of the drugs, which may also compound pain-associated behavioural outcomes. A broader range of assessment is required to determine the full behavioural effect of these drugs.<sup>38</sup> It was outside the scope of this systematic review and meta-analysis to assess the possible side effects of these treatments, but to assist future reviews, studies assessing pharmacokinetics, locomotor activity, anxiety, and depression were annotated.

Rigorous preclinical design is required for internal validity.<sup>28</sup> To limit threats to validity, we endorse conducting and reporting animal experiments as suggested by Andrews et al.,<sup>2</sup> Vollert et al.,<sup>31,54</sup> and in accordance with the ARRIVE guidelines.<sup>40</sup> To improve external validity, researchers should balance the sexes as stipulated by funding bodies including the U.S. National Institutes of Health.<sup>8,49</sup> Historically, emphasis has been placed upon reflex withdrawal responses rather than measuring more complex, ethologically relevant behaviours, although the clinical relevance of these measures to analgesia remains uncertain. The development of sensory profiling for rodents and complex behavioural animal models specific to each pain condition is required to improve face and predictive validity and better reflect the clinical situation (Rice et al., 2017). Multicentre testing may offer a method to improve the generalisability of preclinical findings by increasing environmental heterogeneity and study samples.<sup>53,55</sup> To address the potential issue of publication bias, we recommend researchers make available prespecified protocols and publish all results (ie, positive, null, and negative data). To assist in the optimisation of experimental design, we encourage primary researchers to conduct prospective systematic reviews and use the U.K. National Centre for the Replacement, Refinement and Reduction of Animals of Research Experimental Design Assistant (<https://eda.nc3rs.org.uk/>) to inform their research design and protocol development.<sup>12</sup>

### 6.6. Limitations

Our systematic review has several limitations. First, we can only rely upon what has been reported in publications. There were 99 studies that met the inclusion criteria but could not be included in the meta-analysis due to not reporting key information, eg, variance, sample size, or not having suitable controls. For the included studies, it is possible that methods were used to reduce the risk of bias, but not reported; conversely, these methodological quality criteria may have been reported but not performed. There are also other experimental design factors that will influence behavioural outcomes but are not included in our analyses, eg, housing, diet, handling, habituation,<sup>37</sup> and the sedative effect of the drugs.

We collected all pain-associated behavioural outcomes of all interventions of any dose or route of administration that were being assessed for antinociception. We did not collect information on the aim of the studies. The behavioural studies may have been conducted not to determine efficacy, but for a different rationale.

### 6.6.1. Pain-associated behavioural outcomes and outcome data extraction

In our meta-analysis, we grouped together models by the same underlying biology. Similarly, we grouped together behavioural outcome measures. Most of the outcomes measured were limb withdrawal in response to mechanical or thermal stimuli and despite not having the same underlying biology, these were grouped together if a cohort was assessed in both. There is large variation in how these studies are reported, and it is challenging to identify differences in study design when these are not often reported in detail. In addition, we chose to extract pain-associated behavioural outcome data at the time point where the difference was largest between vehicle and treatment groups. This allowed us to calculate treatment effects independent of the intervention's half-life, particularly pertinent as many studies did not report time course data. Although not feasible within this review, given the number of interventions, doses etc., a future approach may be to calculate area under the curve and percentage maximum possible effect for each experiment at all reported doses allowing more information to be gleaned.

### 6.6.2. Crowd science and machine learning

This review demonstrates that crowd science and machine learning are viable strategies to improve the feasibility of conducting a large review. Our experience supports a recent Cochrane study that demonstrates the feasibility of study identification for inclusion in a systematic review using crowd science and machine learning.<sup>39</sup> Crowd science offers a reduction in individual input by sharing the labour-intensive stages of the review (screening for study selection and annotation and data extraction phases). Errors by crowd members were detected during the reconciliation process and ~20% of studies required re-review by expert reviewers. The limitations of individual crowd members therefore have not undermined the findings of this review. Using a crowd has several hard-to-quantify benefits including increasing diversity, reducing bias, community engagement, training, and education. Importantly, this review has also demonstrated the usefulness of machine learning for study selection, albeit for error analysis. The machine algorithm performed with an eventual sensitivity of 95% and specificity of 94%. This high sensitivity made it unlikely to miss relevant literature.

The search for this review was conducted nearly 2 years ago; given the size of this review, it is unlikely that the incorporation of more recent studies will change the overall conclusions. The continued development of online platforms and automation technologies is required to improve the feasibility of preclinical systematic reviews. This will create the possibility to incorporate the most recent data as it becomes available in the form of a living systematic review,<sup>17</sup> thereby addressing the challenge for the future appraisal of the rapid and exponentially increasing volume of published preclinical literature.

## 7. Conclusion

This systematic review and meta-analysis provides a comprehensive summary of studies in which cannabinoids, cannabis-based

medicines, and endocannabinoid system modulators were assessed for antinociceptive efficacy in animal models of injury-related or pathological persistent pain. The behavioural data effect sizes are significant, and the evidence supports the hypothesis of cannabinoid-induced analgesia. Most drugs tested in animal models were small molecules, which is converse to the clinical situation where cannabis extracts have been evaluated most in clinical trials. The differences between the animal and clinical population highlight the importance for the development of better validated animal models. Behavioural assessments that have greater clinical relevance may also improve the likelihood of the development of effective therapeutic interventions. There is also a need to continually improve clinical trial design in a manner that is informed by high-quality, mechanism-based preclinical research. Despite the “unknown” predictive value of many animal studies, there is value in conducting a prospective systematic review to aid clinical trial decision making. The findings of this review support the need for preclinical living systematic reviews and closer, multidisciplinary, cross-sector collaboration to ensure that animal studies are rigorous to identify potential candidates and more accurately inform clinical trial design.

### 7.1. Glossary

A glossary (Table 8) provides brief explanations for the terms used throughout this systematic review.

### Conflict of interest statement

S. Haroutounian reports grants from Disarm Therapeutics, personal fees from Medoc Ltd, personal fees from Rafa Laboratories, and personal fees from Vertex Pharmaceuticals, outside the submitted work. A.G. Hohmann has a patent Methods of using cannabinoid CB2 cannabinoid receptor agonist compositions to suppress and prevent opioid tolerance and withdrawal US Application Serial No: 16/256,787 (IURTC Ref. 2018-072-02) Inventors: Andrea Grace Hohmann, Ken Paul Mackie, Xiaoyan Lin, Amey Dhopeshwarkar Assignees: Indiana University Research and Technology Corporation pending, and a patent A novel mechanism for decreasing opioid reward US 62/675,944 pending. J. Vollert reports personal fees from Vertex Pharmaceuticals, outside the submitted work. A. Barakat reports funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 814244. N.M. Malewicz reports grants from DFG Deutsche Forschungsgemeinschaft, and nonfinancial support from DGSS Deutsche Schmerzgesellschaft e.V., outside the submitted work. E.M. Pogatzki-Zahn reports personal fees from Mundipharma, personal fees from Grünenthal, grants from Mundipharma, and grants from Grünenthal (to the institution) outside the submitted work; and E.M. Pogatzki-Zahn receives scientific support (to the institution) from the DFG, the BMBF, and the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 777500. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. J.A. Yanes was supported by the National Institute on Drug Abuse under award number: F31DA044013. D.P. Finn reports an Industry-Academia research grant from Alkermes, Inc. and Science Foundation Ireland, outside of the submitted work. He also reports research grants in the area of cannabinoids or the endocannabinoid system from Shionogi Ltd (Shionogi Science Programme), from B. Braun Ltd jointly with Science Foundation Ireland, and from the Irish Research Council, CNPq Brazil, and EU INTERREG Programmes. A.S.C. Rice reports other from IASP, during the



**Table 8****Glossary.**

Systematic Review	Use predefined methods to identify, select, and critically appraise all available literature to address a specific research question
Meta-analysis	The statistical combination of quantitative results (pain-associated behavioural outcomes) of 2 or more studies. The methods included in the meta-analysis are the calculation of effect sizes, the pooling of the effects so that the range and distribution of effects can be observed
Study	In this instance, a study refers to the publication. A publication can have multiple experiments in which an intervention is tested in a cohort of animals and a pain-associated behaviour is measured. There can be multiple outcome measures per cohort. Similarly, a study can have multiple experiments.
Comparison and nested comparison	The outcome measure of a treatment group compared to a control (vehicle-treated) group is a comparison. Often the same cohort of animals undergo multiple pain-associated behavioural outcome measurements. In these instances, the comparisons are combined to give one outcome statistic (a nested comparison) that represents the global measure of the outcomes in that comparison.
Effect size	For each comparison, an effect size is calculated using standardised mean difference (SMD). The difference between group means (mean of control group – mean of experimental group) is divided by the pooled variance, which converts all outcome measures into a standardised scale. (A correction factor, 1 or –1 is used to define the direction of the effect size, whether the outcome is better or worse in comparison to the control).
Heterogeneity	Study heterogeneity denotes the variability in outcomes that are not due to measurement errors but other influencing factors (eg, study characteristics). We have estimated heterogeneity using both Cochran's Q and $I^2$ and explored sources of heterogeneity with stratified meta-analysis.
Estimating heterogeneity with Cochran's Q	Q is an estimate of between-study heterogeneity and is calculated from effect sizes. It is based on a chi-squared distribution. A larger Q value denotes larger variation across studies rather than within subjects within a study. The P value of Q is used to indicate the presence or absence of heterogeneity.
Estimating heterogeneity with $I^2$	$I^2$ is the proportion of total variance between studies that is due to true differences in effect sizes, not differences that are due to chance. If $I^2 = 0\%$ all variation is due to chance alone, 100% all variation is due to differences between the true effect sizes between studies. 0–25%—very low heterogeneity 25–50%—low heterogeneity 50–75%—moderate heterogeneity >75%—high heterogeneity
Stratified analysis	Studies that share a particular characteristic, eg, sex, strain, animal model, will be more similar than studies that do not share the same characteristic. Stratified analysis allows us to partition the heterogeneity between groups of similar studies and between groups of studies to determine whether the differences are statistically significant.
Animal model	Whole in vivo animal models of pathological or injury-related persistent pain, eg, tissue injury, cancer, chemotherapy-induced, inflammation, or nerve damage. Persistent pain was defined as studied over a period of hours, days, weeks, or months.
Pain-associated behavioural outcome	These were when pain was declared the reason for assessment by the authors. Behavioural outcomes include: Evoked limb withdrawal to mechanical, heat, or cold stimuli Spontaneous, eg, weight-bearing difference, spontaneous foot lifting, grimace scale, and nocifensive behaviour Complex, eg, open-field test (thigmotaxis) and burrowing
Antinociception	Attenuation of pain-associated behaviour

conduct of the study, personal fees from Imperial College Consultants, and other from Spinifex, outside the submitted work. In addition, A.S.C. Rice has a patent WO 2005/079771 pending, and a patent WO 2013/110945 pending. The remaining authors have no conflicts of interest to declare.

## Acknowledgements

The authors thank the following staff from the International Association for the Study of Pain for their unwavering support throughout in particular for their organisation and assistance in the recruitment of crowd scientists: Colleen Eubanks, Matthew D'Uva, Yulanda Grant, Diana Bender-Bier, Luke Waldron, and Neil Andrews. The authors also thank Dr Andrew Moore for his comments on the manuscript. Finally, the authors thank the crowd members who did not qualify for authorship but still contributed to the investigation (the screening and data extraction phases): Danilyn Amerna, Rita Bertani, Ines Boujelben, Blake A. Castetter, Conor Clonan, Dr Yogesh Dhakal, Adele Dobry, Dr Marieke H.J. van den Beuken-van Everdingen, Dr McKenzie

C. Ferguson, Alexander T. Gordan, Sonal Gupta, Dr Helen C. Laycock, Dr Visnja Muzika, Dr Pooya Rostami, Raluca Scarlat, Dr Aidan C. Tan, and Dr Ere Vicencio.

Author contributions: (visually represented in supplemental digital content 8, available at <http://links.lww.com/PAIN/B366>) Funding acquisition: A. S. C. Rice; Conceptualisation: N. Soliman, S. Haroutounian, A.G. Hohmann, D.P.Finn, and A. S. C. Rice; Methodology: N. Soliman, S. Haroutounian, A.G. Hohmann, E. Krane, J. Liao, M. Macleod, J. Thomas, J. Vollert, K. Wever, and D.P.Finn, A. S. C. Rice; Software Development: N. Soliman, C. Sena; Investigation: N. Soliman, H. Alaverdyan, A. Barakat, T. Barthlow, A. Bozer, A. Davidson, M. Diaz-delCastillo, A. Dolgorukova, M. Ferdousi, C. Healy, S. Hong, M. Hopkins, A. James, H. Leake, N. Malewicz, M. Mansfield, A. Mardon, D. Mattimoe, D. McLoone, G. Noes-Holt, E. Pogatzki-Zahn, E. Power, B. Pradier, E. Romanos-Sirakis, A. Segelcke, D. Segelcke, R. Vinagre, J. Yanes, J. Zhang, X.Y. Zhang, Project administration: N. Soliman; Supervision: N. Soliman, D.P.Finn, and A. S. C. Rice; Data curation: N. Soliman; Formal analysis: N. Soliman, S. Haroutounian, A.G. Hohmann, M. Macleod, J. Vollert, D.P.Finn,

and A. S. C. Rice; Visualisation: N. Soliman and D. Segelcke; Writing—original draft: N. Soliman; Writing—review and editing: N. Soliman, S. Haroutounian, A.G. Hohmann, E. Krane, J. Liao, M. Macleod, D. Segelcke, C. Sena, J. Thomas, J. Vollert, K. Wever, H. Alaverdyan, H. Alaverdyan, A. Barakat, T. Barthlow, A. Bozer, A. Davidson, M. Diaz-delCastillo, A. Dolgorukova, M. Ferdousi, C. Healy, S. Hong, M. Hopkins, A. James, H. Leake, N. Malewicz, M. Mansfield, A. Mardon, D. Mattimoe, D. McLoone, G. Noes-Holt, E. Pogatzki-Zahn, E. Power, B. Pradier, E. Romanos-Sirakis, A. Segelcke, R. Vinagre, J. Yanes, J. Zhang, X.Y. Zhang, D. P.Finn, and A. S. C.Rice.

Open Science Practise Statement: The systematic review protocol was formally pre-registered on PROSPERO (CRD42019124804; [https://www.crd.york.ac.uk/prospero/display\\_record.php?ID=CRD42019124804](https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019124804)) and published in PAIN Reports (Soliman et al., 2019). The data have been made available on the Open Science Framework DOI 10.17605/OSF.IO/2QDE5.

This work is part of the effort of the International Association for the Study of Pain Presidential Taskforce on Cannabis and Cannabinoid Analgesia. The work is also supported by the BBSRC (grant number BB/M011178/1).

## Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at <http://links.lww.com/PAIN/B332>, <http://links.lww.com/PAIN/B333>, <http://links.lww.com/PAIN/B334>, <http://links.lww.com/PAIN/B335>, <http://links.lww.com/PAIN/B336>, <http://links.lww.com/PAIN/B337> and <http://links.lww.com/PAIN/B338>.

## Article history:

Received 25 January 2021

Received in revised form 12 March 2021

Accepted 12 March 2021

Available online 15 March 2021

## References

- Alexander SPH, Kelly E, Mathie A, Peters JA, Veale EL, Armstrong JF, Faccenda E, Harding SD, Pawson AJ, Sharman JL, Southan C, Buneman OP, Cidlowski JA, Christopoulos A, Davenport AP, Fabbro D, Spedding M, Striessnig J, Davies JA, Collaborators C. The concise guide to pharmacology 2019/20: introduction and other protein targets. *Br J Pharmacol* 2019;176:S1–S20.
- Andrews NA, Latremoliere A, Basbaum AI, Mogil JS, Porreca F, Rice ASC, Woolf CJ, Currie GL, Dworkin RH, Eisenach JC, Evans S, Gewandter JS, Gover TD, Handwerker H, Huang W, Iyengar S, Jensen MP, Kennedy JD, Lee N, Levine J, Lidster K, Machin I, McDermott MP, McMahon SB, Price TJ, Ross SE, Scherrer G, Seal RP, Sena ES, Silva E, Stone L, Svensson CI, Turk DC, Whiteside G. Ensuring transparency and minimization of methodologic bias in preclinical pain research: PPRECISE considerations. *PAIN* 2016;157:901–9.
- Bannach-Brown A, Przybyla P, Thomas J, Rice ASC, Ananiadou S, Liao J, Macleod MR. Machine learning algorithms for systematic review: reducing workload in a preclinical review of animal studies and reducing human screening error. *Syst Rev* 2019;8:23.
- Baron R, Binder A, Wasner G. Neuropathic pain: diagnosis, pathophysiological mechanisms, and treatment. *Lancet Neurol* 2010;9:807–19.
- Bate S, Karp NA. A common control group—optimising the experiment design to maximise sensitivity. *PLoS One* 2014;9:e114872.
- Berman P, Futoran K, Lewitus GM, Mukha D, Benami M, Shlomi T, Meiri D. A new ESI-LC/MS approach for comprehensive metabolic profiling of phytocannabinoids in Cannabis. *Sci Rep* 2018;8:14280.
- Clark JD. Preclinical pain research: can we do better? *Anesthesiology* 2016;125:846–9.
- Clayton JA, Collins FS. Policy: NIH to balance sex in cell and animal studies. *Nature* 2014;509:282–3.
- Crequit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. *Bmc Med* 2016;14:8.
- Currie GL, Angel-Scott HN, Colvin L, Cramond F, Hair K, Khandoker L, Liao J, Macleod M, McCann SK, Morland R, Sherratt N, Stewart R, Tanriver-Ayder E, Thomas J, Wang QY, Wodarski R, Xiong R, Rice ASC, Sena ES. Animal models of chemotherapy-induced peripheral neuropathy: a machine-assisted systematic review and meta-analysis. *PLoS Biol* 2019;17:e3000243.
- de Vries RBM, Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Updated version of the Embase search filter for animal studies. *Lab Anim* 2014;48:88.
- du Sert NP, Bamsey I, Bate ST, Berdoy M, Clark RA, Cuthill IC, Fry D, Karp NA, Macleod M, Moon L, Stanford SC, Lings B. The Experimental Design Assistant. *Nat Methods* 2017;14:1024–25. Available at: <https://doi.org/10.1038/nmeth.4462>.
- du Sert NP, Rice ASC. Improving the translation of analgesic drugs to the clinic: animal models of neuropathic pain. *Br J Pharmacol* 2014;171:2951–63.
- Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000;56:455–63.
- Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tolleit J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *PAIN* 2005;113:9–19.
- Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997;315:629–34.
- Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, Salanti G, Meerpohl J, MacLehose H, Hilton J, Tovey D, Shemilt I, Thomas J. Living Systematic Review N. Living systematic review: 1. Introduction—the why, what, when, and how. *J Clin Epidemiol* 2017;91:23–30.
- Federico CA, Mogil JS, Ramsay T, Fergusson DA, Kimmelman J. A systematic review and meta-analysis of pregabalin preclinical studies. *PAIN* 2020;161:684–93.
- Finn DP, Haroutounian S, Hohmann A, Krane E, Soliman N, Rice ASC. Cannabinoids, the endocannabinoid system and pain: a review of preclinical studies. *PAIN* 2021;162(S1):S5–S25.
- Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Dworkin RH, Gilron I, Haanpää M, Hansson P, Jensen TS, Kamerman PR, Lund K, Moore A, Raja SN, Rice AS, Rowbotham M, Sena E, Siddall P, Smith BH, Wallace M. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. *Lancet Neurol* 2015;14:162–73.
- Fisher E, Moore RA, Fogarty AE, Finn DP, Finnerup NB, Gilron I, Haroutounian S, Krane E, Rice ASC, Rowbotham M, Wallace M, Eccleston C. Cannabinoids, cannabis, and cannabis-based medicine for pain management: a systematic review of randomised controlled trials. *PAIN* 2021;162(S1):S45–S66.
- Guindon J, Hohmann AG. The endocannabinoid system and pain. *CNS Neurol Disord Drug Targets* 2009;8:403–21.
- Harrer R, Cuijpers P, Furukawa TA, Ebert DD. Doing meta-analysis in R: a hands-on guide. PROTECT Lab Erlangen 2019. doi: 10.5281/zenodo.2551803 [Epub ahead of print].
- Hartung J. An alternative method for meta-analysis. *Biometric J* 1999;41:901–16.
- Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat Med* 2001;20:1771–82.
- Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCL's risk of bias tool for animal studies. *BMC Med Res Methodol* 2014;14:43.
- Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Enhancing search efficiency by means of a search filter for finding all studies on animal experimentation in PubMed. *Lab Anim* 2010;44:170–5.
- Huang W, Percie du Sert N, Vollert J, Rice ASC. General principles of preclinical study design. In: Bepalov A, Michel MC, Steckler T, editors. Good Research Practice in Non-Clinical Pharmacology and Biomedicine. Cham: Springer International Publishing, 2020. pp. 55–69.
- Huggins JP, Smart TS, Langman S, Taylor L, Young T. An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of the knee. *PAIN* 2012;153:1837–46.

- [30] Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005;2:696–701.
- [31] Knopp KL, Stenfors C, Baastrup C, Bannan AW, Calvo M, Caspani O, Currie G, Finnerup NB, Huang W, Kennedy JD, Lefevre I, Machin I, Macleod M, Rees H, Rice ASC, Rutten K, Segerdahl M, Serra J, Wodarski R, Berge OG, Treede RD. Experimental design and reporting standards for improving the internal validity of pre-clinical studies in the field of pain: consensus of the IML-Europain consortium. *Scand J Pain* 2015;7:58–70.
- [32] Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubitza AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetz W, Krainc D, Lazic SE, Levine MS, Macleod MR, McCall JM, Moxley RT III, Narasimhan K, Noble LJ, Perrin S, Porter JD, Steward O, Unger E, Utz U, Silberberg SD. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 2012;490:187–91.
- [33] Macleod Malcolm R, O'Collins T, Howells David W, Donnan Geoffrey A. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* 2004;35:1203–8.
- [34] Magnusson K. Interpreting Cohen's d effect size: An interactive visualization (Version 2.5.0) [Web App]. *R Psychologist*, 2021. Available at: <https://rpsychologist.com/cohend/>.
- [35] Maier C, Baron R, Tölle TR, Binder A, Birbaumer N, Birklein F, Gierthmühlen J, Flor H, Geber C, Hüge V, Krumova EK, Landwehrmeyer GB, Magerl W, Maihöfner C, Richter H, Rolke R, Scherens A, Schwarz A, Sommer C, Tronnier V, Üçeyler N, Valet M, Wasner G, Treede DR. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. *PAIN* 2010;150:439–450.
- [36] Mogil JS. Animal models of pain: progress and challenges. *Nat Rev Neurosci* 2009;10:283–94.
- [37] Mogil JS. Laboratory environmental factors and pain behavior: the relevance of unknown unknowns to reproducibility and translation. *Lab Anim* 2017;46:136–41.
- [38] Negus SS. Core outcome measures in preclinical assessment of candidate analgesics. *Pharmacol Rev* 2019;71:225.
- [39] Noel-Storr A, Dooley G, Affengruber L, Gartlehner G. Citation screening using crowdsourcing and machine learning produced accurate results: evaluation of Cochrane's modified Screen4Me service. *J Clin Epidemiol* 2021;130:23–31.
- [40] Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, Clark A, Cuthill IC, Dirnagl U, Emerson M, Garner P, Holgate ST, Howells DW, Karp NA, Lidster K, MacCallum CJ, Macleod M, Petersen O, Rawle F, Reynolds P, Rooney K, Sena ES, Silberberg SD, Steckler T, Würbel H. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLOS Biology* 2020;18: e3000410.
- [41] Pogatzki-Zahn E, Segelcke D, Zahn P. Mechanisms of acute and chronic pain after surgery: update from findings in experimental animal models. *Curr Opin Anesthesiol* 2018;31:575–85.
- [42] Pogatzki-Zahn EM, Segelcke D, Schug SA. Postoperative pain-from mechanisms to treatment. *Pain Rep* 2017;2:e588.
- [43] Rice ASC, Cimino-Brown D, Eisenach JC, Kontinen VK, Lacroix-Fralish ML, Machin I, Mogil JS, Stöhr T. Animal models and the prediction of efficacy in clinical trials of analgesic drugs: a critical appraisal and call for uniform reporting standards. *PAIN* 2008;139:243–7.
- [44] Rice ASC, Finnerup NB, Kemp HI, Currie GL, Baron R. Sensory profiling in animal models of neuropathic pain: a call for back-translation. *PAIN* 2018;159:819–24.
- [45] Rice ASC, Smith BH, Blyth FM. Pain and the global burden of disease. *PAIN* 2016;157:791–6.
- [46] Sideli L, Trotta G, Spinazzola E, La Cascia C, Di Forti M. Adverse effects of heavy cannabis use: even plants can harm the brain. *PAIN* 2021;162(S1): S97–S104.
- [47] Soliman N, Hohmann AG, Haroutounian S, Wever K, Rice ASC, Finn DP. A protocol for the systematic review and meta-analysis of studies in which cannabinoids were tested for antinociceptive effects in animal models of pathological or injury-related persistent pain. *Pain Rep* 2019;4:e766.
- [48] Starowicz K, Finn DP. Cannabinoids and pain: sites and mechanisms of action. *Adv Pharmacol* 2017;80:437–75.
- [49] Tannenbaum C, Greaves L, Graham ID. Why sex and gender matter in implementation research. *BMC Med Res Methodol* 2016;16:145.
- [50] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, Kuss O, Higgins JP, Langan D, Salanti G. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* 2016;7:55–79.
- [51] Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, Antonic A, Howells DW, Macleod MR. Meta-analysis of data from animal studies: a practical guide. *J Neurosci Methods* 2014;221:92–102.
- [52] Viechtbauer W. Conducting meta-analyses in R with the metafor package. 2010;36(3):48.
- [53] Voelkl B, Vogt L, Sena ES, Würbel H. Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biol* 2018;16:e2003693.
- [54] Vollert J, Schenker E, Macleod M, Bepalov A, Wuerbel H, Michel M, Dirnagl U, Potschka H, Waldron AM, Wever K, Steckler T, van de Casteele T, Altevogt B, Sil A, Rice ASC. Systematic review of guidelines for internal validity in the design, conduct and analysis of preclinical biomedical experiments involving laboratory animals. *BMJ Open Sci* 2020;4:e100046.
- [55] Wodarski R, Delaney A, Ultenius C, Morland R, Andrews N, Baastrup C, Bryden LA, Caspani O, Christoph T, Gardiner NJ, Huang W, Kennedy JD, Koyama S, Li D, Ligocki M, Lindsten A, Machin I, Pekcec A, Robens A, Rotariu SM, Voß S, Segerdahl M, Stenfors C, Svensson CI, Treede RD, Uto K, Yamamoto K, Rutten K, Rice ASC. Cross-centre replication of suppressed burrowing behaviour as an ethologically relevant pain outcome measure in the rat: a prospective multicentre study. *PAIN* 2016;157:2350–65.
- [56] Woodhams SG, Chapman V, Finn DP, Hohmann AG, Neugebauer V. The cannabinoid system and pain. *Neuropharmacology* 2017;124:105–20.
- [57] Zwetsloot PP, Van Der Naald M, Sena ES, Howells DW, Int'Hout J, De Groot JAH, Chamuleau SAJ, MacLeod MR, Wever KE. Standardized mean differences cause funnel plot distortion in publication bias assessments. *eLife* 2017;6:e24260.