

https://dspace.org/

# Reporting of Statistical Results in the Field of Pain Research

**Chinaka, Tapiwa Tsitsi**

**2025**
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/49875
Chinaka, Tapiwa Tsitsi . (2025). Reporting of Statistical Results in the Field of Pain Research [Master's dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/49875

Downloaded from a DSpace repository. https://dspace.org/

# REPORTING OF STATISTICAL RESULTS IN THE FIELD OF PAIN RESEARCH

**Tapiwa Tsitsi Chinaka**

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in the fulfilment of the requirements for the degree of Master of Science (Medicine)

Johannesburg, South Africa, August 2025



# DECLARATION

I, Tapiwa Tsitsi Chinaka, declare that this dissertation is my own work, except where otherwise specified. It is submitted for the degree of MSc Medicine in the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg. The work herein has not been submitted before for any degree or examination at any other University.



_____

Tapiwa Tsitsi Chinaka

____25th ____ Day of _____August_____2025

# DEDICATION

*To God be the glory for the great things He has done*

# ABSTRACT

Pain research aims to understand the mechanisms, treatments, and impact of pain on individuals and populations. Experimental studies, where researchers manipulate variables and measure outcomes, are crucial in this field. Statistical tests are used to interpret the data and minimize false claims about the effects of manipulations, and the primary output of these tests that is used to minimise such claims in the probability (P) value. Technically, a P-value represents the probability of obtaining the observed data, or more extreme data, assuming the null hypothesis (no effect) is true. Because of this definition, the P-value is frequently interpreted as the strength of evidence against the null hypothesis. A smaller P-value indicates stronger evidence against the null hypothesis.

Null hypothesis significant testing uses a P-value cutoff, typically 0.05, to decide if results are statistically significant. However, reliance on this cutoff can lead to misinterpretations about the importance of results. A low P-value is not a guarantee against false positives. Indeed, the threshold creates the problem of false positives. Nor does a small P-value indicate the biological or clinical importance of a result.

In my research, I aimed to assess the quality and integrity of statistical reporting in the journal *Pain* by reviewing P-value reporting in original experimental research articles from 2009-2010 and from 2019-2020 (to account for possible changes with time). My investigation focused on the distribution of P-values according to the type of experimental research (e.g., animal, human), the type of p-value (exact or approximate), and the period when articles were published. I also examined the reporting of test statistics and sample sizes in relation to the type of p-value being reported. Finally, I analysed the distribution of exact P-values and the presence of possible P-hacking (a marker of questionable research practices) using P-curves. My goal was to provide insights into the statistical reporting practices in a leading pain journal.

All manuscripts published in the journal *Pain* from 1 January 2009 to 31 December 2010, and 1 January 2019 to 31 December 2020, were reviewed. The articles underwent a screening process based on predefined inclusion and exclusion criteria. Original experimental research papers with at least one identifiable primary hypothesis/objective, and at least one P-value associated with that hypothesis/objective, were included in our study sample. A standardized form developed using REDCap was used for capturing article details, including author names, article title, publication year, volume, issue, page numbers, DOI, study type, objectives, and statistical information. The articles that met our selection criterion were classified into the following categories: basic science (human), basic science (animal), basic science (other), and clinical science. We assessed the quality of P-value reporting and the evidence for P-hacking using the P-curve.

Out of 1,342 articles in *Pain*, 738 met the selection criteria. Of these articles, 98% of experimental papers from 2009-2010 and 97% from 2019-2020 clearly stated a primary hypothesis and had a P-value associated with that hypothesis. Analysis of P-values over time revealed a shift towards reporting exact P-values, with the reporting increasing from 33% in 2009-2010 to 42% in 2019-2020. There also was a positive relationship between the use of exact P-values, instead of approximate P-values, and the reporting of sample size and test statistics with the P-value. The distribution of significant ($P < 0.05$) and non-significant P-values showed no significant differences across article types. However, a higher proportion of "highly" significant P-values ($P < 0.001$) was observed in 2019-2020 compared to in 2009-2010, suggesting either improved research rigor or more selective reporting of significant data. P-curve analysis indicated potential P-hacking in 2009-2010 but not in 2019-2020, highlighting possible changes in research practices over the decade.

The findings of this study have important implications for future research in pain management. I assessed data for the leading journal (by impact factor ranking) in the field of pain. The

improvement in the statistical reporting, as shown by the general leaning towards more precise statistical reporting and the absence of P-hacking or publication bias in 2019-2020, as compared to 2009-2010, enhances the credibility of the research in this field. Whether such credibility can be applied across other journals still needs to be established.

# ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisors, Professor Peter Kamerman and Professor Andrea Fuller, for their invaluable guidance, support, and encouragement throughout this research journey. Their expertise and insights have been instrumental in shaping this thesis. I am truly grateful for many hours we spend in discussion and sometimes having to repeat some concepts many times for me to grasp. I would like to thank them for not giving up on me, even though it would have been understandable for them to do so. This work is a total sum of their hard work, patience, and dedication.

I am also grateful to the Head of the School of Physiology, Professor Willie Daniels, for giving me the time to work on my research. Your support and advice have been greatly appreciated.

To Felicity Siboza, Courtney Jeffery, and Dhuriya Cajee, thank you for your help in quality control of my work. Your assistance has been invaluable.

To my fellow "Pain Lab" students, thank you for both the formal and informal engagements. The coffee writing dates, and several brainstorming sessions, always pushed me to work harder. Your support and team spirit have been invaluable. Special thanks to Dhuriya, your acts of kindness, especially during my ‘final push’, meant the world to me. Who knew that consistent encouragement, thoughtful gifts, and camaraderie could work wonders?

To my colleagues in Brain Function Research Group and the School of Physiology, thank you for your camaraderie, assistance, and the countless discussions that have enriched my understanding and kept me motivated. Your support has been a cornerstone of my research journey. To Lynne Pritchard, thank you for always being available to help with any request, and for doing so with a big smile.

A heartfelt thank you to my dear friends, brothers and brothers in Christ, who have been my pillars of support, offering motivation and understanding during the challenging times. Your belief in me and your prayers have been a constant source of strength.

To my personal person, my soulmate Eric, my heart in human form Tawananyasha and Tanaka Tiara, my sisters Mhamha Lecy and Maikuru Chenge, my Mum and Dad, and my family at large, you have been my backbone. Thank you for your unconditional love, sacrifices, and encouragement. Thank you for allowing me to vent my frustrations when things were not going on well. When I sometimes forgot to believe in myself, you always believed in me, cheered me up, and reminded me that all things are possible through God who strengthens us.

To everyone who has contributed in one way or another to the successful completion of this work, whether small or big, I am truly grateful.

Lastly, I am grateful to the Almighty God for the gift of life, His blessings, guidance, and for giving me the strength and perseverance to complete this thesis.

## Table of Contents

**DECLARATION** ..... i
**DEDICATION** ..... ii
**ABSTRACT** ..... iii
**ACKNOWLEDGEMENTS** ..... vi
**PRESENTATIONS ARISING FROM THIS PROJECET** ..... xii
**Chapter 1: Introduction** ..... 1
**1.1 Background** ..... 1
**1.2 Problem Statement** ..... 2
**1.3 Significance of Study** ..... 4
**1.4 Aims of the Study** ..... 5
**Chapter 2: Literature review** ..... 6
**2.1 Use of statistical analysis in research** ..... 6
**2.2 Background of null hypothesis significant testing, its use and misconception** ..... 7
**2.3 Background of P value use, misuse and misinterpretation** ..... 13
**2.4 The P-curve and its application in assessing research reliability** ..... 15
**2.5 The impact of publication bias on P-curve analysis** ..... 17
**2.6 The impact of P-hacking on P-curve analysis** ..... 19
**2.7 Prevalence of P-hacking across academic disciplines** ..... 22
**2.8 American Statistical Association (ASA) statement on P-values** ..... 23
**2.9 Reporting of P-Values** ..... 26
**Chapter 3: Methodology** ..... 28
**3.1 Data Acquisition** ..... 28
**3.2 Inclusion criteria** ..... 28
**3.3 Exclusion criteria** ..... 28
**3.4 Data Extraction Tool** ..... 29
**3.5 Data Extraction** ..... 29
**3.6 Data Analysis** ..... 31
**3.7 Ethics Clearance** ..... 32
**Chapter 4: Results** ..... 33
**4.1 Study selection** ..... 33
**4.2 Distribution of all articles** ..... 35
**4.3 Outcome analysis** ..... 37
**Chapter 5: Discussion** ..... 58
**References** ..... 67
**Appendix I: Data capturing tool** ..... 77
**Appendix II : Ethics Clearance** ..... 82

**Appendix III: Frequency tables.....83**
**Appendix IV: List of articles in study.....89**

## LIST OF FIGURES

**Figure 2.1 : The effect of publication bias on the distribution of P-values around the significance threshold of 0.05 ..... 17**

**Figure 2.2 : The effect of P-hacking on the distribution of P-values in the range of significance. ..... 21**

**Fgure 4.1 : Flowchart of the selection process of the articles, illustrating the number of articles remaining after each step (See Appendix III, Table A1 for the count of the distribution of articles with a P-values linked to the primary hypothesis) ..... 34**

**Figure 4.2 : Frequency distribution of experimental research article types (477articles in 2009-2010 and 414 articles in 2019-2020). (See Appendix III, Table A2 for actual count of articles) ..... 38**

**.Figure 4.3 : Frequency distribution of significant ($P < 0.05$) and non-significant P-values ($P \geq 0.05$) in 738 experimental studies, stratified by research type (330 significant P-values, 64 non-significant P-values in 2009-2010; 280 significant P-values and 64 non-significant P-values in 2019-2020). (See Appendix III, Table A3 for the article counts) ..... 40**

**Figure 4.4 : Frequency distribution of the type of P-values (exact or approximate) in 738 experimental studies, stratified by research type (131 exact P-values, 263 approximate P-values in 2009-2010 and 146 exact P-values, 198 approximate P-values in in 2019-2020) (See Appendix III, Table A4 for the counts of approximate and exact P-values by article type) ..... 42**

**Figure 4.5 : Frequency distribution of approximate P-values per article type. In 2009-2010, 123 approximate P-values were recorded for basic science (animal), 52 basic science (humans), 8 basic science (other) and 80 for clinical science. In 2019-2020, 94 approximate P-values were recorded for basic science (animal), 29 basic science (humans), 13 basic science (other) and 62 for clinical science. ..... 44**

**Figure 4.6 : Distribution of all exact P-values by article type (131 in 2009-2010 and 146 in 2019-2020. In 2009-2010, 31 articles were recorded for basic science (animal), 38 basic science (humans), 2 basic science (other) and 60 for clinical science. In 2019-2020, 31 articles were recorded for basic science (animal), 27 basic science (humans), 8 basic science (other) and 80 for clinical science. ..... 48**

**Figure 4.7 : Distribution of exact P-values between [0, 0.06) ..... 49**

**Figure 4.8 : P-curve illustrating the frequency distribution of reported exact P-values between [0, 0.06) for 2009-2010 and 2019-2020 period (solid red line). The dotted lines represent simulated P-curves for a null effect, and when statistical power is 60% and 90% to detect a true difference. ..... 51**

## LIST OF TABLES

**Table 2.1 : The prevalence of P-hacking in literature.....23**

**Table 2.2 : Summary of the statement on P-values(Wasserstein & Lazar, 2016).....25**

**Table 3.1 : Classification of experimental research articles.....30**

**Table 4.1 : Distribution of all articles by type and year .....36**

**Table 4.2 : Comparison of significant and non-significant P-values across 2009-2010 and 2019-2020 periods for each article type .....41**

**Table 4.3 : Comparison of approximate and exact P-values for each article type between 2009-2010 and 2019-2020 periods .....42**

**Table 4.4 : Comparison of the distribution of approximate P-values between 2009-2010 and 2019-2020 for each article type .....45**

**Table 4.5 : Frequency distribution of P-values.....52**

**Table 4.6 : Binomial test for 2009-2010 frequency distribution of reported P-values (test for P-hacking).....54**

**Table 4.7 : Binomial test for 2019-2020 frequency distribution of reported P-values (test for P-hacking).....57**

# PRESENTATIONS ARISING FROM THIS PROJECET

1. I presented preliminary data of this project at the 2023 PainSA Congress is held from May 5 to 7, 2023 in Gauteng, South Africa, and was awarded first prize for oral presentations
2. I presented preliminary data of this project at the 4th Annual Brain Function Research Group Research Day on September 23, University of the Witwatersrand, South Africa

# Chapter 1: Introduction

## 1.1 Background

Pain research aims to understand the mechanisms, treatments, and impact of pain on individuals. Experimental studies, ones where an investigator manipulates one or more independent variables and measures changes in an outcome variable, form an important part of the research methods available to biomedical researchers. In this setting, scientists rely on statistical tests to assist in the interpretation of quantitative outcomes, with the goal being to minimise the risk of falsely claiming a manipulation had an effect or did not have an effect. At a simplistic level, the P-value is a probability that measures the strength of evidence against the null hypothesis of no effect of the manipulation, with the smaller the p-value, the greater the evidence against the null hypothesis. More technically, the P-value is the probability of achieving the observed difference, or a greater difference, assuming the null hypothesis is true (Whitley and Ball, 2002; Greenland *et al.*, 2016). While P-values are an important part of interpreting research findings, unfortunately, they are frequently used as the sole output measure from statistical tests to gauge whether an experimental manipulation had a real effect.

P-values range from 0 to 1, with values close to 0 indicating that the null hypothesis of no effect was "wrong" and that the experimental manipulation did have an effect on the outcome measure, and values close to 1 suggesting that the evidence is in support of the null hypothesis (Hung *et al.*, 1997). But the question is, how small must a P-value be before one is comfortable rejecting that the data are consistent with the null hypothesis of no effect of the manipulation?

Null hypothesis significant testing (NHST) is a fundamental statistical procedure used to make decisions about experimental outcomes based on P-values, and it is the most widely used data

analysis method in scientific disciplines (Goodman, 1999). While the P-value is a continuous measure, NHST uses a P-value cutoff value, usually 0.05, to make decisions on whether the data are consistent with the null hypothesis. If the P-value is < 0.05 scientists typically reject the null hypothesis and say the results are statistically significant (often incorrectly interpreted as meaning the result is important) or nonsignificant (often incorrectly interpreted as the result being unimportant).

The use of P-values in statistical inference has increased significantly over time, with the percentage of PubMed abstracts reporting P-values more than doubling from 8% in 1990 to 17% in 2015, and tripling in meta-analyses (Chavalarias *et al.*, 2016). However, a low P-value does not guarantee against false positives (saying that there is an effect of manipulation when the data are consistent with the null hypothesis; Type I statistical error) nor does it indicate the biological or clinical relevance of a result. Indeed, overemphasis on small P-values and arbitrary P-value cutoff values, such as 0.05, has led to the mistaken belief that significant results are more valuable and reliable as P-values fall below the 5% cutoff level, leading to some researcher engaging in questionable research practices when it comes to statistical analysis of their data in order to cross the $P < 0.05$ cutoff, thus "proving" that their experimental manipulation had an effect (Fanelli, 2010a; Head *et al.*, 2015).

## 1.2 Problem Statement

Despite their extensive use, the reliability of P-values and null hypothesis significant testing has been questioned, especially since the reproducibility issue in science – the guarantee that a significant P-value meant that a reliable, and hence reproducible, result had been achieved – became apparent (Goodman, 1992; Ioannidis, 2008; Simmons, Nelson and Simonsohn, 2011; Greenland *et al.*, 2016). This debate about P-values spans multiple scientific disciplines and

focuses on the misuse of statistical inference, often due to misunderstandings of fundamental statistical concepts or, in some cases, intentional deception (Ioannidis, 2005; Simmons et al., 2011; Wasserstein & Lazar, 2016). Such misuse leads to erroneous data processing and misinterpretation of research findings, thus hindering scientific progress. Indeed, the *Journal of Basic and Applied Social Psychology* proposed banning P-values and hypothesis testing (Trafimow and Marks, 2015), and to curtail inappropriate uses or interpretation of P-values, and in 2016, the American Statistical Association issued a statement on P-value interpretation, emphasizing that meeting significance thresholds does not imply a higher probability of true hypotheses (Wasserstein & Lazar, 2016).

The misinterpretation and misuse of P-values has had a major unintended consequence, which also hinders scientific progress, namely, that journals are more likely to publish important findings. Unfortunately, the measure of importance has in many cases been hijacked by low, statistically significant, P-values instead of the importance of the scientific question and the rigor of the investigators in addressing the question, irrespective of the P-value. Indeed, Ioannidis (2019) found that 96% of reported P-values in biomedical literature are $\leq 0.05$, which is concerning as it suggests that there is a bias favoring more "important" results as measured by smaller P-values (Franco, Malhotra and Simonovits, 2014; Ioannidis *et al.*, 2014)..

This bias provides a favorable environment for questionable research practices. Questionable research practices, or P-hacking, involve selectively picking data or using different methods to achieve statistically significant results, and hence "important" results (Simmons, Nelson and Simonsohn, 2011; Bishop and Thompson, 2016; Teixeira, 2018). Examples of P-hacking include excluding "outliers", stopping data collection when significant results are reached, and selective reporting of P-values that fall below the threshold for statistical significance. Questionable research practices can distort the distribution of P-values, leading to over-

representation just below the conventional P = 0.05 threshold for statistical significance (Head *et al.*, 2015) (Brodeur *et al.*, 2016).

## 1.3 Significance of Study

Most information we have on questionable research practices and publication bias comes from studies in the field of psychology. There are only a few reports in the literature on the existence of these biases in pain research, a field of research that includes multiple scientific domains, from single cell preparations through to human psychology studies. In preclinical studies, Soliman *et al.* (2020) and Sena *et al.* (2010) highlight how bias inflates treatment effects. NeuPSIG's review of 229 trials further revealed a 10% overstatement of efficacy due to publication bias (*Finnerup et al.,* 2015), reinforcing the need for transparent reporting and inclusion of unpublished data. However the question of publication bias and questionable research practices is very relevant to pain clinical practice and patient well-being because statistical validity is fundamental to ethical clinical research (Emanuel,*et.al.,* 2000). Scientifically unsound research not only breaches ethical standard (Council for International Organizations of Medical Sciences, 2009), but it can also have detrimental effects on human health (Fanelli, 2010). Questionable research practices and publication bias could affect the conclusions drawn from pain studies. For example, publication of mainly positive results could lead to an overestimation of the efficacy of new treatments or new devices and the underestimation of their adverse effects (McGauran *et al.*, 2010). Reporting bias in clinical trials will also misinform researchers, clinicians and policymakers, which may affect patients directly and indirectly by affecting evidence-based medicine. For example, a meta-analysis of randomized controlled trials of pharmacological agents for managing neuropathic pain reported that the effect size of treatments is over-estimated by approximately 10% (Finnerup

*et al.*, 2015). Reduced publication of negative results means that the information does not reach the greater science community, thus influencing the information available for interpretation.

## 1.4 Aims of the Study

The main objective of my study was to assess the integrity of statistical reporting in the journal *Pain* by reviewing P-value reporting in the journal to establish how P-values were being presented, if there is evidence of good reporting practices, and whether there were any questionable research practices. The assessment will focus on original experimental research articles published in *Pain* in two periods 2009-2010 and 2019-2020. *Pain* was chosen as the target journal based on it being the premier specialist pain journal (based on journal impact factor) for both periods of assessment. To the knowledge of my colleagues and I, no analyses have been done on the reporting and distribution of P-values in any specialist pain journals.

I investigated:

i) The distribution of all the article types published during the study period

ii) The distribution of the experimental research articles

iii) The distribution of significant and non-significant P-values

iv) The quality of the null hypothesis significant testing

   a. The reporting of the P value in experimental research articles (approximate vs exact)

   b. The reporting of test statistics with P-values

   c. The reporting of sample size with P-values

v) The distribution of all exact P-values

vi) the presence of possible p-hacking using P-curves

# Chapter 2: Literature review

## 2.1 Use of statistical analysis in research

Statistical reasoning is a fundamental part of our daily lives. For example, it is found in everyday activities such as political pooling and government policies shaped by census data (American Statistical Association, 2005). Indeed, consistent with these uses, the word, "statistic" was coined by Gottfried Achenwall in the mid-eighteenth century to describe data about the state (Ostasiewicz, 2014). It was only in the late 19$^{th}$ century that statistics became fully integrated with probability theory and a mathematical discipline in its own right (Kennedy-Shaffer, 2019). The prevalence of statistics in modern society owes much to its role in the research process, where data are crucial for answering questions of interest. Researchers observe phenomena, formulate questions, collect data, and analyse the data. After analysing the data, researchers can develop conclusions in response to their research questions. (Schwab *et al.*, 2011; Agresti *et al*., 2018).

Researchers must comprehensively understand the concepts and reasoning underlying the techniques used in statistical analysis and inferential methods (Hoekstra *et al.*, 2014). Failing to do so can threaten the vital objective of science, which is to validate knowledge. One inferential method widely used in research is the statistical hypothesis test (Goodman, 1999; Hoekstra *et al*., 2006). Statistical hypothesis testing enables researchers to draw inferences about a population based on the data collected from a sample. The context in which the hypothesis test is applied plays a crucial role in informing the procedure and interpretation of results (Supino *et.al.,* 2012).

## 2.2 Background of null hypothesis significant testing, its use and misconception

Null hypothesis significance testing is a fundamental statistical procedure widely used across scientific disciplines to make inferences and guide decision-making based on study data (Goodman, 1999; Gigerenzer, 2004; Hoekstra, Finch and Johnson, 2006; Mezones-Holguín *et al.*, 2024). As one of the most employed data analysis techniques, null hypothesis significant testing involves testing a hypothesized factor against a null hypothesis, which posits that there is no effect or relationship between the variables under investigation (Pernet, 2017). The null hypothesis suggests that any observed differences or relationships are due to random variation, whereas the alternative hypothesis proposes that there is a true association between the factors being studied. Null Hypothesis significant testing is used to draw conclusions about populations based on sample data and to estimate measures of association, The null hypothesis significant testing evaluates hypotheses by calculating a P-value for a given population (Mezones-Holguín *et al.*, 2024). By using empirical data, null hypothesis significant testing allows researchers to assess whether the observed outcomes are consistent with the null hypothesis or if they provide sufficient evidence to reject the null hypothesis. The purpose of the null hypothesis significant testing is to reject or accept a null hypothesis. The concept of statistical significance evolved from the concepts of Ronald Fisher and Jerzy Neyman (Perezgonzalez, 2015).

Ronald Fisher first proposed the significance testing method, which involved calculating the probability (P-value) to measure strength of evidence against the null hypothesis (Johnstone, 1987). Fisher's approach was flexible and exploratory, allowing for conclusions to be drawn after multiple experiments and interpretation of the information according to the accumulated knowledge on the research question (Perezgonzalez, 2015). He emphasized the interpretation

of the P-value as a continuous measure, without explicitly defining error rates (the probability of making false positive or false negative decisions).

Neyman and Pearson then developed the hypothesis testing framework, which involves comparing two hypotheses (null and alternative) and making a decision based on pre-defined error rates (Type I and Type II errors; see below for explanation) (Perezgonzalez, 2015; Mertens and Recker, 2020). Their approach was more rigid and decision oriented. In the Neyman-Pearson approach, researchers set a null hypothesis ($H_0$) of no effect and an alternative hypothesis ($H_1$) of an effect. They also decide on an alpha level ($\alpha$), the acceptable false-positive rate. The Neyman-Pearson approach helps measure Type I and Type II errors and statistical power (Levine *et al*., 2008). Type I error, also known as false positives, is when researchers mistakenly reject the null hypothesis when it is true. Type II error, also known as false negatives, is when researchers fail to reject the null hypothesis when the alternative hypothesis is true. The alpha level ($\alpha$) reflects how much risk researchers are willing to take in making Type I errors. This alpha level is reflected in the traditional cut-off value for a calculated P-value of 0.05 (Lakens, 2022). An $\alpha$ of 0.05 means that if a researcher calculates a P-value of less than $\alpha$, they can decide to reject the null hypothesis. However, the cost of using a threshold ($\alpha$) is that there is, in this example, a 0.05 probability that the researcher rejects the null hypothesis, but the data are in fact consistent with the null hypothesis and the researcher is merely seeing a rare event. Researchers can choose to set the $\alpha$ level and then use a lower value like 0.01 or 0.001 for stricter control of false positives. However, lowering the $\alpha$ level reduces statistical power, which is the probability of correctly rejecting the null hypothesis (Elliss, 2010). As the $\alpha$ level decreases from 0.05 to 0.01, it becomes more "difficult" to reject the null hypothesis and so statistical power decreases. A study with low power may then fail to reject

the null hypothesis when the alternative is valid, or in other words, increases the chance of making a Type II error, where researchers miss a true effect.

Discussions about the proper use of null hypothesis testing in scientific studies date back to the proposals made by Ronald Fisher for significance testing. Various accounts have been given of how this technique evolved as an empirical inference method (Pernet, 2017; Szucs and Ioannidis, 2017). Several reviews and analyses of the null hypothesis significance testing properties have also been published (Amrhein et al., 2019; McShane et al., 2019; Wasserstein & Lazar, 2016).

Null hypothesis significance testing used by research community members, such as educators, authors, reviewers, publishers, and editors, has become ritualistic and established, making it hard to break. This can threaten the quality of research. Some practices commonly used with null hypothesis significance testing, such as interpreting the size of the p-value, have been prone to misunderstandings and misuse (Dixon, 2003; Cohen, 1994).

The misinterpretations of the null hypothesis significance testing cause problems in scientific practice, threatening research efficiency, validity, reproducibility, and robustness (Ioannidis, 2005; Simmons et al., 2011; Wasserstein & Lazar, 2016).

- Research efficiency: This refers to conducting scientific studies in a manner that maximizes the knowledge gained while minimizing wasted resources such as time, money, and effort.
- Validity: This pertains to the accuracy or truthfulness of the conclusions drawn from a study. A valid study accurately measures what it intends to measure and produces trustworthy results.

- Reproducibility: This is the ability to obtain a similar result when a study is repeated under the same conditions. Reproducibility is crucial in scientific practice as it ensures that findings are not random or specific to a single study's circumstances.
- Robustness: This refers to the consistency and reliability of results across different conditions or methodologies. A robust scientific finding should remain consistent under various assumptions, settings, or analytical approaches.

The use of null hypothesis significance testing and its role in decision-making, including healthcare research, has been controversial since its inception (Gigerenzer, 2004; Schwab *et al.*, 2011; Greenland *et al.*, 2016; Sedgwick *et al.*, 2022). Some longstanding challenges to research caused by using null hypothesis significance testing, which have remained relevant over time, and new challenges that have become prominent in recent years will be discussed below.

1. The null hypothesis significance testing is difficult to understand and often misinterpreted. Null hypothesis significance testing uses a P-value, a complex statistical measure that summarizes how well a data set fits a proposed model. The most common use of null hypothesis significance testing involves constructing a null hypothesis and an alternate under a set of assumptions. Null hypothesis significance testing shows whether the results of a tested effect are not reliable based on whether a test statistic is significant by using a P-value threshold (Masicampo and Lalande, 2012; Leek and Pen, 2015; Sanfilippo *et al.*, 2016). Key terms like "statistical significance" and "P-value" are often misunderstood (Greenland *et al.*, 2016; Amrhein *et al*., 2019; Cohen, 1994). For example, the P-value does not show how strong or big an effect is (Gelman and Carlin, 2014). Misinterpreting the P-value can lead to incorrect assumptions about the strength, reliability, and probability of the effect under study. The belief that a small P-value indicates a high likelihood of replicating

results is inaccurate (Ioannidis, 2005; Gelman and Carlin, 2014; Ronald L. Wasserstein and Lazar, 2016). In studies, a low P-value can suggest a misfit of the null hypothesis, but does not support a specific alternative hypothesis (Miller, 2009). Incorrectly interpreting null hypothesis significance testing can lead researchers to ignore evidence that does not reach statistical significance or to give too much importance to evidence that does reach statistical significance. Interventions or treatments based on misinterpreted evidence might not work or be harmful (McShane *et al.*, 2019). Additionally, false findings may be published, spreading unverified theoretical claims.

2. Null hypothesis significance testing is sensitive to sampling strategy and sample size. Null hypothesis significance testing relies on an appropriate sampling strategy, emphasizing random sampling to draw valid conclusions about a population. Suppose samples are not chosen randomly or are not representative of the population; the conclusions drawn from this testing are invalid because it is impossible to correct for sampling bias, which statistical significance testing assumes is non-existent (Leahey, 2005). Large sample sizes, which increase the probability of statistically significant findings, may lead to accepting trivial differences as significant (Sedgwick *et al.*, 2022). Sample size sensitivity is particularly problematic when dealing with hypotheses that assume no effect or difference, which are common in social sciences, where directional statements are prevalent (Mertens and Recker, 2020). A researcher who gathers a large enough sample has increased chance of falsely rejecting the null hypothesis since as sample size increases, precision of the estimates increase for any effect size, so the confidence interval around the null effect decreases, increasing the probability of rejecting the null hypothesis (Schwab *et al.*, 2011). Thus, reliance on null hypothesis significance testing with large sample sizes can result in wrong inferences (O'Brien, 2007) and biased estimations of effects, particularly in

observational studies where controlling for all relevant variables is challenging (Bruns and Ioannidis, 2016). Therefore, null hypothesis significance testing guarantees that any difference, whether small or irrelevant, will be statistically significant if the sample size is large enough (Sedgwick et al., 2022). For the opposite reason, smaller sample sizes have less statistical power for a given effect size. Indeed, for a statistically significant conclusion to be drawn for a small, under-powered, study, the effect size must be large. When conducting studies, it is important to consider the sample size to ensure that the results obtained genuinely reflect the population under study ( Sedgwick et al., 2022).

3. Null hypothesis significance testing is susceptible to questionable research practices that blur the ethical lines of research (Bedeian *et al*., 2010; Butler *et al*., 2017). Questionable research practices can be caused by factors such as pressure to publish significant results and publication bias (Chittaranjan Andrade, 2021). Changes in the academic environment, the emphasis on scholarly performance metrics, and the regulations on research impact measurement have pressured academics to publish 'significant' results. This phenomenon is known as publish or perish/ publish under pressure (Davies and Felappi, 2017). This pressure arises from the need to meet expectations for promotions (van Dalen, 2021) and to demonstrate research impact (Davies and Felappi, 2017). The adoption of these practices is often understated, but evidence shows that they are prevalent in academic settings (Bedeian, Taylor and Miller, 2010). These practices include p-hacking and HARKing (Hypothesizing After Results are Known). P-hacking involves extensive data manipulations or calculations until a statistically significant result is found (Simmons *et al*., 2011; Bishop and Thompson, 2016; Chittaranjan, 2021). HARKing is a questionable research practice where researchers formulate hypotheses based on the data they have already observed, and then present these hypotheses as if they were established before the

study began (Simmons, Nelson and Simonsohn, 2011; Chittaranjan Andrade, 2021). This practice can lead to biased results and undermine the reproducibility and validity of scientific findings Fanelli, D. (2010)

There have been frequent calls for a statistics reform regarding the use of null hypothesis significance testing in decision-making, including the removal of the concept of statistical significance (Sedgwick et al., 2022). Nonetheless, the binary approach to decision-making is convenient, and its use has remained pervasive. Regardless of the future of statistical significance, there are calls for greater focus on the magnitude and accuracy of the observed effects and their clinical importance (Sedgwick et al., 2022).

## 2.3 Background of P value use, misuse and misinterpretation

Researchers need to measure the strength of the evidence they observe and minimise the risk that a result occurred by chance (Whitley and Ball, 2002). To do the latter, most researchers use the P-value, which has been a cornerstone in statistical hypothesis testing for over a century, serving as a measure of evidence against a null hypothesis. Since their introduction, P-values have become widely used in scientific research across various disciplines for making statistical inferences from experimental data (Ioannidis, 2019). The percentage of abstracts reporting P-values for all PubMed abstracts more than doubled from 8% in 1990 to 17% in 2015, and the rate tripled for meta-analysis (Chavalarias *et al.*, 2016).

Being a probability, P can take any value from 0 to 1 and is interpreted based on their proximity to 0 or 1 (Hung *et al.*, 1997). P-values close to 0 suggest that the observed differences in the data are unlikely to result by chance indicating, by the Fisherian concept, strong evidence

against the null hypothesis, whereas P-values close to 1 suggest there are no differences between groups of data beyond random variation (Hung *et al.*, 1997). P-values are therefore used to determine statistical significance. Statistically significant or positive findings are those that are less than the cut-off value (usually set at P = 0.05) and they show the statistical incompatibility of data with the null hypothesis. Researchers can interpret this incompatibility as casting doubt on or providing evidence against the null hypothesis or the underlying assumptions and can conclude that it supports the alternate hypothesis (Lu & Belitskaya-Levy, 2015; Wasserstein & Lazar, 2016). Research findings with P-values equal to or larger than 0.05 are considered non-significant or negative. Non-significant results mean that there is little evidence to support the alternate hypothesis, although it is not ruled out as being false. Importantly, P-values larger than the threshold also do not mean that the null hypothesis is true because the hypothesis being tested assumes that the null is true. Rather P-values above the threshold indicate that there is not enough evidence to reject the null hypothesis (Wasserstein & Lazar, 2016). A large P-value simply means that the observed data is not sufficiently inconsistent with the null hypothesis to warrant its rejection. While the P-value undoubtedly provides useful information on probabilistic significance, a low P-value does not guarantee a positive result, nor does it indicate the biological or clinical relevance of a result (Sedgwick et al., 2022; Wasserstein & Lazar, 2016). The P-value alone does not provide information regarding the size of the treatment effect. Instead, P-values should be used together with other statistical measures such as confidence intervals, which indicate both the direction and precision of the measured treatment effect. By providing a range of plausible effects sizes, confidence intervals enable an assessment of the clinical significance of a treatment effect, supplementing the evaluation of its statistical significance (Sedgwick, 2015).

Most fields in biomedical, life, social, and physical sciences report their results using P-values to show statistical significance (Ioannidis, 2019). Despite the extensive use of P-values, some researchers have started to question the usefulness and reliability of using P-values, null hypothesis significance testing , and their interpretation as a measure of evidence (Goodman, 1992; Ioannidis, 2008; Simmons, Nelson and Simonsohn, 2011; Greenland *et al.*, 2016). Indeed the *Journal of Basic and Applied Social Psychology* has proposed a ban on the use of any P-values and hypothesis testing from articles being published in the journal (Trafimow and Marks, 2015). They declared null hypothesis significance testing was "invalid" and "We believe that the $p < 0.05$ bar is too easy to pass and sometimes serves as an excuse for lower quality research"

The publication of statistically significant results is linked to the successful track records of many researchers whereas results that fail to confirm the findings of previous studies are seen to be of less scientific value and generate less interest from publishers (Dirnagl and Lauritzen, 2010). As a result of the immense pressure on researchers to publish and increase citation scores coupled with an overemphasis on the importance of statistical significance as a condition for publication, the P-value has become vulnerable to misapplication and misinterpretation (Dirnagl & Lauritzen, 2010; Ioannidis, 2005; Reynolds & Waldrop, 2024; Wasserstein & Lazar, 2016).

## 2.4 The P-curve and its application in assessing research reliability

A p-curve, which represents the distribution of statistically significant P-values from a set of studies can be used to assess the reliability of published research (Head *et al.*, 2015). According to Simonson and colleagues, the p-curve offers a way to differentiate between true effects and

selective reporting (Simonsohn, et.al., 2014). By examining the distribution of P-values, particularly those between 0 and 0.05, researchers can determine the evidential value of findings, helping them to conclude whether the effect size is nonzero. Under the null hypothesis, where no effect exists, a uniform distribution of P-values is expected, suggesting that each P-value is equally likely to occur as per Figure 1 below (Black line, Fig. 1A ) (Head *et al.*, 2015; Hartgerink *et al.*, 2016). However, as shown by a black line in Figure 1B, when a true effect exist, the P-curve becomes right-skewed P, with a higher concentration of P-values closer to zero (Head *et al.*, 2015). A right-skewed P-curve suggests a collection of independent findings with decreasing P-values, providing stronger evidence for the presence of an effect (Head *et al.*, 2015; Hartgerink *et al.*, 2016). The degree of skewness increases with the statistical power of the analysis, with higher power leading to steeper p-curves.

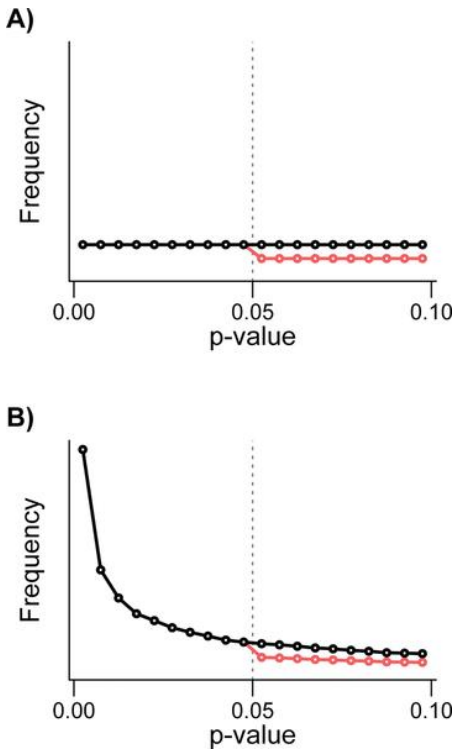



**Figure 2.1: The effect of publication bias on the distribution of P-values around the significance threshold of 0.05**

**.**A) Black line shows distribution of P-values when there is no evidential value, and the red line shows how publication bias influences this distribution. B) Black line shows distribution of P-values when there is evidential value, and red line shows how publication bias influences this distribution.

Source: https://doi.org/10.1371/journal.pbio.1002106.g001

## 2.5 The impact of publication bias on P-curve analysis

One of the primary uses of the p-curve is to identify potential publication bias by plotting the distribution of P-values around the commonly accepted significance threshold of 0.05. Publication bias refers to selective publication of large amounts of statistically significant outcomes, which means that articles reporting significant results are more likely to be published than those reporting negative results (Franco *et al.,* 2014; Ioannidis *et al.*, 2014). The concept of the publication bias, is also known as the file drawer effect, occurs when scientific work that exceeds the 5% level ends up not being submitted for publication, ending up being discarded

in the metaphorical “file drawer” where they are forgotten about. Publication bias causes a huge problem in research since it implies that there is a greater inclination to publish only significant results leading to a skewed representation of evidence in the scientific literature (Head *et al.*, 2015). Despite the importance of publishing all the results in research, many research organizations and industry entities fail to do so. In pain research, more specifically clinical trials, failure to publish null results could lead to adverse effects for patients since it distort the perceived efficacy of interventions, treatments, and diagnostic tests, ultimately impacting clinical and policy decisions and also violating ethical (Munafò and Neill, 2016; Brassington, 2017). A meta-analysis on therapeutic trials of selective serotonin reuptake inhibitors found that the efficacy of selective serotonin reuptake inhibitors was overestimated by up to 115% when only considering published studies (Treanor *et al.*, 2020).

Positive results are more likely to be published, cited and accepted, especially by high-ranking journals, and negative results are disappearing from many disciplines (Fanelli, 2010a; Duyx *et al.*, 2017). Treanor *et al.*, 2020 found that conference abstracts with positive conclusions were more likely to be published as full-text articles compared to those with neutral or negative conclusions. However, Okike *et al.* (2008) found no evidence of publication bias in the review of manuscripts for publication by *The Journal of Bone & Joint Surgery* (American Volume) but they showed concern over the dearth of non-positive (i.e., not statistically significant) studies, attributing this to researcher-driven publication bias. The unwillingness of researchers to publish negative results and the bias or demand from journals to publish statistically significant results can lead researchers into wilfully or unintentionally manipulating analyses to achieve statistically significant results. Because of publication bias, P-values may provide an inaccurate measure of strength of the evidence against the null hypothesis and leading to questionable conclusions being reached. A decrease in the P-values above 0.05 on a P-curve

(as shown by the red line) can be concluded as an indication for publication bias (Fig 1A and Fig 1B) (Head et al., 2015; Masicampo & Lalande, 2012). However, this pattern alone does not differentiate between selective publication bias and p-hacking.(Gelman and Carlin, 2014; Simonsohn, et.al 2014).

## 2.6 The impact of P-hacking on P-curve analysis

Questionable research practices, also known as P-hacking, occur when non-significant results are manipulated to appear significant, producing an overabundance of P-values just below the 0.05 threshold. Ideally, researchers must make decisions on how data will be collected, when to stop data collection, the methods to be used in data analysis as well as the data that need to be included in the analysis before the beginning of a study. If such decisions are made during data analysis it leads to potential bias and P-hacking. P-hacking, is a result of researchers selectively picking data or using different methods to analyse the data, until a statistically significant result is obtained (Simmons *et al.,* 2011; Bishop and Thompson, 2016; Teixeira, 2018). Some examples of questionable research practices include; (1) exclusion of outliers, (2) checking results and stopping data collection when a significant result is reached (data peeking), (3) analysing data extensively without a specific hypothesis in mind (data dredging) (4) testing multiple hypothesis , (5) failing to report null results, and (6) selective reporting of P-values (cherry picking) (Stefan and Schönbrodt, 2023). However, in some cases, researchers inadvertently exercise the "researcher degrees of freedom" and carry out multiple data analyses, not realising that without a prespecified analysis plan, every analysis decision may have an influence on whether statistical significance is achieved, with no intention of P-hacking (Gelman and Carlin, 2014).

One possible manifestation of P-hacking is over-representation of P-values just below the 0.05 threshold thus increasing the probability of false positives, leading to the publications of results that indicate an effect exists when there is no effect of an experimental intervention (Head *et al.*, 2015). False positives from P-hacking can mislead data-driven decisions in critical areas such as healthcare and economics. The failure to replicate significant findings is also a major consequence of P-hacking. In the article "Estimating the reproducibility of psychological science" by the Open Science Collaboration, of the 97% of original studies that had significant results ($P < .05$), only 36% of replications could be replicated. P-hacking distorts the scientific record by emphasizing findings that appear significant but are statistically or methodologically unsound. Researchers use questionable research practices to manipulate the reporting of P-values so that they just cross-over the significance cut-off, hence the higher frequency of just significant P-values. P-hacking distorts the shape of the p-curve, particularly near the significance cutoff, resulting in a misleading representation of the data. As a result, the P-curve will exhibit an unusual clustering of P-values just below 0.05. This distortion alters the shape of the P-curve, making it less reliable as an indicator of true effects. The P-curve will shift from flat to left skewed, when there is no true effect as shown by the red line in Figure 2A. With true effect, P-hacking results in an over representation of P-values in the tail of the curve just below 0.05 as shown by the red line in Figure 2B (Head *et al.*, 2015). Ultimately, the exact shape of the P-curve is determined by the true effect size and the extent of the P-hacking, with steeper and more pronounced right skewness suggesting stronger evidence for a true effect

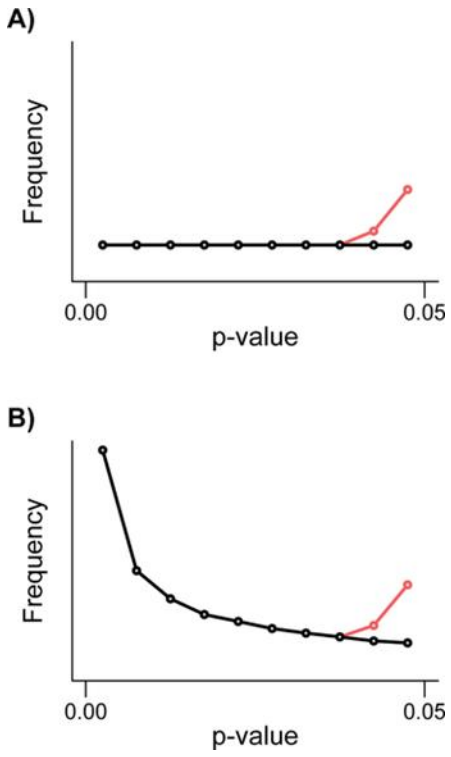



**Figure 2.2: The effect of P-hacking on the distribution of P-values in the range of significance.**

A) Black line shows distribution of P-values when there is no evidential value, and the red line shows how P-hacking influences this distribution. B) Black line shows distribution of P-values when there is evidential value, and the red line shows how P-hacking influences this distribution.

Source: https://doi.org/10.1371/journal.pbio.1002106.g002

However, the use of the P-curve to detect publication bias and P-hacking has been criticised by some authors. Dorothy V.M. Bishop and Paul A. Thompson (2016) argue that the shape of the P-curve depends on whether dependent variables are intercorrelated. For uncorrelated variables, p-hacked data do not show the "p-hacking bump" just below .05. However, there is a negative skew when variables are intercorrelated. The authors argue that p-curve analysis requires careful selection of P-values and considerable control over the data to be effective. Alternative methods like z-curve analysis have been recommended because they address these

shortcomings of the P-curve by converting P-values into z-scores and providing a clearer picture of P-hacking and selection bias (Simmons, *et al.*, 2014). Z-curve also provides measures like the expected replication rate (ERR) and false positive risk, offering more robust insights into the reliability of findings (Simmons, *et al.*, 2014)

## 2.7 Prevalence of P-hacking across academic disciplines

While publication bias removes true or false negatives from literature, P-hacking bring true or false positives to literature (Gelman and Carlin, 2014; Head *et al.*, 2015; Ronald L Wasserstein and Lazar, 2016). Nevertheless, both questionable research practices and publication bias lead to the computation of P-values which provide an inaccurate strength of evidence against the null hypothesis (Hartgerink et al., 2016; Masicampo & Lalande, 2012). Table 2 provides an overview of the prevalence of P-hacking in the literature. This table summarizes findings from various studies on the frequency of P-hacking across different disciplines, shedding light on its widespread impact and potential implications for research integrity.

**Table 2.1: The prevalence of P-hacking in literature**

| Study | Field | Objective | Key finding |
|---|---|---|---|
| Simmons, Nelson and Simonsohn, 2011 | Psychology | To demonstrate how researchers can manipulate data to achieve significant results. | Showed that P-hacking is common and coined the term "P-hacking". Highlighted the importance of preregistration and replication. |
| Head *et al.*, 2015 | General Science | To investigate the prevalence and impact of P-hacking using text-mining techniques. | Found that P-hacking is widespread, but its effect is relatively weak compared to real effect sizes. P-hacking does not drastically alter scientific consensuses drawn from meta-analyses. |
| Bin Abd Razak *et al.*, 2016 | Orthopaedic Literature | To investigate if the phenomenon of P-hacking was evident in in three top orthopaedic journals in 2015. | Frequency plots for all three journals revealed evidence of P-hacking. Binomial tests for all three journals were significant for evidence of evidential value However, the binomial test* for P-hacking was significant only for one journal |
| Berman *et al.*, 2018 | Marketing | To explore P-hacking in A/B testing. | Showed how companies might manipulate data to present favorable outcomes. |
| Baum and Bromiley, 2019 | Management | To examine P-hacking in top-tier management journals. | Revealed instances of P-hacking, compromising the integrity of published findings. |
| Adda *et al.*, 2020 | Clinical Trials | To investigate P-hacking in clinical trials and how economic incentives shape the distribution of results. | Found no widespread P-hacking but identified selective reporting by small industry sponsors. Highlighted the role of economic incentives. |
| Rooprai *et al.*, 2023 | Imaging research | To investigate P-hacking in Imaging research | Found no evidence of P-hacking in abstracts published in over 100 imaging journals since 1972 |

*A binomial test is a statistical method used to determine if there is a significant difference in proportions between an observed sample and a hypothesized sample. Bin Abd Razak et al. (2016) ran binomial tests by analyzing the frequency of reported P-values in studies from three orthopaedic journals to determine whether the proportion of significant results ($p < 0.05$) was greater than expected by chance, indicating evidential value or potential P-hacking

## 2.8 American Statistical Association (ASA) statement on P-values

Despite the general criticism due to its misuse, the P-value continues to be used in many scientific papers (Pondrom, 2016). However, some experts argue that while P-values are valid statistical measures, the problem lies in overreliance on P-values by the scientific community, and misinterpretation of P-values by investigators (Ioannidis, 2005; Pondrom, 2016; Reynolds

and Waldrop, 2024). In response to the misuse and misunderstanding of P values in scientific literature, the American Statistical Association issued a statement clarifying that P values should not replace scientific reasoning (Ronald L Wasserstein and Lazar, 2016). They emphasize the importance of good statistical practice and the need for proper analysis and reporting. According to the statement, good statistical practice should involve using multiple methods and techniques, including creating a comprehensive understanding of the data presented and interpreting results in a given context. The statement also emphasized the importance of well-designed statistical arguments, which are more than just a number, and whether the number exceeds an arbitrary, in this case, $P < 0.05$. The American Statistical Association should be used to guide research and ensure that all the results are published. Table 1 below summarises six principles underpinning the appropriate use, the interpretations of P-values, and the decisions made on the P-values.

**Table 2.2: Summary of the statement on P-values(Wasserstein & Lazar, 2016)**

| Principle | Use | Misuse | Misconception |
|---|---|---|---|
| P-values can indicate how incompatible the data are with a specified statistical model | P-values help determine whether observed data deviate from expectations under a specific model | Interpreting a P-value as the probability that the null hypothesis is true, or that the data were produced by random chance alone | Believing that a P-value can provide a definite measure of evidence regarding a hypothesis |
| P-values do not measure the probability that the studied hypothesis is true or the probability that the data were produced by random chance alone | P-values show the compatibility of data with a specific hypothesis/model | Treating P-values as direct evidence to determine the degree to which the hypothesis is true | Assuming that a low P-value implies a high probability that the alternative hypothesis is true |
| Scientific conclusions and business or policy decisions should not be based only on whether a P-value passes a specific threshold | P-values should be one of many pieces of evidence considered in decision-making | Solely relying on P-values to make critical decisions without considering other statistical factors. With large test sets, the P-values will always be small or even small enough to indicate statistical significance | Believing that crossing a P-value threshold (i.e. 0.05) is sufficient for significance |
| Proper inference requires full reporting and transparency | Report all relevant statistical details, including P-values, effect sizes, and confidence intervals to allow proper interpretation and reproducibility of work | Selectively reporting P-values that support a desired outcome while ignoring others | Thinking that P-values alone can provide a complete picture of the data analysis |
| A p-value, or statistical significance, does not measure the size of an effect or the importance of a result | Recognize that P-values indicate whether an effect exists, not its magnitude or practical significance | Interpreting a significant P-value as indicating a large or important effect | Assuming that statistical significance implies practical or scientific significance |
| By itself, a P-value does not provide a good measure of evidence regarding a model or hypothesis | Use p-values in conjunction with other statistical measures like confidence, credibility, or prediction intervals and Bayesian methods to get a more comprehensive understanding of data | Relying on P-values as the sole measure of evidence for or against a hypothesis | Believing that a P-value can substitute for comprehensive statistical analysis and scientific judgment |

These principles, however, do not necessarily discourage the use of P-values in empirical research or devalue them, but advocate the reporting of test results more transparently, not just with P-values themselves. According to the American Statistical Association, the P-value became a gatekeeper in some fields that prevent work from being published over time.

According to Jessica Utts (Wasserstein & Lazar, 2016), the president of the American Statistical Association in 2016 the editorial bias in scientific literature has led to the "file-drawer effect," where studies with statistically significant outcomes are more likely to be published than those with non-significant results. Additionally, overemphasis on small P-values over other scientific and statistical reasoning as main indicators of successful research have resulted in questionable research practices such as "data dredging" and "P-hacking,"(Fanelli, 2010b).

## 2.9 Reporting of P-Values

The reporting of P-values in academic literature is a critical aspect of statistical analysis and hypothesis testing. P-values help researchers determine the significance of their results, indicating whether the observed effects are likely due to chance or represent true relationships. Proper reporting of P-values ensures transparency, reproducibility, and accurate interpretation of research findings. The article "Standardized Reporting of Statistical Results in APA Format: Enhancing Clarity, Transparency, and Reproducibility in Research" by Sathyanarayana S and T. Mohanasundaram (2025) emphasizes the importance of accurately reporting statistical results to ensure clarity and reproducibility in research. According to the American Psychological Association (APA) it is essential to follow the guidelines discussed below when reporting P-values.

- P-values larger than 0.01 should be reported to two decimal places (e.g., P = 0.17).
- P-values between 0.01 and 0.001 should be reported to three decimal places (e.g., P = 0.004).
- P-values less than 0.001 should be reported as P < 0.001

Researchers should report exact P-values rather than using approximate P-values (e.g., p <.05). Exact P-values enhance transparency and allow readers to better understand the statistical

significance of the result. Additionally, it is important to report the test statistic along with the P-value to provide complete information to the reader. For example, "A Chi-Square Test was used to assess if there was a significant difference between the two groups. There was a significant difference between the two groups, $X^2 = 0.145$, df (1) P = 0.004".

To enhance the reporting of P-values, researchers should also include effect sizes, confidence intervals, and the null hypothesis. This provides a more comprehensive understanding of the results and their practical significance. accurate and transparent reporting of P-values is essential for the integrity of scientific research. By adhering to standardized guidelines and providing additional statistical information, researchers can ensure their findings are clear, reproducible, and meaningful.

Our study therefore aimed to evaluate the integrity of statistical reporting in the journal *Pain* by examining P-value reporting in original experimental research articles from 2009-2010 and 2019-2020. It addressed concerns about the reliability of P-values and null hypothesis significance testing, which have been questioned due to issues of reproducibility and misuse across various disciplines. The investigation focused on the distribution of article types, significant and non-significant P-values, and the quality of null hypothesis significance testing, including the reporting of approximate versus exact P-values, test statistics, and sample sizes. Additionally, the study analyzed the distribution of exact P-values and potential P-hacking using P-curves, aiming to provide insights into the statistical reporting practices in this leading pain journal and highlight areas for improvement to enhance scientific rigor and transparency

# Chapter 3: Methodology

## 3.1 Data Acquisition

All publications published in the journal *Pain* from 1 January 2009 to 31 December 2010, and 1 January 2019 to 31 December 2020 were eligible for inclusion. The journal *Pain* was chosen as the target journal based on it being the premier specialist pain journal (based on journal impact factor) for both periods of assessment. All articles for the review periods were downloaded from the journal website https://journals.lww.com/pain/Pages/issuelist.aspx?mode=cover. The articles then underwent a screening process based on predefined inclusion and exclusion criteria (see below).

## 3.2 Inclusion criteria

- Original experimental research papers with at least one identifiable primary hypothesis/objective, and at least one p-value associated with that hypothesis/objective.

## 3.3 Exclusion criteria

- Original experimental research papers that did not have clearly identifiable primary hypotheses/objectives and/or at least one p-value associated with the primary hypothesis/objective;
- Non-experimental studies (observational studies);
- Reviews, including systematic reviews and meta-analyses;
- Editorials;
- Commentaries;
- Clinical notes;
- Letters to the editor;

- Errata/corrigenda;
- Obituaries;
- Book reviews.

Systematic reviews, meta-analyses, editorials and all the other articles which did not meet our inclusion criteria were excluded in our study because they synthesize existing research findings rather than generate original experimental results, which means they do not involve primary hypothesis testing or direct P-value reporting in the same way original experimental studies do.

## 3.4 Data Extraction Tool

A standardized data extraction form was used for data capture (Appendix I). The form was developed based on the preliminary assessment of studies in the journal such that the first thirty articles from years 2009 and 2019 were used to inform design of the data capture form. The form was operationalized using the Research Electronic Data Capture (REDCap) application, a web-based electronic data capturing application. The data extraction form captured the following information:

- **Article Details**: Author name(s), article title, year of publication, and volume, issue, page numbers, and DOI.
- **Type of study:** The type of study based on predefined groups (see below).
- **Study Objectives:** Whether there was a primary hypothesis/objective.
- **Statistical Information:** Whether there was a P-value associated with the primary hypothesis/objective, and the format of the P-value (see below).

## 3.5 Data Extraction

All articles in the review periods were screened. Bibliometric data and type of article classification (e.g., narrative reviews, editorials, and letters to the editor, experimental research,

systematic reviews) was recorded for every article in the four years considered in the review. No further extraction was done on all the other articles except the original experimental research papers. Experimental research papers were then classified into sub-categories as per Table 2.1. Only articles fitting the classification of being a research paper underwent further review and had data on P-values extracted.

**Table 3.1: Classification of experimental research articles**

| Category | Description |
|---|---|
| Basic Science Human/Animals | Publications with research on a particular topic focussing on knowledge creation in science without any particular application or use in view. |
| Basic Science (other) | Publications with research done outside a living organism using bacteria, animal and human cells, or tissues in culture. |
| Clinical science | Publications with research that evaluates the effectiveness and safety of medication or medical devices by monitoring their effects in people. Reviews of patients' clinical data were also recorded under clinical science. |

The primary hypothesis for each reviewed article was found by perusing the abstract and introduction section of the manuscript. The results section of each article, including figures and tables, was read, and the P-value for the primary hypothesis recorded. If multiple objectives were specified, only the first objective was noted, and only first P-value corresponding to this objective was recorded. To be eligible for selection, P-values had to be specified with the letter 'P' or 'p', and the value described by an exact numeric value or a threshold value preceded by an inequality sign. General phrases such as 'no change' or 'significant difference' were not accepted.

If the primary hypothesis was not clear, the article was noted accordingly, and no statistical data was recorded. Articles were also excluded if the primary hypothesis, even when clearly

indicated, did not require null hypothesis significance testing to be answered and hence no P-value could be expected. Primary extraction was performed by me, and then the extraction verified by colleagues (Felicia Siboza, Dhuriya Cajee, and Courtney Jeffery). Conflicts were resolved through discussions with the research assistants. If an agreement could not be reached, the issue was escalated to Peter Kamerman, for resolution. The complete reference list of the selected articles is attached at the end of this dissertation in Appendix IV.

## 3.6 Data Analysis

Data from REDCap was exported into an Excel spreadsheet. The R Statistical Computing Platform (v4.4.2) was used for data cleaning, data analysis, and plotting.

Descriptive statistics were used to present the results, and frequency distributions were used to describe the analysis done for 2009-2010 and 2019-2020 periods. A P-value of less than 0.05 was considered the threshold for statistical significance in our analysis. P-curves, which show the distribution of exact P-values was analyzed to assess evidential value and detect the presence of P-hacking. To validate the observed P-curves, we conducted a binomial test at 60% and 90% power for specific P-value bins ([0.03, 0.04), [0.04, 0.05), [0.05, 0.06)) using R packages. A binomial test is a statistical method which is used to determine if there is a significant difference in proportions between an observed sample and a hypothesized sample. A significant P-value with a positive difference in probability indicates that there is a significant increase in observed frequencies from expected frequencies suggesting potential P-hacking or selective reporting.

## 3.7 Ethics Clearance

The articles for this study are made available to the public on the Pain website https://journals.lww.com/pain/Pages/issuelist.aspx?mode=cover. Therefore, an ethics waiver was granted from the University of Witwatersrand Human Research Ethics Committee (Medical), reference number W-CP-191108-1(Appendix II).

# Chapter 4: Results

## 4.1 Study selection

Figure 4.1 illustrates the study selection process. We extracted all articles published in the journal *Pain* from 2009 to 2010 and 2019 to 2020, yielding a total of 1,342 articles. The number of publications was 746 for 2009-2010 and 596 for 2019-2020. After excluding 451 non-experimental research articles (2009-2010: 269; 2019-2020: 182), 26 articles without a clear primary hypothesis (2009-2010: 9; 2019-2020: 17) and 127 articles without a P-value linked to the primary hypothesis (2009-2010: 74; 2019-2020: 53), 738 articles met our selection criteria. This study sample comprised of 394 articles from 2009-2010 and 344 from 2019-2020. Of the 738 articles with a P-value linked to the primary hypothesis, 277 articles recorded exact P-values, while 461 recorded approximate P-values: 146 exact P-values in 2009-2010, and131 exact P-values in 2019-2020. The exact P-values were used to construct the P-curve for each period, which were used to detect the presence of possible P-hacking.

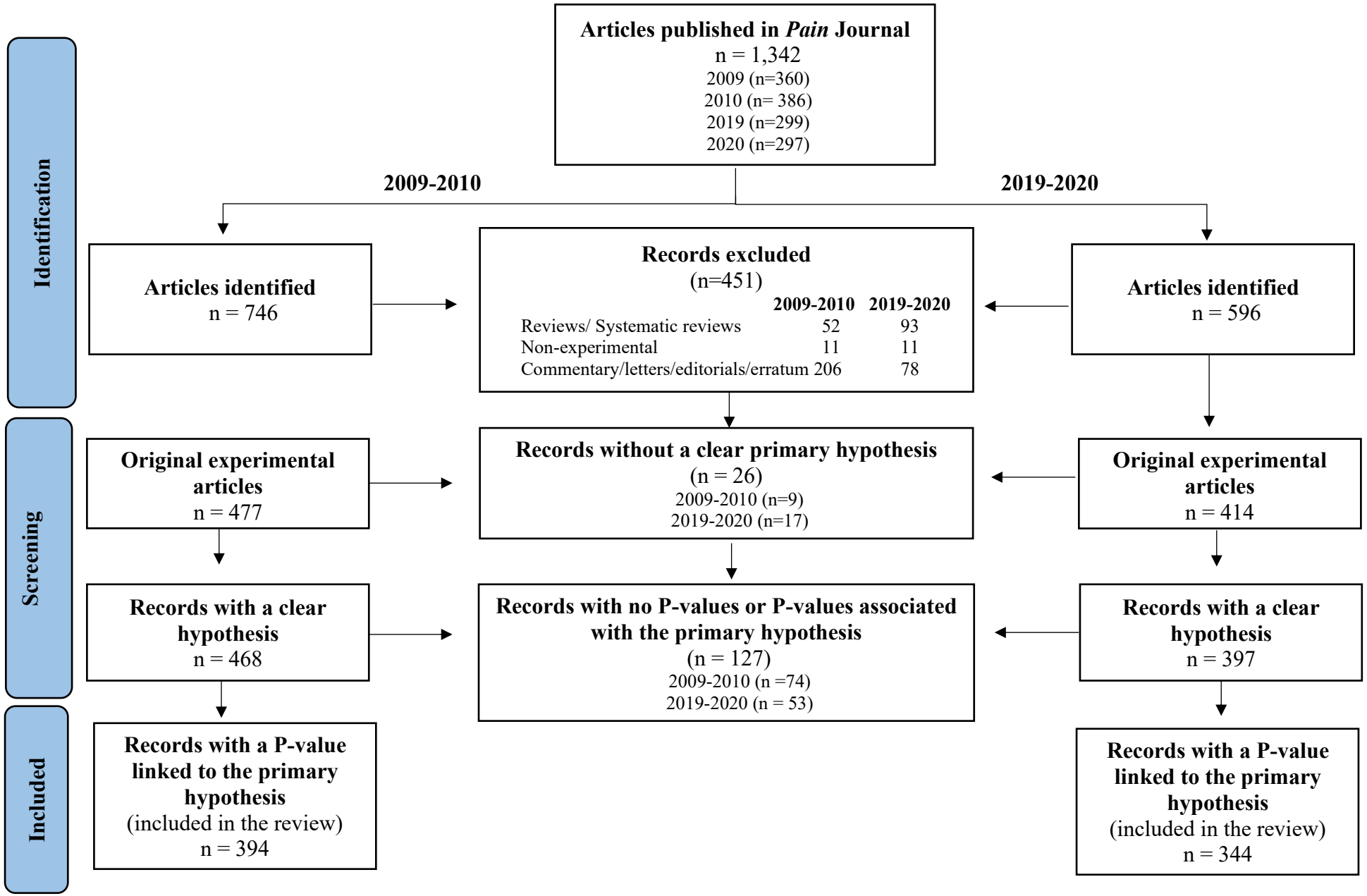



**Figure 4.1: Flowchart of the selection process of the articles, illustrating the number of articles remaining after each step (See Appendix III, Table A1 for the count of the distribution of articles with a P-values linked to the primary hypothesis)**

## 4.2 Distribution of all articles

The descriptive data and results for all 1,342 articles are presented in Table 4.1, categorizing the articles by type and year of publication (2009, 2010, 2019, and 2020). There was no systematic pattern of change in article type, expressed as a percentage of articles published each year. Clinical science articles constituted the largest proportion of all experimental research articles accounting for 26%. Clinical science also makes up the largest proportion of articles each year, peaking at 30% in 2019. However, in 2009 25% of the articles were basic science (animals) making it the largest proportion for that year. Basic science (animal) articles fluctuated slightly, with a low of 22% in 2010, peaking at 26% in 2019. Basic science (human) articles decreased from 18% in 2009 to 8.8% in 2020. Basic science (other) articles increased from 1.1% in 2009 to 5.4% in 2020. Commentaries, editorials, letters, and errata collectively comprised 21% of all articles, with a notable drop from 29% in 2009 to 12% in 2019, before slightly increasing to 14% in 2020. Narrative reviews increased from 4.7% in 2009 to 10% in 2020 with a low of 4.1% in 2010. Non-experimental (human) articles were 0.3% in 2009, 2.6% in 2010, and none were published in 2019, but 3.7% of published articles in 2020 were non-experimental (human) studies. Systematic reviews (animals) were very low, making up 0.3% of articles in both 2019 and 2020, with none in 2009 and 2010. Systematic reviews (human) increased from 2.5% in 2009 to 5.7% in 2020.

**Table 4.1: Distribution of all articles by type and year**

| Characteristic | Overall N = 1,342[1] | 2009 N = 360[1] | 2010 N = 386[1] | 2019 N = 299[1] | 2020 N = 297[1] |
|---|---|---|---|---|---|
| Article type | | | | | |
| Basic science (animal) | 320 (24%) | 90 (25%) | 84 (22%) | 77 (26%) | 69 (23%) |
| Basic science (human) | 184 (14%) | 63 (18%) | 53 (14%) | 42 (14%) | 26 (8.8%) |
| Basic science (other) | 43 (3.2%) | 4 (1.1%) | 12 (3.1%) | 11 (3.7%) | 16 (5.4%) |
| Clinical science | 344 (26%) | 70 (19%) | 101 (26%) | 89 (30%) | 84 (28%) |
| Commentary/editorial/letter/erratum | 284 (21%) | 106 (29%) | 100 (26%) | 36 (12%) | 42 (14%) |
| Narrative review | 91 (6.8%) | 17 (4.7%) | 16 (4.1%) | 27 (9.0%) | 31 (10%) |
| No experimental (human) | 22 (1.6%) | 1 (0.3%) | 10 (2.6%) | 0 (0%) | 11 (3.7%) |
| Systematic review (animal) | 2 (0.1%) | 0 (0%) | 0 (0%) | 1 (0.3%) | 1 (0.3%) |
| Systematic review (human) | 52 (3.9%) | 9 (2.5%) | 10 (2.6%) | 16 (5.4%) | 17 (5.7%) |

[1]n (%)

## 4.3 Outcome analysis

### **4.3.1** Distribution of experimental research article type

Figure 4.2 shows the distribution of experimental research article types within and across the two periods.

*Within period:* In 2009-2010, basic science (animal) had the highest number of articles (174), followed by clinical science (171). Basic science (human) had 116 articles and basic science (other) had the least number of articles at 16. In 2019-2020, clinical science had the highest number (173), followed by basic science (animal) with 146 articles. Basic science (human) 68 articles and basic science (other) had the least number with 27 articles.

*Across periods:* The number of articles in basic science (animal) decreased from 174 in the 2009-2010 to 146 in the 2019-2020, reflecting a reduction of 2% over the decade from 37% to 35%. Basic science (human) articles saw the biggest decline, dropping from 116 articles in 2009-2010 to 68 articles in 2019-2020, which corresponds to a reduction of 8% from 24% to 16%. In contrast, basic science (other) articles increased from 16 in 2009-2010 to 27 in 2019-2020, indicating a rise of 4% over the decade from 3% to 7%. Clinical science articles also increased with 171 articles in 2009-2010 and 173 articles in 2019-2020, showing an increase of 6% from 36% to 42%.

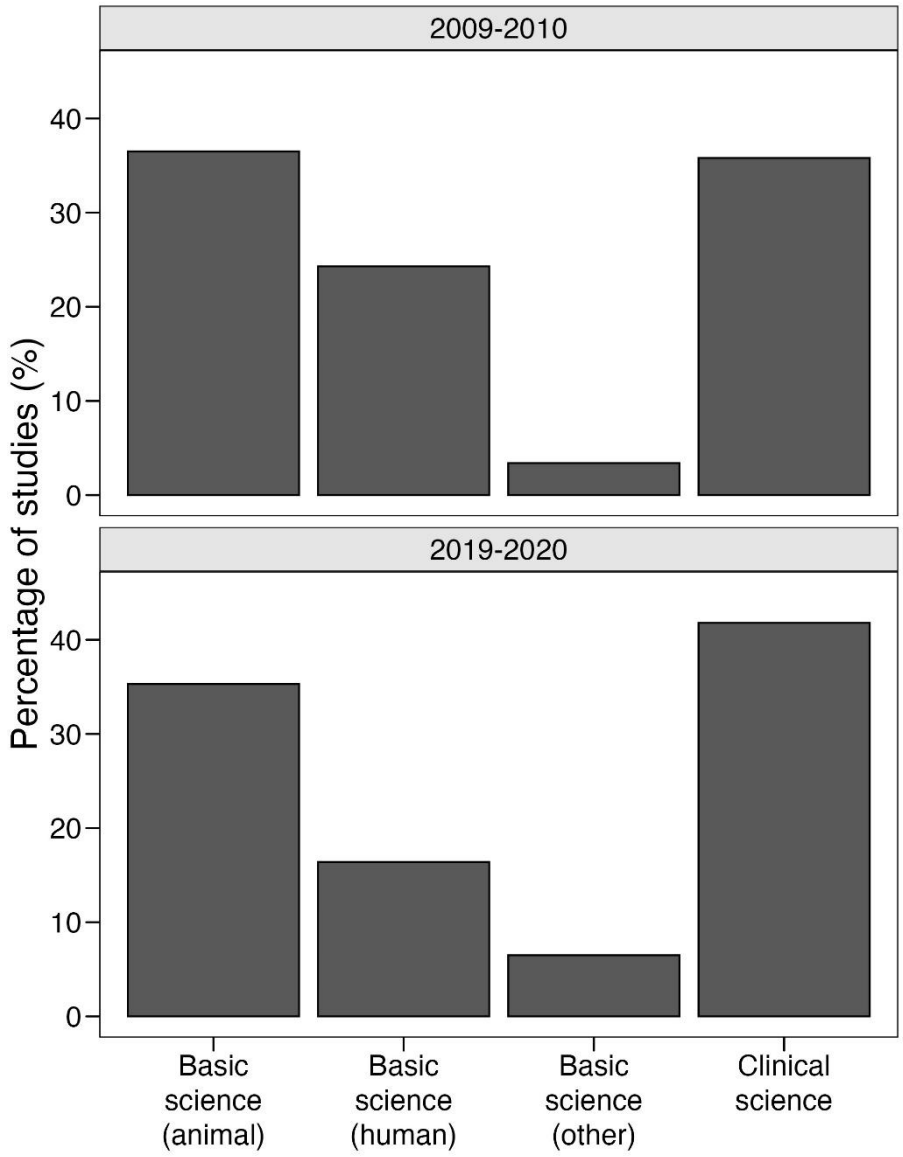



Figure 4.2: Frequency distribution of experimental research article types (477 articles in 2009-2010 and 414 articles in 2019-2020). (See Appendix III, Table A2 for actual count of articles)

### 4.3.2 Analysis of P-value reporting in experimental research articles

*4.3.2.1 Distribution of significant and non-significant P-values*

Figure 4.3 compares the number of articles reporting significant ($P < 0.05$) and non-significant ($P \geq 0.05$) P-values across different article types for 2009-2010 and 2019-2020. In 2009-2010, 84% of P-values were significant compared to 81% in 2019-2020. Basic science (animal)

maintained a high proportion of significant results, increasing from 87% in 2009-2010 to 90% in 2019-2020. Basic science (human) consistently reported 82% significant results during both periods. Basic science (other) showed an increase from 80% in 2009-2010 to 95% in 2019-2020. Clinical science studies showed a higher proportion of non-significant results in 2019-2020 (29%) compared to 2009-2010 (19%). Table 4.2 shows the statistical analysis of non-significant and significant P-values across both periods per article type. Chi-squared tests were used for basic science (human), basic science (animal), and clinical sciences groups. Our analysis revealed that none of the article types showed statistically significant differences after applying the Holm P-value adjustment for multiple comparisons. We used the Fisher's exact test for basic science (other) because our sample size was less than ten, and the results also support that there is no significant difference in this group of articles. The overall statistical analysis of non-significant and significant P-values across both periods did not show a significant association, $\chi^2 = 0.559$, df = 1, P = 0.455.

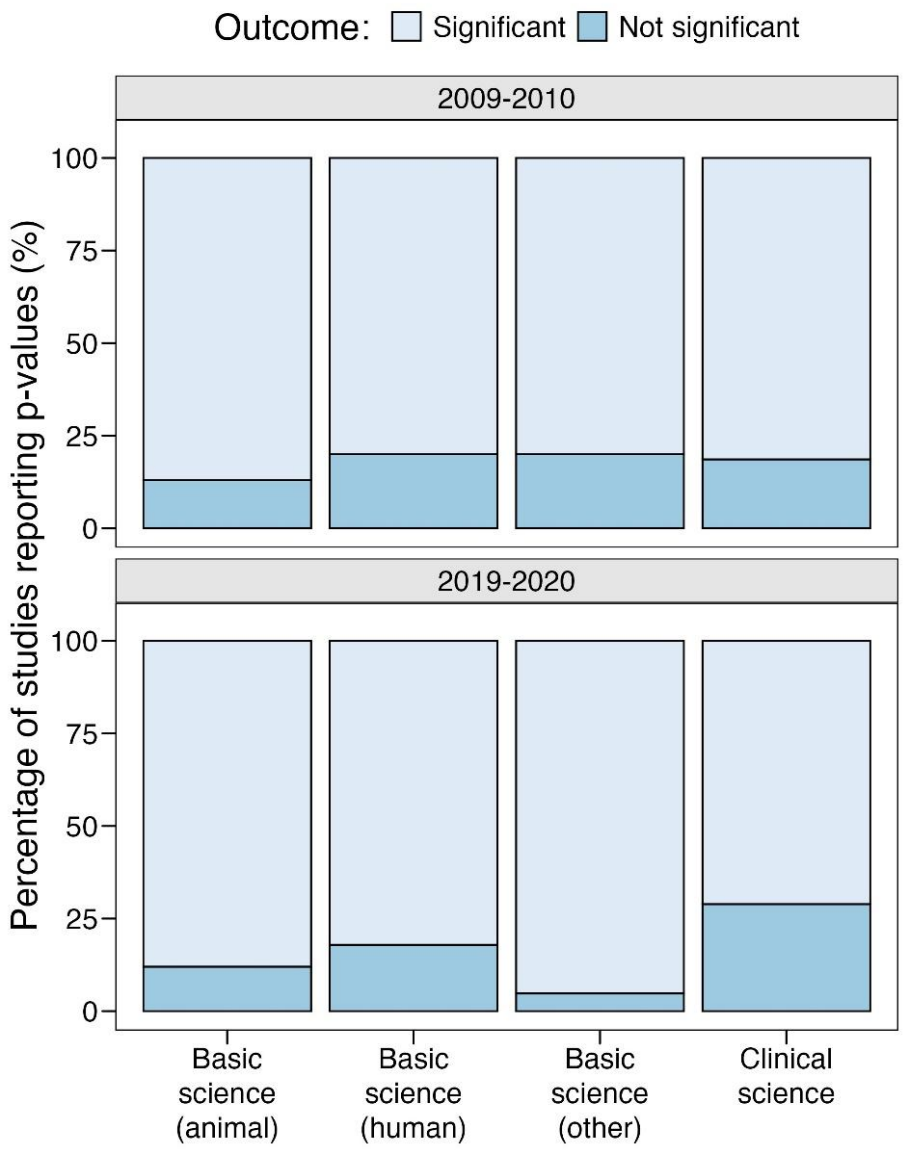



**Figure 4.3:** Frequency distribution of significant ($P < 0.05$) and non-significant P-values ($P \geq 0.05$) in 738 experimental studies, stratified by research type (330 significant P-values, 64 non-significant P-values in 2009-2010; 280 significant P-values and 64 non-significant P-values in 2019-2020). (See Appendix III, Table A3 for the article counts)

**Table 4.2:** Comparison of significant and non-significant P-values across 2009-2010 and 2019-2020 periods for each article type

| Article type | χ2 (df=1) | P-value | Holm-adjusted P-value |
|---|---|---|---|
| Basic science (animal) | 0.004 | 0.948 | ≈ 1 |
| Basic science (human) | 0.011 | 0.918 | ≈ 1 |
| Clinical science | 3.581 | 0.058 | 0.232 |
| Basic science (other) (Fisher's test) | | 0.237 | 0.711 |

*4.3.2.2 Distribution of the type of P-values*

Figure 4.4 shows the comparison of articles reporting approximate (P-values reported as an inequality) and exact P-values across different article types for 2009-2010 and 2019-2020. In this study, all P-values reported as $P < 0.001$ were treated as approximate P-values. There is a shift towards reporting exact P-values from 2009-2010 (33%) to in 2019-2020 (42%). The results indicated a significant shift towards reporting exact P-values compared to approximate P-values over the decade, $\chi^2 = 6.234$, df = 1, $P = 0.013$. Basic science (animal) showed a slight increase in exact P-values from 20% to 25% in 2019-2020. Basic science (human) increased from 42% to 48%. Basic science (other) increased in the proportion of exact P-values from 20% to 38%, and clinical science increased from 43% to 56%. The overall trend indicates a move towards the recommended method of reporting P-values as exact values. Table 4.3 compares the distribution of approximate and exact P-values between 2009-2010 and 2019-2020 for each article type. The P-values for all article types was not statistically significant after adjusting for multiple comparisons.

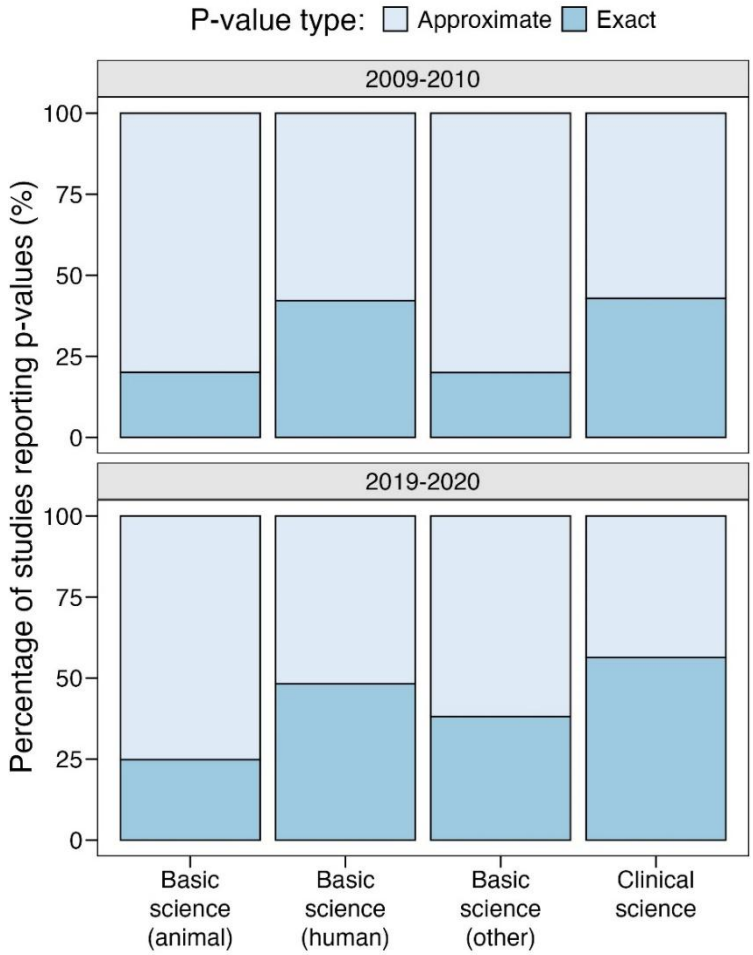



**Figure 4.4**: Frequency distribution of the type of P-values (exact or approximate) in 738 experimental studies, stratified by research type (131 exact P-values, 263 approximate P-values in 2009-2010 and 146 exact P-values, 198 approximate P-values in in 2019-2020) (See Appendix III, Table A4 for the counts of approximate and exact P-values by article type).

**Table 4.3:** Comparison of approximate and exact P-values for each article type between 2009-2010 and 2019-2020 periods

| **Article type** | **$\chi 2$** | **P-value (df=1)** | **Holm-adjusted P-value** |
|---|---|---|---|
| Basic science (animal) | 0.621 | 0.431 | ≈ 1 |
| Basic science (human) | 0.289 | 0.591 | ≈ 1 |
| Clinical science | 4.600 | 0.032 | 0.128 |

| Article type | $\chi 2$ | P-value (df=1) | Holm-adjusted P-value |
|---|---|---|---|
| Basic science (other) (Fisher's test) | | 0.429 | ≈ 1 |

*4.3.2.3 Distribution of approximate P-values (per article type)*

Figure 4.5 illustrates the prevalence of reported approximate P-values ($P < 0.001$, $P < 0.01$, $P < 0.05$, and $P > 0.05$) across various article categories during 2009-2010 and 2019-2020. There is a substantial increase in "highly" significant P-values ($P < 0.001$), rising from 39% in 2009-2010 to 55% in 2019-2020. This trend was observed across all article types. Specifically, in basic science (animal) research, the proportion of $P < 0.001$ values increased from 27% to 49%, while basic science (human) increased from 52% to 69%. Basic science (other) articles showed an increase from 38% to 46%, and clinical science articles increased from 50% to 60%. The observed general increase in "highly" significant results could suggest an improvement in the rigor and robustness of research findings over the examined period, or an increase in selective reporting of only "highly" statistically significant studies. Our analysis of the difference between each article types across the 2009-2010 and 2019-2020 periods showed mixed results in terms of statistical significance as shown in Table 4.3. For basic science (animal) the Fisher's exact test yielded a P-value, indicate that the observed increase from 27% in 2009-2010 to 49% in 2019-2020 of basic science (animal) articles showing highly significant p values ($p<0.001$) was statistically significant. The P-values for basic science (human), basic science (other) and clinical science were all not statistically significant (i.e., they did not fall below 0.05). We then compared the distribution of $P < 0.05$ to all the other P-values for basic science (animal) since our results showed that there is a statistically significant difference in the data between the two periods. The Chi-square test yielded a value of $\chi 2=8.601$with a P-value of 0.003, and the Holm-adjusted P-value was 0.012. This indicates that the observed differences are statistically significant even after adjusting for multiple comparisons. The significant difference in the distribution of $P < 0.05$ compared to all other P-values for basic science (animal) suggests that

there could have been notable changes in the research, for example change in methodologies, reporting standards and polices, influencing the outcomes of basic science (animal) studies over the decade.

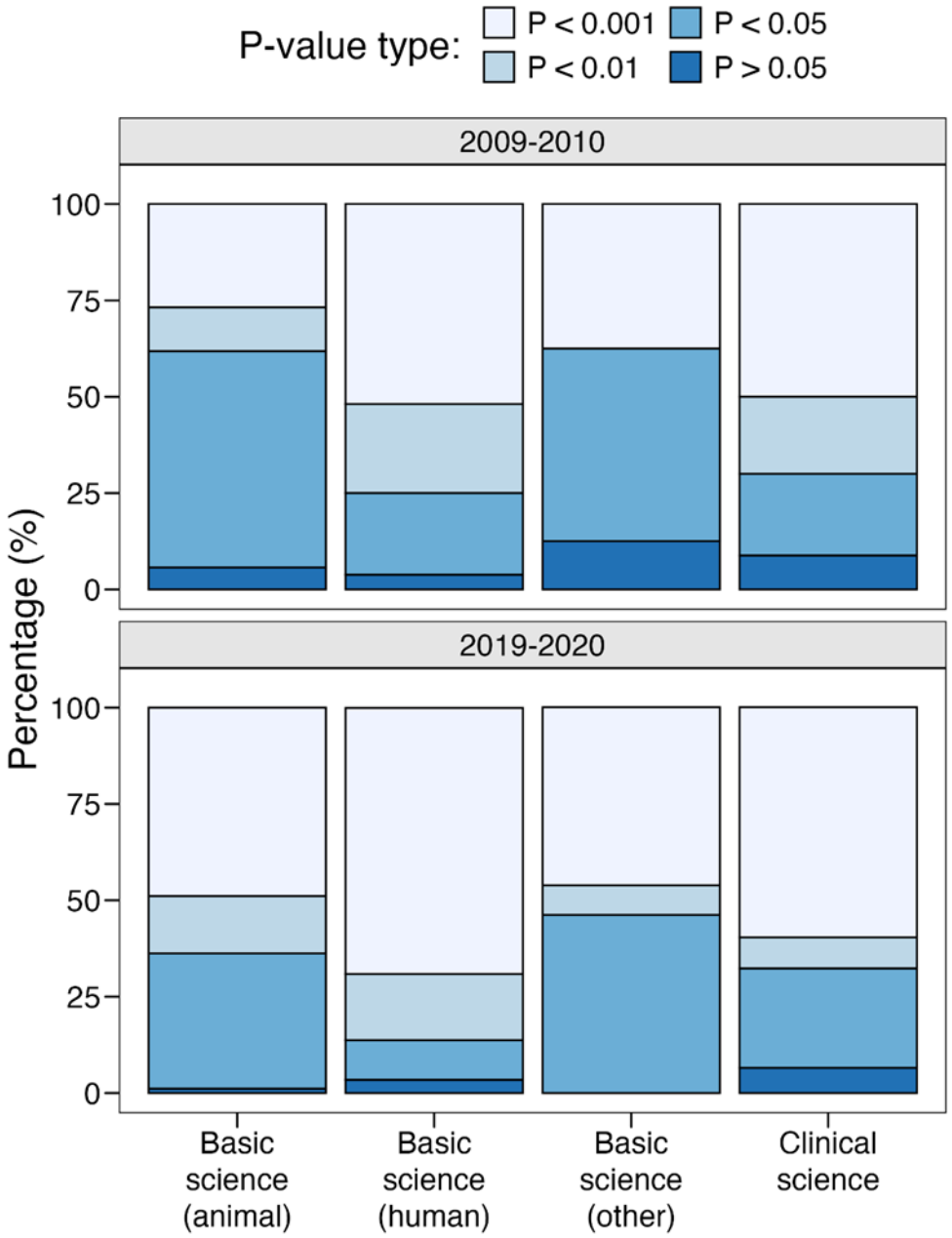



**Figure 4.5:** Frequency distribution of approximate P-values per article type. In 2009-2010, 123 approximate P-values were recorded for basic science (animal), 52 basic science (humans), 8 basic science (other) and 80 for clinical science. In 2019-2020, 94 approximate P-values were recorded for basic science (animal), 29 basic science (humans), 13 basic science (other) and 62 for clinical science. (See Appendix III, Table A5 for the article counts)

**Table 4.4:** Comparison of the distribution of approximate P-values between 2009-2010 and 2019-2020 for each article type

| Article type | P-value (Fisher's test) | Holm-adjusted P-value |
|---|---|---|
| Basic science (animal) | 0.001 | 0.005 |
| Basic science (human) | 0.477 | 0.954 |
| Basic science (other) | 0.766 | 0.954 |
| Clinical science | 0.207 | 0.621 |

*4.3.2.4 Reporting of test statistics with P-values*

In the 2009-2010 period, out of 131 exact P-values included in our study, 45% had a corresponding test statistic, while 32.7% of approximate P-values has a corresponding test statistic reported. A Chi-squared test was performed to determine if there was a significant association between reporting an exact or approximate P-value and including a corresponding test statistic. The results indicated a significant association, $\chi^2 = 5.206$, df = 1, P = 0.023. For the 2019-2020 period, 45.2% of the 146 exact P-values had a test statistic, while 33.3% of 198 approximate P-values has a corresponding test statistic reported. The Chi-squared test results again showed a significant association, $\chi^2 = 4.519$, df = 1, P = 0.0034.

A comparative analysis of 2009-2010 versus 2019-2020 showed that out of 131 exact P-values 2009-2010 included in our study, 45% the P-value had a corresponding test statistic, while 45.2% exact P-values in 2019-2010 had a corresponding test statistic, A Chi-squared test was performed to determine if there was a significant association between reporting an exact P-value in a specific period and including a corresponding test statistic. The results indicated a non-significant association, $\chi^2 = 0.000.$, df = 1, P ≈ 1. For the approximate P-values in 2009-2010, 32.7 % of the 262 approximate P-values had a corresponding test statistic, while 33.3 %

of 198 approximate P-values in 2019-2020. The Chi-squared test results again showed a non-significant association, $\chi^2 = 0.002.$, df = 1, P = 0.966.

These analysis shows that that while the type of P-value is associated with the likelihood of reporting a test statistic, this reporting practice did not significantly change over time.

*4.3.2.5 Reporting of sample size with P-values*

In the 2009-2010 period, out of 131 exact P-values included in our study, 98.5% had a corresponding sample size, while 87.8% of 262 approximate P-values has a corresponding sample size reported. A Chi-squared test was performed to determine if there was a significant association between reporting an exact or approximate P-value and including a corresponding sample size. The results indicated a significant association, $\chi^2 = 11.243$, df = 1, P < 001. For the 2019-2020 period, 96.6 % of the 146 exact P-values had a test statistic, while 93.9% of 198 approximate P-values has a corresponding sample size reported. The Chi-squared test results again showed a non-significant association, $\chi^2 = 0.745$, df = 1, P = 0.388.

A comparative analysis of 2009-2010 versus 2019-2020 showed that out of 131 exact P-values 2009-2010 included in our study, 98.5% had a corresponding sample size, while 96.6% of 146 exact P-values in 2019-2010 had a corresponding sample size. A Fisher's exact test was performed to determine if there was a significant association between reporting an exact P-value in a specific period and including a corresponding sample size. The results indicated a significant association, with P = 0.039. For the approximate P-values in 2009-2010, 87.8% of the 262 approximate P-values had a sample size, while 96.6% of 198 approximate P-values in 2019-2020. The Chi-squared test results again showed a significant association, $\chi^2 = 4.250.$, df = 1, P = 0.039. Our results show an increase in the reporting of samples sizes for the P-values in both period with approximate P-values showing the higher increase. The Chi-squared results for the association between an approximate P-value and a corresponding sample size is

significant showing a change in reporting of statistics over time. There was consistent in the reporting of exact P-values and a corresponding sample size between the two period as shown by the insignificant P value obtained.

*4.3.2.6 Distribution of exact P-values by article type*

The percentage of articles reporting exact P-values across different article types for the 2009-2010 and 2019-2020 periods is shown in Figure 4.6. In the 2009-2010 period, 33% of articles reported exact P-values, which increased to 42% in the 2019-2020 period, indicating a general trend towards more precise P-value reporting. The number of articles in basic science (animal) remained constant at 20% in both periods, basic science (human) saw an increase from 42% in 2009-2010 to 48% in 2019-2020. In basic science (other), the number of articles increased from 20% to 38%, and from 43% to 56% in clinical science.

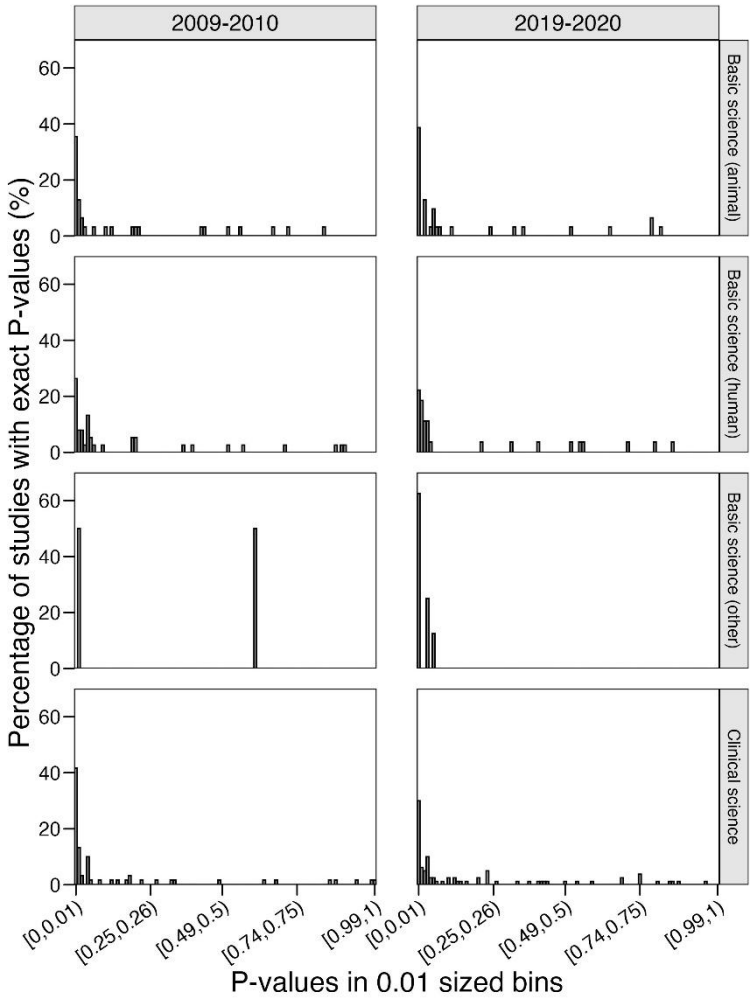



**Figure 4.6 :** Distribution of all exact P-values by article type (131 in 2009-2010 and 146 in 2019-2020. In 2009-2010, 31 articles were recorded for basic science (animal), 38 basic science (humans), 2 basic science (other) and 60 for clinical science. In 2019-2020, 31 articles were recorded for basic science (animal), 27 basic science (humans), 8 basic science (other) and 80 for clinical science**.** (See Appendix III, Table A6 for the article counts)

*4.3.2.7 Distribution of all exact P-values by period*

Figure 4.7 shows the histogram of all exact P-values for 2009-2010 and 2019-2020. The shape of the distribution is similar across both periods with a large proportion of studies reporting small P-values. The insert in each panel shows the distribution of the exact P-values between [0, 0.06). The extremely small P-values (lower than 0.001) were dominant in both periods. The proportion of studies reporting P-values just under 0.05 ([0.04, 0.05)) is higher (10%) in 2009-2010 period compared to 2.5% in 2019-2020.

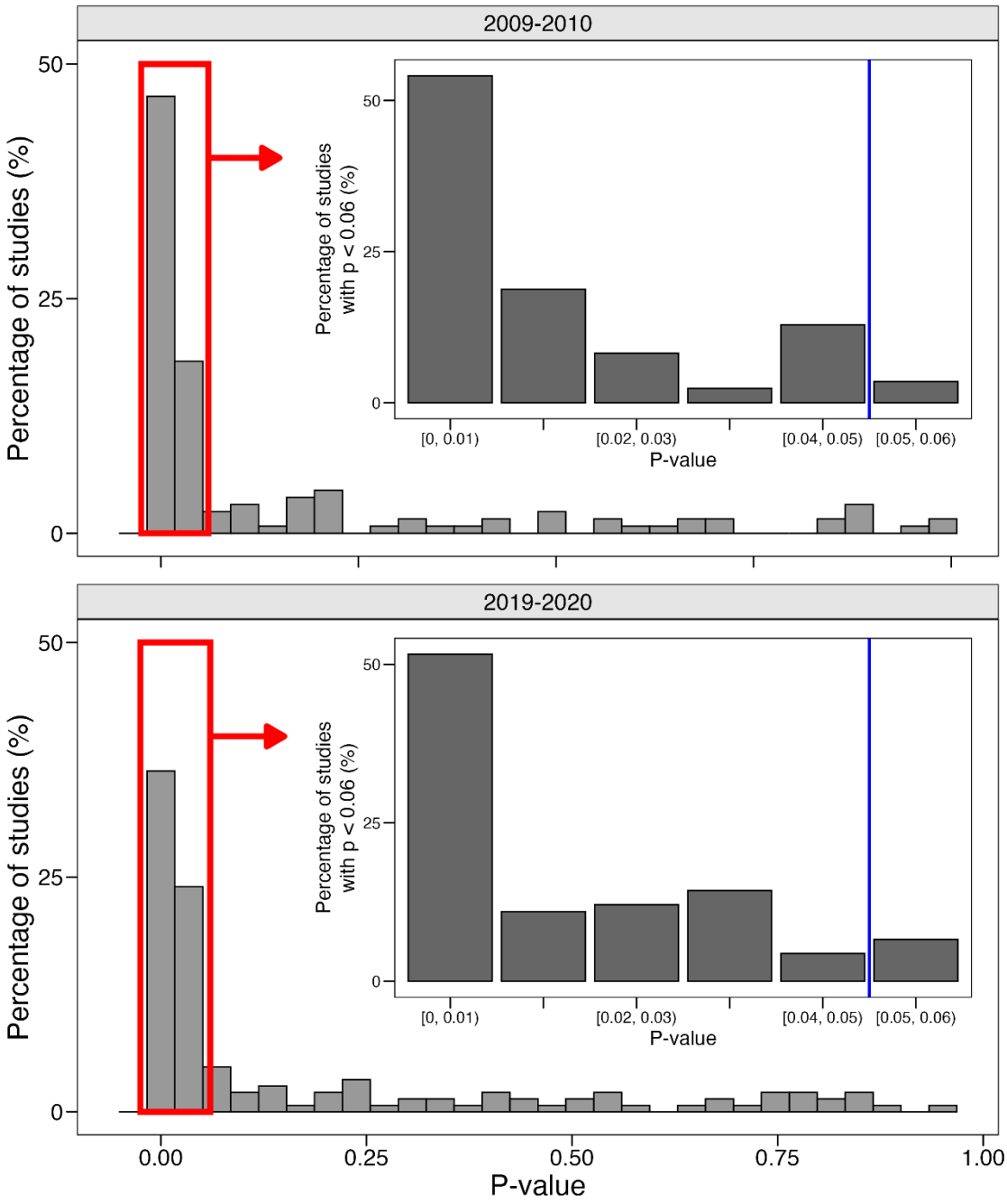



**Figure 4.7 :** Distribution of all exact P-values by period. (Insert) Distribution of exact P-values between [0, 0.06)

The distribution of exact P-values between [0, 0.06) for both periods were plotted on a p-curve. A P-curve traditionally refers to the distribution of statistically significant p-values (usually < 0.05) from a set of studies. The idea is to assess whether the findings reflect true effects or are

likely due to P-hacking or publication bias. However, for our analysis, we included near-significant P-values (e.g., < 0.06) because P-hacking often results in a clustering of P-values just below 0.05. Including values up to 0.06 can help identify this pattern more clearly. The P-curves for both 2009-2010 and 2019-2020 are shown in Figure 4.8. The count of P-values in each bin is expressed as a percentage of the exact P-values for the given period are shown in table 4.5. The count of P-values in each bin per article type per period. are shown in Appendix III, Tables A7(i)-7(iv). The observed P-value distribution in both periods shown in solid red line is right-skewed, several P-values at 0 and 0.001, indicating a high frequency of very small P-values. As discussed earlier, a right skewed P-curve with a high concentration of small P-values and flattening tail as the P-values increases, suggests that the reported P-values are likely to be due to true effects rather than P-hacking. The flattening tail towards higher P-values may indicate publication bias, where studies with non-significant results are less likely to be published (Simonsohn, Nelson and Simmons, 2014). The dotted blue lines show the distribution of simulated P-values at 60%, and 90% power. Power refers to the probability that a test will correctly reject a false null hypothesis (i.e. detect a true effect when there is one). A higher power (closer to 100%) means the test is more likely to detect a true effect. For example, a power of 90%means there's a 90% probability of correctly identifying a true effect. Higher power is important because it reduces the risk of Type II errors, ensuring that true effects are not overlooked. A binomial test, which is a statistical method used to determine if there is a significant difference in proportions between an observed sample and a hypothesized sample was used to assess whether the reported distribution of P-values significantly differed from the hypothesised distribution at 60% and 90% power. The Binomial test was done for the P-value bins of [0.03, 0.04) [0.04, 0.05) and [0.05, 0.06).

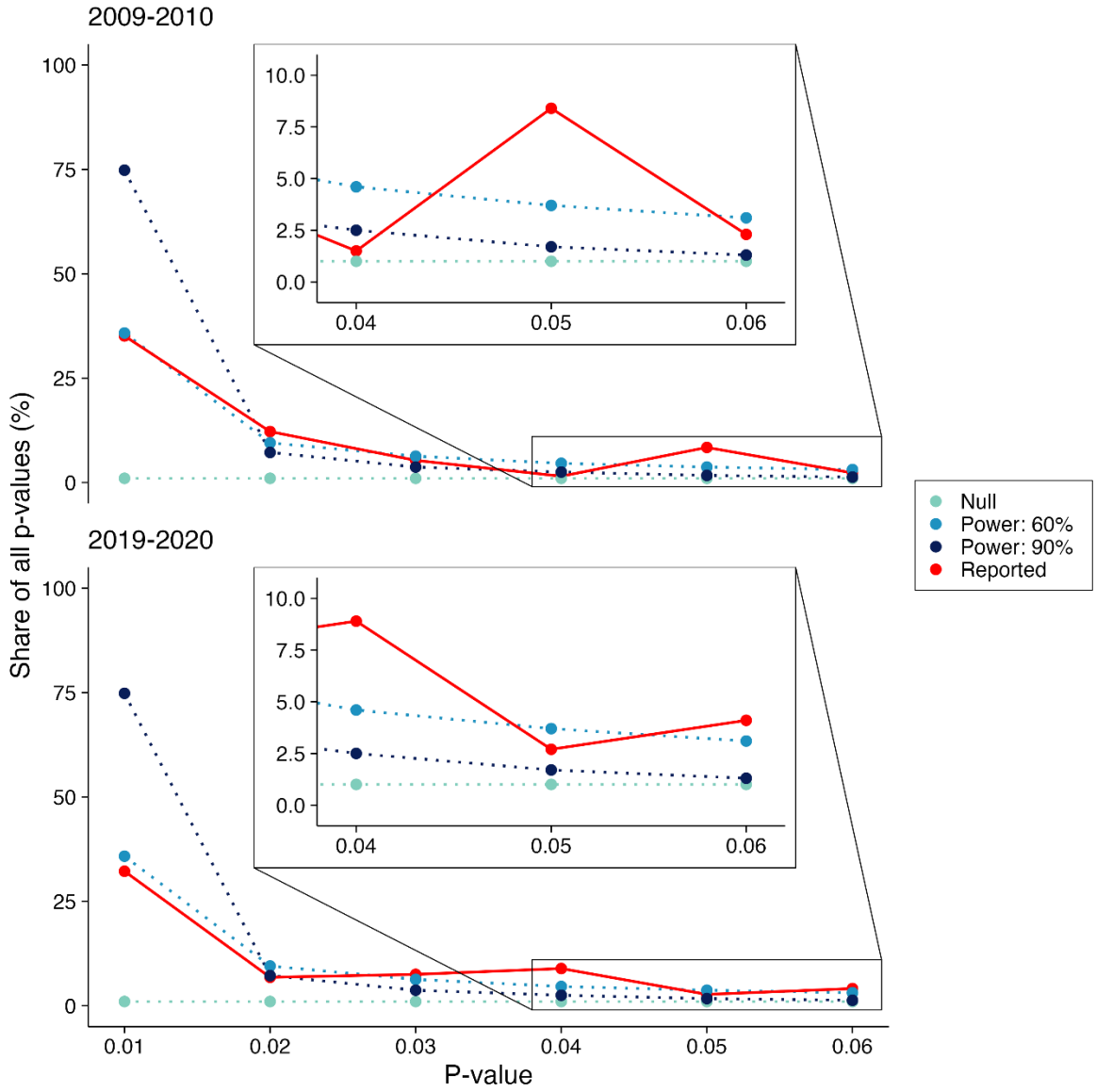



**Figure 4.8 :** P-curve illustrating the frequency distribution of reported exact P-values between [0, 0.06) for 2009-2010 and 2019-2020 period (solid red line). The dotted lines represent simulated P-curves for a null effect, and when statistical power is 60%, and 90% to detect a true difference.

**Table 4.5:** Frequency distribution of P-values

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| 2009-2010 | [0,0.01) | 46 | 131 | 35.1 |
| | [0.01,0.02) | 16 | 131 | 12.2 |
| | [0.02,0.03) | 7 | 131 | 5.3 |
| | [0.03,0.04) | 2 | 131 | 1.5 |
| | [0.04,0.05) | 11 | 131 | 8.4 |
| | [0.05,0.06) | 3 | 131 | 2.3 |
| | [0.06,1] | 46 | 131 | 35.1 |
| 2019-2020 | [0,0.01) | 47 | 146 | 32.2 |
| | [0.01,0.02) | 10 | 146 | 6.8 |
| | [0.02,0.03) | 11 | 146 | 7.5 |
| | [0.03,0.04) | 13 | 146 | 8.9 |
| | [0.04,0.05) | 4 | 146 | 2.7 |
| | [0.05,0.06) | 6 | 146 | 4.1 |
| | [0.06,1] | 55 | 146 | 37.7 |

*Analysis of the 2009-2010 P curve*

Table 4.5 and Figure 4.8 show the distribution of exact P-values between. The distribution of reported P-values for 2009-2010, is right skewed showing more smaller P-values (e.g. 0.01) than the higher ones (e.g. 0.04) which may suggest the presence of a true effect. 35.1% of the reported P-values fall below $P < 0.01$, 19% of the P-values fell within the range $0.01 \leq P < 0.04$. The P-curve shows general decrease of the distribution of the observed P-values; however, it shows an increase of P-values in the bins just below the significance threshold and a decrease just after the significance threshold 0.05. The frequency of the observed P-values is 1.5% [0.03, 0.04), 8.4% [0.04, 0.05), and 2.3% [0.05, 0.06). The high proportion of P-values just under 0.05 and a decrease in the proportion just over 0.05 suggests potential P-hacking and/or publication bias.

Table 4.6 below shows the results of the binomial tests which we performed to assess the distribution of our reported P-values for 2009-2010 to those of a hypothesised distribution at

60% and 90% power. The table shows the observed frequencies versus the expected frequencies for the bins around the significance threshold, the difference in probability, the 95% confidence interval and the associated P-values. At 60% power, bin [0.03,0.04), the observed probability of 0.015 with a 95 % confidence interval of [0.002 to 0.054] is not statistically significantly different to the hypothesized probability of 0.046 (P = 0.097). At 90% power, bin [0.03,0.04), the observed probability of 0.015 with a 95 % confidence interval of [0.002 to 0.054] is not statistically significantly different to the hypothesized probability of 0.025 (P = 0.777). At 60% power, bin [0.04;0.05), the observed probability of 0.084 with a 95 % confidence interval of [0.043 to 0.145] is statistically significantly different to the hypothesized probability of 0.037 (P = 0.001). At 90% power, bin [0.04;0.05), the observed probability of 0.084 with a 95 % confidence interval of [0.0043 to 0.145] is statistically significantly different to the hypothesized probability of 0.017 (P <0.001). The observed positive difference in probability at bin [0.04;0.05), suggests that that the observed frequency of the reported P- value is higher than the excepted frequency. These binomial test results and visual analysis of the P-curve confirm that the reported frequency of P-values is higher than the expected frequency supporting the evidence of potential P-hacking and publication bias obtained by analysing the frequency of P-values in the bins around the threshold of 0.05. At 60% power, bin [0.05,0.06), the observed probability of 0.023 with a 95 % confidence interval of [0.005 to 0.065] is not statistically significantly different to the hypothesized probability of 0.031 (P = 0.802). At 90% power, bin [0.05, 0.06], the observed probability of 0.023 with a 95 % confidence interval of [0.005 to 0.065] is not statistically significantly different to the hypothesized probability of 0.013 (P= 0.260).

**Table 4.6:** Binomial test for 2009-2010 frequency distribution of reported P-values (test for P-hacking)

| | P-value bins | Observed number in bin | Total sample size (number of exact P-values) | Observed probability (95% Confidence Interval) | Hypothesized* probability | Difference in Probability (95% Confidence Interval) | P-value |
|---|---|---|---|---|---|---|---|
| **60% power** | [0.03,0.04) | 2 | 131 | 0.015 [0.002 to 0.054] | 0.046 | -0.031 [-0.044 to 0.008] | 0.097 |
| | [0.04,0.05) | 11 | 131 | 0.084 [0.043 to 0.145] | 0.037 | 0.047 [ 0.006 to 0.108] | 0.001 |
| | [0.05,0.06) | 3 | 131 | 0.023 [0.005 to 0.065] | 0.031 | -0.008 [-0.026 to 0.035] | 0.802 |
| **90% power** | [0.03,0.04) | 2 | 131 | 0.015 [0.002 to 0.054] | 0.025 | -0.010 [-0.023 to 0.029] | 0.777 |
| | [0.04,0.05) | 11 | 131 | 0.084 [0.043 to 0.145] | 0.017 | 0.067 [ 0.025 to 0.128] | <0.001 |
| | [0.05,0.06) | 3 | 131 | 0.023 [0.005 to0.065] | 0.013 | 0.010 [-0.009 to 0.052] | 0.260 |

*Based on 100000 repeats simulation at each power

*Analysis of the 2019-2020 P curve*

Table 4.6 and figure 4.8 show the distribution of exact P-values between [0, 0.06). The distribution of reported P-values for 2019-2020, is right skewed showing more smaller P-values (e.g. 0.001) than the higher ones (e.g. 0.04) which may suggest the present of true effect. 32.2% of the reported P-values fall below P <0.01, 23% P-values fell within the range [0.01 ≤ P < 0.04). Looking closely at the frequency of P-values around the significance threshold, we found more P-values in the (8.9%) P-values in the bin [0.03;0.04), 2.7% in the bin [0.04, 0.05], and 4.1% in the bin [0.05;0.06). These results of our analysis suggests that there is no evidence of p-hacking in the 2019-2020 data but supports the presence of true effects. The increase of the frequence of the P-values just above the significance level suggest that there is an increase in the number of articles published with non-significant results which could be due to changes in journal policies and attitude towards publishing negative results as well as researchers submitting all their research regardless of the outcome.

Table 4.7 below shows the results of the binomial tests which we performed to assess the distribution of our reported P-values for 2019-2020 to those of a hypothesised distribution at 60% and 90% power. The table shows the observed frequencies against the expected frequencies for each bin, the difference in probability, the 95% confidence interval and the associated P-values. At 60% power, bin [0.03,0.04), the observed probability of 0.089 with a 95 % confidence interval of [0.046 to 0.147] is statistically significantly different to the hypothesized probability 0.046 (P=0.026). At 90% power, [0.03,0.04), the observed probability of 0.089 with a 95 % confidence interval of [0.048 to 0.147] is statistically significantly different to the hypothesized probability of 0.025 (P < 0.001). Our results mean that observed frequency has more concentration of P-values

than the expected. Generally, the overrepresentation of P-values appears just below the threshold which is consistent with questionable research practices. Our results show evidence of P-hacking in the 2009-2010 and no evidence of P-hacking a decade later (2019-2020). However, our results show a shift in the high concentration of the P-value from the bin just below the significance [0.04, 0.05) to the lower bin [0.03, 0.04). At 60% power, bin [0.04, 0.05), the observed probability of 0.027 with a 95 % confidence interval of [0.008 to 0.069] is not statistically significantly different to the hypothesized probability of 0.037 (P-=0.825). At 90% power, bin [0.04, 0.05), the observed probability of 0.027 with a 95 % confidence interval of [0.008 to 0.069] is not statistically significantly different to the hypothesized probability of 0.017(P = 0.324). At 60% power, bin [0.05, 0.06), the observed probability of 0.041 with a 95 % confidence interval of [0.015 to 0.087] is not statistically significantly different to the hypothesized probability of 0.031 (P =0.464). At 90% power, bin [0.04, 0.05), the observed probability of 0.041 with a 95 % confidence interval of [0.015 to 0.087] is statistically significantly different to the hypothesized probability of 0.013 (P = 0.015). These findings align with our previous findings from Table 4.8.

**Table 4.7:** Binomial test for 2019-2020 frequency distribution of reported P-values (test for P-hacking)

| | P-value bins | Observed number in bin | Total sample size (number of exact P-values) | Observed probability (95% Confidence Interval) | Hypothesized* probability | Difference in Probability (95% Confidence Interval) | P-value |
|---|---|---|---|---|---|---|---|
| **60% power** | [0.03, 0.04) | 13 | 146 | 0.089 [0.048 to 0.147] | 0.046 | 0.043 [ 0.002 to 0.101] | 0.026 |
| | [0.04, 0.05) | 4 | 146 | 0.027 [0.008 to 0.069] | 0.037 | -0.010 [-0.030 to 0.032] | 0.825 |
| | [0.05, 0.06) | 6 | 146 | 0.041 [0.015 to 0.087] | 0.031 | 0.010 [-0.016 to 0.057] | 0.464 |
| **90% power** | [0.03, 0.04) | 13 | 146 | 0.089 [0.048 to 0.147] | 0.025 | 0.064 [ 0.023 to 0.122] | <0.001 |
| | [0.04, 0.05) | 4 | 146 | 0.027 [0.008 to 0.069] | 0.017 | 0.010 [-0.010 to 0.051] | 0.324 |
| | [0.05, 0.06) | 6 | 146 | 0.041 [0.015 to 0.087] | 0.013 | 0.028 [ 0.002 to 0.074] | 0.015 |

*Based on 100000 repeats simulation at each power

# Chapter 5: Discussion

The reporting of P-values in research has been a topic of significant interest, particularly in the context of ensuring the integrity and reliability of research findings. The aim of my study was to assess the quality and integrity of statistical reporting in the journal *Pain* by reviewing P-value reporting in the journal to establish how P-values were being presented, if there was evidence of good reporting practices, and whether there was evidence of any questionable research practices over two distinct periods (2009-2010 and 2019-2020).

As mentioned earlier, P-hacking is identified by an overabundance of P-values just below the significance threshold of 0.05 compared to those just above it (Head et al., 2015). To detect P-hacking, researchers look for an excess of P-values, compared to what you would expect for a given level of statistical power, in the range [0.04, 0.05). An excess of P-values in the [0.04, 0.05) range suggests either P-hacking, where significant results are favoured for publication. A higher number of P-values just below 0.05 indicates P-hacking, as researchers stop adjusting their analyses once they achieve significance. A significant drop in P -values above the significance threshold of 0.05, is interpreted as evidence of bias. P-curve analysis has also revealed P-hacking psychology (Masicampo and Lalande, 2012).

We plotted the distribution of exact P-values in the range [0, 0.06) to identify if there were anomalous concentrations around the 0.05 threshold. The right-skewed P-curves indicate that the reported findings are likely to reflect true effects.

Our results show evidence of P-hacking in the 2009-2010 and no evidence of P-hacking a decade later (2019-2020). However, our results show a shift in the high concentration of the P-value from the bin just below the significance [0.04, 0.05) to the lower bin [0.03, 0.04). The ongoing discourse in the scientific community regarding the misuse and misinterpretation of P-values has led to significant changes in research practices. One possibility is that in an effort to avoid scrutiny, researchers have shifted their focus from the traditional 0.05 threshold to a slightly lower threshold of 0.04. This adjustment could be as a result of pressure to produce statistically significant results, which can sometimes lead to questionable research practices. However, our sample size is too small to make a definitive interpretation of this result. Future research could involve looking more closely at specific articles that report P-values in this specific range, examining factors such as sample sizes, statistical tests used, article types, and whether funding sources were involved.

The P-curve analysis for 2019-2020 shows that the distribution of P-values is right-skewed, indicating a high frequency of very low P-values, which suggests the presence of true effects in the data rather than P-hacking, indicating a positive change in the reporting of statistics within the decade. The change in results suggests that researchers in the 2009-2010 period may have been more likely to P-hack their analysis methods to achieve statistically significant results. It is impossible to tell if these practices were inadvertent (e.g., because of lack of knowledge about what analytical approaches are appropriate) or deliberate. In contrast, the absence of P-hacking in 2019-2020 implies that research practices may have improved, with researchers adhering more strictly to unbiased and transparent methods. The improvement in the transparency of reporting statistics could be due to increased awareness of P-hacking and its negative impact on scientific integrity, as well as the implementation of stricter guidelines and better practices in the research

community. Our results reflect a positive trend towards more reliable and reproducible research findings.

Enhancing research transparency, one of the suggested ways of dealing with questionable research practices, is critically important in all research, including pain research, which can be achieved by practices such as the use of registered reports, and the public sharing of data and code (Chavalarias *et al.*, 2016). A registered report is a journal article where authors detail their study plan, methods, and analyses. The report goes through peer review, and if approved, the journal commits to publishing the results regardless of the outcome. Journals play a pivotal role in promoting research transparency, especially through the implementation of policies and mandates. The *Pain* journal does not explicitly require the pre-registration of protocols or reports for all submissions. However, it encourages transparency and adherence to ethical guidelines. The journal's guidelines on transparency may be contributing to the improved characteristics of P-value reporting found in our study for 2019-2020. The *Pain* journal's guidelines on clinical trials and data sharing emphasize data transparency, data citation, analytical code and research materials transparency, and adherence to reporting guidelines. The journal encourages pre-registration of animal and human research studies and requires authors to indicate pre-registration in their manuscripts. In psychology, registered report studies have shown significantly fewer positive results (44%) compared to non-registered report studies (96%), where "positive" refers to statistically significant findings (Scheel *et.al* 2021). We did not assess whether studies included in our sample were pre-registered, which could have provided more clarity to our results.

Our P-curve analysis had some limitations; our sample size was small and limited the effectiveness of our analysis. One of the reasons we had a small sample size was because we manually extracted the P-values as opposed to using text mining. This approach was labour-intensive limiting the dates we assessed to two two-year periods. Indeed, our original design only required the assessment of two one-year assessment periods, but the lack of data made us extend our investigation to two two-year periods. However, the use of manual assessment had benefits. Simonsohn *et al.* (2014) stressed that for P-curve analysis to be valid, selected P-values must (1) relate to the hypothesis, (2) be statistically independent, and (3) follow a uniform distribution under the null hypothesis. Since our P-values were manually extracted, they were all linked to a primary hypothesis, a level of accuracy that text mining cannot achieve.

Our study focused exclusively on experimental research articles. We found that 98% of the experimental research articles published in 2009-2010 and 95% of those published in 2019-2020 reported a clear primary hypothesis. The high percentage of articles which stated the primary hypothesis highlights the critical role of a well-defined hypothesis in guiding research design, focusing data collection, facilitating analysis, and enhancing interpretation, collectively contributing to the robustness and reliability of experimental research which is crucial for advancing scientific knowledge (Bulajic, 2012).However, 16% of the studies in 2009-2010 and 13% of studies in 2019-2020 had no P-values linked to the primary hypothesis. Since experimental research, by definition, involves experimenter manipulations with the objective being to assess the effect of the manipulation on a defined outcome, it is surprising that so many primary hypotheses were not accompanied by P-values; an indirect assessment of whether a comparison was made between the interventions.

P-curves have limitations for example; the estimates a of a p curve are not robust when the sample size is small and only statistically significant results are used for analysis. Using a Z-curve method of analysis could have been advantageous for our study as it includes both significant and non-significant results, thus offering a larger sample size. Unlike p-curve, which focuses only on significant p-values and may limit its ability to detect p-hacking, the Z-curve adjusts for publication bias by estimating the proportion of studies missing due to non-significant results. This provides a clearer picture of the data's reliability.

The proportion of significant P-values remained high across both periods (83% overall), with a slight increase in non-significant results from 16% in 2009-2010 to 19% in 2019-2020. The high proportion of significant results suggests potential issues with selective reporting or publication bias. Alternatively, it could mean that most studies are well powered to detect the effect of experimental interventions. Consistent with our data, Head et al. (2015) found that the proportion of significant results in biomedical research has increased over time. They speculated that the elevation could potentially be due to publication bias. However, the increase in non-significant results, particularly in clinical science articles, found in our study may reflect more robust experimental design.

Our study showed an increase in the reporting of exact P-values and towards more precise statistical reporting. These changes in statistical reporting may be driven by evolving guidelines, increased awareness of best practices, and advancements in statistical methods. This aligns with

the recommendations by the American Statistical Association (Wasserstein & Lazar, 2016), emphasizing the need for more accurate and transparent reporting practices. Improved statistical reporting can enhance the reliability and reproducibility of research findings, which is crucial for advancing scientific knowledge. Wasserstein & Lazar (2016) emphasized the importance of context and the limitations of relying solely on P-values for scientific conclusions. However, the trend of better statistical reporting observed over the decade might be influenced by the increasing pressure to produce significant results in the competitive scientific environment the introduction of stricter guidelines (Chavalarias *et al.*, 2016, American Statistical Association, 2022, Hesson-McInnis, 2025 ) ASA&APA guidelines. Reporting the exact P-values helps readers understand the results better, considering the specific details of the research setting or study design. Similar studies in other fields have reported varying trends in P-value reporting. For instance, a study published in the *JAMA Network* analysed the evolution of P-value reporting in the biomedical literature from 1990 to 2015 and found that the proportion of articles reporting at least one statistically significant P-value (≤0.05) remained relatively stable over time. However, the study also noted an increase in the reporting of exact P-values, which aligns with the findings of this study in pain research (Chavalarias *et al.*, 2016). In the field of medical research, (Kruschke, *et al.,* 2012) discussed the importance of reporting exact P-values and the potential issues associated with the misuse of P-values. They emphasized that while P-values themselves are not flawed, their misuse can lead to misleading conclusions.

Our results for the reporting practices for P-values and their corresponding test statistics over the two periods, 2009-2010 and 2019-2020 show that there is a significant association between the type of P-value (exact or approximate) and whether a corresponding test statistic was also reported.

This finding implies that researchers were more diligent in reporting test statistics when exact P-values were provided. For the 2019-2020 period, there was an improvement in reporting practices, with 66% of the exact P-values having a corresponding test statistic which reflects a positive trend towards better reporting standards over the decade. The consistent finding of a significant association between the exact P-value and a corresponding test statistic across both periods highlights the importance of exact P-values in promoting comprehensive reporting of statistical analyses. The improvement in reporting practices over time may be attributed to increased awareness and emphasis on research transparency and reproducibility within the scientific community. However, the fact that a substantial proportion of approximate P-values still lack corresponding test statistics indicates that there is room for further improvement. Future efforts should focus on encouraging researchers to consistently report test statistics for all types of P-values, whether exact or approximate.

A similar trend was found for the reporting of sample size with p-values. The reporting of sample sizes with P-values is of critical importance because for a given effect size, the power to detect a statistically significant difference increases with sample size. Thus, with a large enough sample size, a statistically significant effect may be detected, but the magnitude, and therefore the practical significance of the effect may be minor. This relationship obviously also works in the opposite direction such that large effect sizes may be rejected as being statistically significant (false negatives) because the sample size is inadequate and the statistical power too low. Thus, both too larger sample size (which is an ethically questionable practice in its own right) may lead to false positive and too small sample size may lead to false negatives (use of an inappropriately small sample size is also ethically questionable). It is a limitation of this study that we only looked at the

relationship between P-values and sample size, and did not include effect sizes, but the array of effect sizes used (standardised and unstandardised) made extracting these data difficult.

The distribution of article types showed a shift in research focus over the decade, with an increase in clinical science and a decrease in basic science (human) articles. The increase in clinical research publications may reflect the growing emphasis on clinical applications and translational research in pain management. This trend may also reflect changes in funding priorities, emerging research areas, or evolving scientific questions. Similar trends were observed by Chavalaris *et.al* (2016) who noted an increase in clinical research publications in medical journals. Understanding the shift in research focus can help policymakers and funders for example, to allocate resources more effectively to address current and future research.

In conclusion, our study shows a general improvement in the reporting of statistics over the decade studied, particularly in reporting exact p-values along with their corresponding sample sizes and test statistics. Notably, the potential p-hacking and publication bias observed in the 2009-2010 period were not found in the 2019-2020 period.

While we used p-curve analysis, which has its limitations, our findings highlight potential issues that need to be addressed to maintain the integrity of research. Continuous education on the correct use and reporting of p-values is necessary to eliminate the use of approximate p-values and ensure p-values are reported with corresponding statistics and sample sizes. Researchers should be equipped with the skills to conduct rigorous and transparent studies, interpret statistics accurately,

and report findings responsibly. Journals can play a leading role by providing workshops and comprehensive guidelines on ethical research practices and conduct. They can also make it mandatory for all research to include a detailed research report.

Our study assessed data from the journal *Pain* and focused on two specific periods: 2009-2010 and 2019-2020. While *Pain* may not represent all pain-related research, it was selected for this study due to its status as the leading journal in the field by impact factor ranking. Future studies should include a wider range of journals to provide a more comprehensive overview of statistical reporting in pain-related research. Additionally, future research should compare the p-values of articles with pre-registered reports to those without. Increasing the sample size in future studies is also important to ensure all categories of articles have sufficient sample sizes, as this affects statistical power and the ability to detect differences.

The findings of this study have important implications for future research in pain management. The improvement in the statistical reporting as shown by the general leaning towards more precise statistical reporting and the absence of P-hacking or publication bias in 2019-2020 as compared to 2009-2010 enhance the credibility of the research in this field. Future studies should continue to prioritize transparency and methodological rigor to build on these findings and further advance the field. Exaggerating the effects and distorting evidence for scientific claims can lead to the misallocation of resources, as researchers may pursue areas with minimal evidence. Additionally, promoting unproven or ineffective treatments can harm the public and diminish trust in scientific institutions.

# References

Bin Abd Razak, H.R. *et al.* (2016) 'P-hacking in orthopaedic literature: A twist to the Tail', *Journal of Bone and Joint Surgery - American Volume*, 98(20), p. e91. Available at: https://doi.org/10.2106/JBJS.16.00479.

Agresti, A., Franklin, C. and Klingenberg, B. (2018) *The art and science of learning from data*, *Antimicrobial Agents and Chemotherapy*.

American Statistical Association (2005) *Guidelines for assessment and instruction in statistics education: College report*, *Report*.

Amrhein, V., Trafimow, D. and Greenland, S. (2019) 'Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication', *American Statistician*, 73(sup1), pp. 262–270. Available at: https://doi.org/10.1080/00031305.2018.1543137.

Baum, J. and Bromiley, P. (2019) 'P-hacking in Top-tier Management Journals', *Academy of Management Proceedings*, 2019(1), p. 10810. Available at: https://doi.org/10.5465/ambpp.2019.10810abstract.

Bedeian, A., Taylor, S. and Miller, A. (2010) 'Management science on the credibility bubble: Cardinal sins and various misdemeanors', *Academy of Management Learning and Education*, 9(4), pp. 715–725. Available at: https://doi.org/10.5465/AMLE.2010.56659889.

Berman, R. *et al.* (2018) *p-Hacking and False Discovery in A/B Testing*. Available at: https://ssrn.com/abstract=3204791.

Bishop, D.V.M. and Thompson, P.A. (2016) 'Problems in using p -curve analysis and text-mining to detect rate of p -hacking and evidential value', *PeerJ*, 4, p. e1715. Available at:

https://doi.org/10.7717/peerj.1715.

Brassington, I. (2017) 'The ethics of reporting all the results of clinical trials', *British Medical Bulletin*, 121(1), pp. 19–29. Available at: https://doi.org/10.1093/bmb/ldw058.

Brodeur, A. *et al.* (2016) 'Star wars: The empirics strike back', *American Economic Journal: Applied Economics*, 8(1), pp. 1–32. Available at: https://doi.org/10.1257/app.20150044.

Bruns, S.B. and Ioannidis, J.P.A. (2016) 'p -Curve and p -Hacking in Observational', pp. 1–13. Available at: https://doi.org/10.1371/journal.pone.0149144.

Bulajic, A. (2012) 'The importance of defining the hypothesis in scientific research', *International Journal of Educational Administration and Policy Studies*, 4(8), pp. 170–176. Available at: https://doi.org/10.5897/ijeaps12.009.

Butler, N., Delaney, H. and Spoelstra, S. (2017) 'The gray zone: Questionable research practices in the business school', *Academy of Management Learning and Education*, 16(1), pp. 94–109. Available at: https://doi.org/10.5465/amle.2015.0201.

Chavalarias, D. *et al.* (2016) 'Evolution of Reporting P Values in the Biomedical Literature, 1990-2015', *JAMA*, 315(11), p. 1141. Available at: https://doi.org/10.1001/jama.2016.1952.

Chittaranjan Andrade (2021) 'HARKing, Cherry picking, P-hacking, Fishing expeditions, and Data Dredging and Mining as Questionable Research Practices', *Journal of Clinical and Practical Psychopharmacology*, 82:1(January/February), pp. e1–e3.

Council for International Organizations of Medical Sciences (2009) 'S. Rose, International Ethical Guidelines for Epidemiological Studies. American Journal of Epidemiology.', *International Ethical Guidelines for Epidemiological Research*, 170(11), p. 26. Available at: Geneva, Switzerland, World Health Organization.

van Dalen, H.P. (2021) 'How the publish-or-perish principle divides a science: the case of

economists', *Scientometrics*, 126(2), pp. 1675–1694. Available at: https://doi.org/10.1007/s11192-020-03786-x.

Davies, B. and Felappi, G. (2017) 'Publish or Perish', *Metaphilosophy*, 48(5), pp. 745–761. Available at: https://doi.org/10.1111/meta.12269.

Dirnagl, U. and Lauritzen, M. (2010) 'Fighting Publication Bias: Introducing the Negative Results Section', *Journal of Cerebral Blood Flow & Metabolism*, 30(7), pp. 1263–1264. Available at: https://doi.org/10.1038/jcbfm.2010.51.

Dixon, P. (2003) 'The p-value fallacy and how to avoid it.', *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, 57(3), pp. 189–202. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14596477 (Accessed: 5 July 2019).

Duyx, B. *et al.* (2017) 'Scientific citations favor positive results: a systematic review and meta-analysis', *Journal of Clinical Epidemiology*, pp. 92–101. Available at: https://doi.org/10.1016/j.jclinepi.2017.06.002.

Elliss, P. (2010) *The Essential Guide to Effect Sizes*.

Emanuel, E.J., Wendler, D. and Grady, C. (2000) 'What Makes Clinical Research Ethical?', *JAMA*, 283(20), pp. 2701–2711. Available at: https://doi.org/10.1001/jama.283.20.2701.

Erin Leahey (2005) 'Alphas and Asterisks : The Development of Statistical Significance Testing Standards in Sociology Author ( s ): Erin Leahey Published by : Oxford University Press Stable URL : https://www.jstor.org/stable/3598292 Alphas and Asterisks : The Development of', 84(1), pp. 1–24.

Erômeerˆerôme Adda, J.′, Decker, C. and Ottaviani, M. (2020) 'P-hacking in clinical trials and how incentives shape the distribution of results across phases', 117(24). Available at: https://doi.org/10.7910/DVN/NBLYSW).y.

Fanelli, D. (2010a) ‘Do Pressures to Publish Increase Scientists’ Bias? An Empirical Support from US States Data’, *PLoS ONE*. Edited by E. Scalas, 5(4), p. e10271. Available at: https://doi.org/10.1371/journal.pone.0010271.

Fanelli, D. (2010b) ‘Do Pressures to Publish Increase Scientists’ Bias? An Empirical Support from US States Data’, *PLoS ONE*. Edited by E. Scalas, 5(4), p. e10271. Available at: https://doi.org/10.1371/journal.pone.0010271.

Finnerup, N.B. *et al.* (2015) ‘Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis.’, *The Lancet. Neurology* [Preprint]. Available at: https://doi.org/10.1016/S1474-4422(14)70251-0.

Fisher, R.A. and Johnstone, D.J. (1987) *Tests of Significance following R. A. Fisher*. Available at: https://www.jstor.org/stable/687354.

Franco, A., Malhotra, N. and Simonovits, G. (2014) ‘Publication bias in the social sciences: Unlocking the file drawer’, *Science*, 345(6203), pp. 1502–1505. Available at: https://doi.org/10.1126/science.1255484.

Gelman, A. and Carlin, J. (2014) ‘Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors’, *Perspectives on Psychological Science*, 9(6), pp. 641–651. Available at: https://doi.org/10.1177/1745691614551642.

Gigerenzer, G. (2004) ‘Mindless statistics’, *Journal of Socio-Economics*, 33(5), pp. 587–606. Available at: https://doi.org/10.1016/j.socec.2004.09.033.

Goodman, S.N. (1992) ‘A comment on replication,P-values and evidence’, *Statistics in Medicine*, 11(7), pp. 875–879. Available at: https://doi.org/10.1002/sim.4780110705.

Goodman, S.N. (1999) ‘Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy’, *Annals of Internal Medicine*, 130(12), p. 995. Available at: https://doi.org/10.7326/0003-4819-

130-12-199906150-00008.

Greenland, S. *et al.* (2016) 'Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations', *European Journal of Epidemiology*, 31(4), pp. 337–350. Available at: https://doi.org/10.1007/s10654-016-0149-3.

Hartgerink, C.H.J. *et al.* (2016) 'Distributions of p -values smaller than .05 in psychology: what is going on?', *PeerJ*, 4, p. e1935. Available at: https://doi.org/10.7717/peerj.1935.

Head, M.L. *et al.* (2015) 'The Extent and Consequences of P-Hacking in Science', *PLOS Biology*, 13(3), p. e1002106. Available at: https://doi.org/10.1371/journal.pbio.1002106.

Hoekstra, R. *et al.* (2014) 'Robust misinterpretation of confidence intervals', *Psychonomic Bulletin and Review*, 21(5), pp. 1157–1164. Available at: https://doi.org/10.3758/s13423-013-0572-3.

Hoekstra, R., Finch, S. and Johnson, A. (2006) 'Probability as certainty: Dichotomous thinking and the misuse of p values', *Psychonomic Bulletin and Review*, 13(6), pp. 1033–1037. Available at: https://doi.org/10.3758/BF03213921.

Hullett, C., Park, H.S. and Lindsey, L.L.M. (2008) 'A Critical Assessment of Null Hypothesis Significance Testing in Quantitative Communication Research', 34(May 2003), pp. 171–187. Available at: https://doi.org/10.1111/j.1468-2958.2008.00317.x.

Hung, H.M. *et al.* (1997) 'The behavior of the P-value when the alternative hypothesis is true.', *Biometrics*, 53(1), pp. 11–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9147587 (Accessed: 5 July 2019).

Ioannidis, J.P.A. (2005) 'Why most published research findings are false', *Plos Medicine*, 2(8), pp. 0696–0701. Available at: https://doi.org/10.1371/journal.pmed.0020124.

Ioannidis, J.P.A. (2008) 'Why Most Discovered True Associations Are Inflated', *Epidemiology*,

19(5), pp. 640–648. Available at: https://doi.org/10.1097/EDE.0b013e31818131e7.

Ioannidis, J.P.A. *et al.* (2014) 'Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention', *Trends in Cognitive Sciences*, 18(5), pp. 235–241. Available at: https://doi.org/10.1016/j.tics.2014.02.010.

Ioannidis, J.P.A. (2019) 'What Have We (Not) Learnt from Millions of Scientific Papers with P Values?', *The American Statistician*, 73(sup1), pp. 20–25. Available at: https://doi.org/10.1080/00031305.2018.1447512.

Kennedy-Shaffer, L. (2019) 'Before $p < 0.05$ to Beyond $p < 0.05$: Using History to Contextualize p-Values and Significance Testing', *American Statistician*, 73(sup1), pp. 82–90. Available at: https://doi.org/10.1080/00031305.2018.1537891.

Kruschke, J.K., Aguinis, H. and Joo, H. (2012) 'The Time Has Come', *Organizational Research Methods*, 15(4), pp. 722–752. Available at: https://doi.org/10.1177/1094428112457829.

Lakens, D. (2022) 'Sample Size Justification', *Collabra: Psychology*, 8(1), pp. 1–28. Available at: https://doi.org/10.1525/collabra.33267.

Leek, J. and Pen, R. (2015) 'P values are just the tip of the iceberg', *Nature*, 520(30 April 2015), p. 2015.

Lu, Y. and Belitskaya-Levy, I. (2015) 'The debate about p-values.', *Shanghai archives of psychiatry*, 27(6), pp. 381–5. Available at: https://doi.org/10.11919/j.issn.1002-0829.216027.

Masicampo, E. J. and Lalande, D.R. (2012) 'A peculiar prevalence of p values just below .05.', *Quarterly journal of experimental psychology (2006)*, 65(11), pp. 2271–9. Available at: https://doi.org/10.1080/17470218.2012.711335.

Masicampo, E.J. and Lalande, D.R. (2012) 'A peculiar prevalence of p values just below .05', *Quarterly Journal of Experimental Psychology*, 65(11), pp. 2271–2279. Available at:

https://doi.org/10.1080/17470218.2012.711335.

McGauran, N. *et al.* (2010) 'Reporting bias in medical research - a narrative review', *Trials*, 11(1), p. 37. Available at: https://doi.org/10.1186/1745-6215-11-37.

McShane, B.B. *et al.* (2019) 'Abandon Statistical Significance', *American Statistician*, 73(sup1), pp. 235–245. Available at: https://doi.org/10.1080/00031305.2018.1527253.

Mertens, W. and Recker, J. (2020) 'New guidelines for null hypothesis significance testing in hypothetico-deductive is research', *Journal of the Association for Information Systems*, 21(4), pp. 1072–1102. Available at: https://doi.org/10.17705/1jais.00629.

Mezones-Holguín, E. *et al.* (2024) 'THE NULL HYPOTHESIS SIGNIFICANCE TEST AND THE DICHOTOMIZATION OF THE P-VALUE: Errare Humanum Est', *Revista Peruana de Medicina Experimental y Salud Publica*, 41(4), pp. 422–430. Available at: https://doi.org/10.17843/RPMESP.2024.414.14285.

Miller, J. (2009) 'What is the probability of replicating a statistically significant effect ?', 16(4), pp. 617–640. Available at: https://doi.org/10.3758/PBR.16.4.617.

Munafò, M. and Neill, J. (2016) 'Null is beautiful: On the importance of publishing null results', *Journal of Psychopharmacology*, 30(7), p. 585. Available at: https://doi.org/10.1177/0269881116638813.

O'Brien, R.G.O. (no date) 'Sample-Size Analysis for Traditional Hypothesis Testing : Concepts and Issues', (January 2007), pp. 237–271.

Okike, K. *et al.* (2008) 'Publication bias in orthopaedic research: An analysis of scientific factors associated with publication in The Journal of Bone and Joint Surgery (American volume)', *Journal of Bone and Joint Surgery*, 90(3), pp. 595–601. Available at: https://doi.org/10.2106/JBJS.G.00279.

Ostasiewicz, W. (2014) ‘The emergence of statistical science’, *Śląski Przegląd Statystyczny* [Preprint], (12 (18)). Available at: https://doi.org/10.15611/sps.2014.12.04.

Perezgonzalez, J.D. (2015) ‘Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing’, *Frontiers in Psychology*, 6(MAR). Available at: https://doi.org/10.3389/fpsyg.2015.00223.

Pernet, C. (2017) ‘Null hypothesis significance testing: a short tutorial’, *F1000Research*, 4, p. 621. Available at: https://doi.org/10.12688/f1000research.6963.1.

Phyllis G Supino Jeffrey S. Borer (2012) *Principles of Research Methodology*.

Pondrom, S. (2016) ‘The Value of P’, *American Journal of Transplantation*, 16(10), pp. 2769–2770. Available at: https://doi.org/10.1111/ajt.14026.

Reynolds, S.S. and Waldrop, J. (2024) ‘Misuse of the P Value Using Quality Improvement Analyses to Identify Clinically Significant Improvements’, *Dimensions of Critical Care Nursing*, 43(2), pp. 96–101. Available at: https://doi.org/10.1097/DCC.0000000000000623.

Rooprai, P. *et al.* (2023) ‘Is There Evidence of P-Hacking in Imaging Research?’, *Canadian Association of Radiologists Journal*, 74(3), pp. 497–507. Available at: https://doi.org/10.1177/08465371221139418.

Sanfilippo, P.G. *et al.* (2016) ‘Review of null hypothesis significance testing in the ophthalmic literature: Are most “significant” P values false positives?’, *Clinical and Experimental Ophthalmology*, pp. 52–61. Available at: https://doi.org/10.1111/ceo.12570.

Schwab, A. *et al.* (2011) ‘PERSPECTIVE — Researchers Should Make Thoughtful Assessments Instead of Null-Hypothesis Significance Tests Researchers Should Make Thoughtful Assessments Instead of Null-Hypothesis Significance Tests’, (February 2024). Available at: https://doi.org/10.1287/orsc.1100.0557.

Sedgwick, P. (2015) ‘Confidence intervals, P values, and statistical significance’, *BMJ (Online)*, 350(August). Available at: https://doi.org/10.1136/bmj.h1113.

Sedgwick, P.M. *et al.* (2022) ‘Current controversies : Null hypothesis significance testing’, (February), pp. 624–627. Available at: https://doi.org/10.1111/aogs.14366.

Sena E.S., van der Worp H.B., Bath P.M.W., Howells D.W., and Macleod M.R. (2010) Publication Bias in Reports of Animal Stroke Studies Leads to Major Overstatement of Efficacy. *PLOS Biology* 8(3): e1000344. https://doi.org/10.1371/journal.pbio.1000344

Simmons, J.P., Nelson, L.D. and Simonsohn, U. (2011) ‘False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant’, *Psychological Science*, 22(11), pp. 1359–1366. Available at: https://doi.org/10.1177/0956797611417632.

Simonsohn, U., Nelson, L.D. and Simmons, J.P. (2014) ‘P-curve: A key to the file-drawer.’, *Journal of Experimental Psychology: General*, 143(2), pp. 534–547. Available at: https://doi.org/10.1037/a0033242.

Soliman, N., Rice, A.S.C., and Vollert, J. (2020) A practical guide to preclinical systematic review and meta-analysis. *PAIN* 161(9):pp 1949-1954, https://10.1097/j.pain.0000000000001974

Stefan, A.M. and Schönbrodt, F.D. (2023) ‘Big little lies: A compendium and simulation of p-hacking strategies’, *Royal Society Open Science*, 10(2). Available at: https://doi.org/10.1098/rsos.220346.

Szucs, D. and Ioannidis, J.P.A. (2017) ‘Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature’, *PLoS Biology*, 15(3), pp. 1–18.

Available at: https://doi.org/10.1371/journal.pbio.2000797.

Teixeira, P.M. (2018) 'p-Hacking – A call for ethics', *Pulmonology*, 24(3), pp. 207–208. Available at: https://doi.org/10.1016/j.pulmoe.2018.03.001.

Trafimow, D. and Marks, M. (2015) 'Editorial', *Basic and Applied Social Psychology*, 37(1), pp. 1–2. Available at: https://doi.org/10.1080/01973533.2015.1012991.

Treanor, L. *et al.* (2020) 'Publication bias in diagnostic imaging: conference abstracts with positive conclusions are more likely to be published', *European Radiology*, 30(5), pp. 2964–2972. Available at: https://doi.org/10.1007/s00330-019-06568-z.

Wasserstein, Ronald L. and Lazar, N.A. (2016) 'The ASA's Statement on p-Values: Context, Process, and Purpose', *American Statistician*, 70(2), pp. 129–133. Available at: https://doi.org/10.1080/00031305.2016.1154108.

Wasserstein, Ronald L and Lazar, N.A. (2016) 'The ASA Statement on p -Values: Context, Process, and Purpose', *The American Statistician*, 70(2), pp. 129–133. Available at: https://doi.org/10.1080/00031305.2016.1154108.

Whitley, E. and Ball, J. (2002) 'Statistics review 3: Hypothesis testing and P values', *Critical Care* [Preprint]. Available at: https://doi.org/10.1186/cc1493.

# Appendix I: Data capturing tool



## Journal Details

| | |
|---|---|
| Record ID | _____ |
| **Journal Details** | |
| Journal | ◯ The Journal of pain<br>◯ Pain |
| First Author | _____ |
| Publication Date | _____ |
| Volume Number | _____ |
| Issue Number | _____ |
| Page Numbers | _____ |
| DOI Number | _____ |
| Journal Type | ◯ Basic Science Humans<br>◯ Basic Science Animals<br>◯ Basic Science Invitro/ Ex vivo<br>◯ Clinical trials<br>◯ Systematic Review Humans<br>◯ Systematic Review Animals<br>◯ Observational Humans<br>◯ Observational Animals<br>◯ Commentary/Editorial/ Letter<br>◯ Narrative Reviews<br>(Choose one) |
| Clear Primary Hypothesis | ◯ Yes<br>◯ No |

# P Values

| **P-values** | |
|---|---|
| P-values For Primary Hypothesis In The Results | ◯ Yes<br>◯ No |
| Exact P-value | ◯ Yes<br>◯ No |
| P-value For The Primary Hypothesis (Exact ) | _____ |
| P-value For The Primary Hypothesis | ☐ Range<br>☐ Calculated |
| P-value For The Primary Hypothesis (Approximate Value) | ☐ P< 0.001<br>☐ P< 0.01<br>☐ P< 0.05<br>☐ P>0.05<br>(Choose one) |
| P-value For the Primary Hypothesis (Calculated Value) | _____ |
| Sample Size For The Primary Hypothesis | ◯ Exact<br>◯ Range<br>◯ Unclear<br>(Choose one) |
| Exact Sample Size | _____ |
| Sample Size Range | _____ |

# Statistics

| **Statistics** | |
|---|---|
| Test Statistics | ◯ Yes<br>◯ No |
| Hypothesis Test (Test Statistics Used) | ☐ t-test<br>☐ ANOVA<br>☐ Chi-squared<br>☐ Pearson's Correlation<br>☐ Spearman's Correlation<br>☐ Linear Regression<br>☐ Logistic Regression<br>☐ Fisher's Exact Test<br>☐ Ancova<br>☐ General Linear Model<br>☐ General Linear Mixed Model<br>☐ Other<br>☐ Unclear/ Not Specified |
| T test value | _____ |
| Degrees Of Freedom T-test | _____ |
| ANOVA Value | _____ |
| Degrees of Freedom Anova (Between Groups) | _____ |
| Degrees of Freedom Anova (Within Groups) | _____ |
| Chi-Squared Value | _____ |
| Degrees Of Freedom Chi-Squared | _____ |
| Pearson's Correlation Value | _____ |
| Spearman's Correlation | _____ |
| Linear Regression Value | _____ |
| Logistic Regression Value | _____ |

| | |
|---|---|
| Fisher's Exact test value | _____ |
| Ancova Value | _____ |
| General Linear Model Value | _____ |
| General Linear Mixed Model Value | _____ |
| Other | _____ |

# Supervisor Comments

| **Notes For Supervisors** |
| --- |
| Comments<br>_____ |
| The Publication |

# Appendix II : Ethics Clearance

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Ref: W-CP-191108-1

08 November 2019

***TO WHOM IT MAY CONCERN***:

**Waiver:** This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical).

| | |
|---|---|
| **Investigator:** | **Ms Tapiwa Tsitsi Chinaka (student no 342873)** |
| **Supervisor:** | **Prof P Kamerman and Prof A Fuller** |
| **Faculty:** | Health Sciences |
| **School:** | Physiology |
| **Department:** | Brain Function Research Group |
| **Project title:** | **Reporting of statistical results in the field of pain research** |
| **Reason:** | Literature Review |

Dr Clement Penny

Chair: Human Research Ethics Committee (Medical)

Copy – HREC (Medical) Secretariat: Zanele Ndlovu, Rhulani Mkansi,Iain Burns, Mapula Ramaila

Research Office Secretariat: Faculty of Health Sciences, Phillip Tobias Building, 3rd Floor, Office 301, 29 Princess of Wales Terrace, Parktown, 2193 Tel +27 (0)11-717-1252 /1234/2656/2700 Private Bag 3, Wits 2050, email: zanele.ndlovu@wits.ac.za Office email: hrec-medical.researchoffice@wits.ac.zaWebsite:

# Appendix III: Frequency tables

**Table A1: Frequency distribution of articles with a P-value linked to the primary hypothesis by article type and period (Data for Figure 4-1)**

| | P-value for primary hypothesis | No (%) | | Yes (%) | | Total |
|---|---|---|---|---|---|---|
| **2009-2010** | **Total** | **74** | **(16%)** | **394** | **84%** | **468** |
| | Basic science (animal) | 15 | (9%) | 154 | (91%) | 169 |
| | Basic science (human) | 23 | (20%) | 90 | (80%) | 113 |
| | Basic science (other) | 6 | (38%) | 10 | (62%) | 16 |
| | Clinical science | 30 | (18%) | 140 | (82%) | 170 |
| **2019-2020** | **Total** | **53** | **(13%)** | **344** | **(87%)** | **397** |
| | Basic science (animal) | 9 | (7%) | 125 | (93%) | 134 |
| | Basic science (human) | 12 | (18%) | 56 | (82%) | 68 |
| | Basic science (other) | 5 | (19%) | 21 | (81%) | 26 |
| | Clinical science | 27 | (16%) | 142 | (84%) | 169 |
| **Total** | **All Periods** | **127** | **(15%)** | **738** | **(85%)** | **865** |

**Table A2: Frequency distribution of experimental research article types (Data for Figure 4-2)**

| | Clear hypothesis | No (%) | | Yes (%) | | Total |
|---|---|---|---|---|---|---|
| **2009-2010** | **Total** | **9** | **(2%)** | **468** | **(98%)** | **477** |
| | Basic science (animal) | 5 | (3%) | 169 | (97%) | 174 |
| | Basic science (human) | 3 | (3%) | 113 | (97%) | 116 |
| | Basic science (other) | 0 | (0%) | 16 | (100%) | 16 |
| | Clinical science | 1 | (1%) | 170 | (99%) | 171 |
| **2019-2020** | **Total** | **17** | **(4%)** | **397** | **(96%)** | **414** |
| | Basic science (animal) | 12 | (8%) | 134 | (92%) | 146 |
| | Basic science (human) | 0 | (0%) | 68 | (100%) | 68 |
| | Basic science (other) | 1 | (4%) | 26 | (96%) | 27 |
| | Clinical science | 4 | (2%) | 169 | (98%) | 173 |
| **Both Periods** | | **26** | **(3%)** | **865** | **(97%)** | **891** |

**Table A3: Frequency distribution of significant and non-significant results by article type and period (Data for Figure 4-3)**

| | | | | | | |
|---|---|---|---|---|---|---|
| **2009–2010** | **All Articles** | **64** | **(16%)** | **330** | **(84%)** | **394** |
| | Basic Science (Animal) | 20 | (13%) | 134 | (87%) | 154 |
| | Basic Science (Human) | 16 | (18%) | 74 | (82%) | 90 |
| | Basic Science (Other) | 2 | (20%) | 8 | (80%) | 10 |
| | Clinical Science | 26 | (19%) | 114 | (81%) | 140 |
| **2019–2020** | **All Articles** | **64** | **(19%)** | **280** | **(81%)** | **344** |
| | Basic Science (Animal) | 13 | (10%) | 112 | (90%) | 125 |
| | Basic Science (Human) | 10 | (18%) | 46 | (82%) | 56 |
| | Basic Science (Other) | 1 | (5%) | 20 | (95%) | 21 |
| | Clinical Science | 41 | (29%) | 101 | (71%) | 142 |
| **Both Periods** | | **128** | **(17%)** | **610** | **(83%)** | **738** |

**Table A4: Frequency distribution of approximate and exact P-values results by article type and period (Data for Figure 4-4)**

| Time Period | Article Type | Approximate (%) | | Exact (%) | | Total |
|---|---|---|---|---|---|---|
| **2009–2010** | **All Articles** | **263** | **(67%)** | **131** | **(33%)** | **394** |
| | Basic Science (Animal) | 123 | (80%) | 31 | (20%) | 154 |
| | Basic Science (Human) | 52 | (58%) | 38 | (42%) | 90 |
| | Basic Science (Other) | 8 | (80%) | 2 | (20%) | 10 |
| | Clinical Science | 80 | (57%) | 60 | (43%) | 140 |
| **2019–2020** | **All Articles** | **198** | **(58%)** | **146** | **(42%)** | **344** |
| | Basic Science (Animal) | 94 | (75%) | 31 | (25%) | 125 |
| | Basic Science (Human) | 29 | (52%) | 27 | (48%) | 56 |
| | Basic Science (Other) | 13 | (62) | 8 | (38%) | 21 |
| | Clinical Science | 62 | (44%) | 80 | (56%) | 142 |
| **Both Periods** | | **461** | **(62%)** | **277** | **(38%)** | **738** |

**Table A5: Frequency distribution of approximate P-values results by article type and period (Data for Figure 4.5)**

| Period | | P < 0.001 (%) | | P < 0.01 (%) | | P < 0.05 (%) | | P > 0.05 (%) | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| **2009-2010** | | **103** | **(39%)** | **42** | **(%16)** | **101** | **(38%)** | **17** | **(6%)** | **263** |
| | Basic science (animal) | 33 | (27%) | 14 | (11%) | 69 | (56%) | 7 | (6%) | 123 |
| | Basic science (human) | 27 | (52%) | 12 | (23%) | 11 | (21%) | 2 | (4%) | 52 |
| | Basic science (other) | 3 | (38%) | 0 | (0%) | 4 | (50%) | 1 | (13%) | 8 |
| | Clinical science | 40 | (50%) | 16 | (20%) | 17 | (21%) | 7 | (9%) | 80 |
| **2019-2020** | | **109** | **(55%)** | **25** | **(13%)** | **58** | **(29%)** | **6** | **(3%)** | **198** |
| | Basic science (animal) | 46 | (49%) | 14 | (15%) | 33 | (35%) | 1 | (1%) | 94 |
| | Basic science (human) | 20 | (69%) | 5 | (17%) | 3 | (10%) | 1 | (3%) | 29 |
| | Basic science (other) | 6 | (46%) | 1 | (8%) | 6 | (46%) | 0 | (0%) | 13 |
| | Clinical science | 37 | (60%) | 5 | (8%) | 16 | (26%) | 4 | (6%) | 62 |
| **Both Periods** | | **212** | **(46%)** | **67** | **(15%)** | **159** | **(34%)** | **23** | **(5%)** | **461** |

**Table A6: Frequency distribution of significant and non-significant exact P-values results by article type and period (Data for Figure 4.6)**

| Period | | Not Significant (%) | | Significant (%) | | Grand Total |
|---|---|---|---|---|---|---|
| **2009-2010** | **Total** | **47** | **(36%)** | **84** | **(64%)** | **131** |
| | Basic science (animal) | 13 | (42%) | 18 | (58%) | 31 |
| | Basic science (human) | 14 | (37%) | 24 | (63%) | 38 |
| | Basic science (other) | 1 | (50%) | 1 | (50%) | 2 |
| | Clinical science | 19 | (32%) | 41 | (68%) | 60 |
| **2019-2020** | **Total** | **58** | **(40%)** | **88** | **(60%)** | **146** |
| | Basic science (animal) | 12 | (39%) | 19 | (61%) | 31 |
| | Basic science (human) | 9 | (33%) | 18 | (67%) | 27 |
| | Basic science (other) | 0 | (0%) | 8 | (100%) | 8 |
| | Clinical science | 37 | (46%) | 43 | (54%) | 80 |
| **Grand Total** | **All Periods** | **105** | **(38%)** | **172** | **(62%)** | **277** |

**Table A7(i)-7(iv) Frequency distribution of significant exact P-values results by article type and period (bins)**

*A7(i). Basic science (human)*

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| **2009-2010** | [0,0.01) | 10 | 38 | 26.3 |
| | [0.01,0.02) | 3 | 38 | 7.9 |
| | [0.02,0.03) | 3 | 38 | 7.9 |
| | [0.03,0.04) | 1 | 38 | 2.6 |
| | [0.04,0.05) | 5 | 38 | 13.2 |
| | [0.05,0.06) | 2 | 38 | 5.3 |
| | [0.06,1] | 14 | 38 | 36.8 |
| **2019-2020** | [0,0.01) | 6 | 27 | 22.2 |
| | [0.01,0.02) | 5 | 27 | 18.5 |
| | [0.02,0.03) | 3 | 27 | 11.1 |
| | [0.03,0.04) | 3 | 27 | 11.1 |
| | [0.04,0.05) | 1 | 27 | 3.7 |
| | [0.05,0.06) | 0 | 27 | 0.0 |
| | [0.06,1] | 9 | 27 | 33.3 |

*A7(ii). Basic science (animal)*

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| **2009-2010** | [0,0.01) | 11 | 31 | 35.5 |
| | [0.01,0.02) | 4 | 31 | 12.9 |
| | [0.02,0.03) | 2 | 31 | 6.5 |
| | [0.03,0.04) | 1 | 31 | 3.2 |
| | [0.04,0.05) | 0 | 31 | 0.0 |

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| | [0.05,0.06) | 0 | 31 | 0.0 |
| | [0.06,1] | 13 | 31 | 41.9 |
| **2019-2020** | [0,0.01) | 12 | 31 | 38.7 |
| | [0.01,0.02) | 0 | 31 | 0.0 |
| | [0.02,0.03) | 4 | 31 | 12.9 |
| | [0.03,0.04) | 0 | 31 | 0.0 |
| | [0.04,0.05) | 1 | 31 | 3.2 |
| | [0.05,0.06) | 3 | 31 | 9.7 |
| | [0.06,1] | 11 | 31 | 35.5 |

***A7(iii). Basic science (other)***

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| **2009-2010** | [0,0.01) | 0 | 2 | 0.0 |
| | [0.01,0.02) | 1 | 2 | 50.0 |
| | [0.02,0.03) | 0 | 2 | 0.0 |
| | [0.03,0.04) | 0 | 2 | 0.0 |
| | [0.04,0.05) | 0 | 2 | 0.0 |
| | [0.05,0.06) | 0 | 2 | 0.0 |
| | [0.06,1] | 1 | 2 | 50.0 |
| **2019-2020** | [0,0.01) | 5 | 8 | 62.5 |
| | [0.01,0.02) | 0 | 8 | 0.0 |
| | [0.02,0.03) | 0 | 8 | 0.0 |
| | [0.03,0.04) | 2 | 8 | 25.0 |
| | [0.04,0.05) | 0 | 8 | 0.0 |
| | [0.05,0.06) | 1 | 8 | 12.5 |

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| | [0.06,1] | 0 | 8 | 0.0 |

### *A7(iv). Clinical science*

| Period | P-value bin | Count | Total | Percent |
|---|---|---|---|---|
| **2009-2010** | [0,0.01) | 25 | 60 | 41.7 |
| | [0.01,0.02) | 8 | 60 | 13.3 |
| | [0.02,0.03) | 2 | 60 | 3.3 |
| | [0.03,0.04) | 0 | 60 | 0.0 |
| | [0.04,0.05) | 6 | 60 | 10.0 |
| | [0.05,0.06) | 1 | 60 | 1.7 |
| | [0.06,1] | 18 | 60 | 30.0 |
| **2019-2020** | [0,0.01) | 24 | 80 | 30.0 |
| | [0.01,0.02) | 5 | 80 | 6.2 |
| | [0.02,0.03) | 4 | 80 | 5.0 |
| | [0.03,0.04) | 8 | 80 | 10.0 |
| | [0.04,0.05) | 2 | 80 | 2.5 |
| | [0.05,0.06) | 2 | 80 | 2.5 |
| | [0.06,1] | 35 | 80 | 43.8 |

# Appendix IV: List of articles in study

| First Author | Publication Year | Volume Number | Issue Number | Page Numbers | doi |
|---|---|---|---|---|---|
| Ho T.W. *et. al.,* | 2009 | 141 | 1 | 19-24 | doi:10.1016/j.pain.2008.07.013 |
| Masheb R.M. *et. al.,* | 2009 | 141 | 1 | 31-40 | doi.10.1016/j.pain.2008.09.031 |
| Banik R.K. *et. al.,* | 2009 | 141 | 1 | 41-51 | doi:10.1016/j.pain.2008.10.004 |
| Wrigley P.J. *et. al.,* | 2009 | 141 | 1 | 52-59 | doi:10.1016/j.pain.2008.10.007 |
| Gazerani P, *et. al.,* | 2009 | 141 | 1 | 60-69 | doi:10.1016/j.pain.2008.10.005 |
| Wang S. *et. al.,* | 2009 | 141 | 1 | 97-103 | doi:10.1016/j.pain.2008.10.021 |
| Velduijzen D. et. al., | 2009 | 141 | 1 | 104-113 | doi:10.1016/j.pain.2008.10.027 |
| Walker S.M.*et. al.,* | 2009 | 141 | 1 | 79-87 | doi.10.1016/j.pain.2008.10.012 |
| Cabanero D. *et. al.,* | 2009 | 141 | 1 | 88-96 | doi.10.1016/j.pain.2008.10.011 |
| Sullivan M.J.*et. al.,* | 2009 | 141 | 1 | 70-78 | doi.10.1016/j.pain.2008.10.017 |
| Campbell C.M. et al., | 2009 | 141 | 1 | 114-118 | doi.10.1016/j.pain.2008.10.023 |
| Oliveira M.C.G. *et al.,* | 2009 | 141 | 1 | 127-134 | doi:10.1016/j.pain.2008.10.024 |
| Keeble J.E. *et al.,* | 2009 | 141 | 1 | 135-142 | doi:10.1016/j.pain.2008.10.025 |
| Sandor K. *et al.,* | 2009 | 141 | 1 | 143-150 | doi:10.1016/j.pain.2008.10.028 |
| Tucker K.J. *et al.,* | 2009 | 141 | 1 | 151-155 | doi:10.1016/j.pain.2008.10.029 |
| Olechowski C.J. et al., | 2009 | 141 | 1 | 156-164 | doi.10.1016/j.pain.2008.11.002 |
| Wollgarten-Hadamek I. *et al.,* | 2009 | 141 | 1 | 165-172 | doi:10.1016/j.pain.2008.11.008 |
| Hollins M. *et al.,* | 2009 | 141 | 3 | 215-221 | doi:10.1016/j.pain.2008.10.003 |
| McDougall J.J. *et al.,* | 2009 | 141 | 3 | 222-232 | doi:10.1016/j.pain.2008.10.013 |
| Knabl J. et al., | 2009 | 141 | 3 | 233-238 | doi:10.1016/j.pain.2008.10.015 |
| Brenchat A. et al., | 2009 | 141 | 3 | 239-247 | doi. 10.1016/j.pain.2008.11.009 |
| Wicksell R. K. *et al.,* | 2009 | 141 | 3 | 248-257 | doi:10.1016/j.pain.2008.11.006 |
| Ogawa A. *et al.,* | 2009 | 141 | 3 | 269-275 | doi:10.1016/j.pain.2008.11.015 |
| Trost Z. *et al.,* | 2009 | 141 | 3 | 276-282 | doi:10.1016/j.pain.2008.11.016 |
| Obara I. *et al.,* | 2009 | 141 | 3 | 283-291 | doi:10.1016/j.pain.2008.12.006 |
| Agostini S. *et al.,* | 2009 | 141 | 3 | 292-299 | doi:10.1016/j.pain.2008.12.007 |
| Honore P. *et al.,* | 2009 | 142 | 1 | 27-35 | doi:10.1016/j.pain.2008.11.004 |
| Mewes R. *et al.,* | 2009 | 142 | 1 | 36-41 | doi:10.1016/j.pain.2008.11.007 |
| Khatibi A. *et al.,* | 2009 | 142 | 1 | 42-47 | doi:10.1016/j.pain.2008.11.020 |
| Amaya F. *et al.,* | 2009 | 142 | 1 | 59-67 | doi:10.1016/j.pain.2008.11.013 |
| Frew A.K. *et al.,* | 2009 | 142 | 1 | 68-74 | doi:10.1016/j.pain.2008.11.014 |
| Peng H. et al., | 2009 | 142 | 1 | 75-88 | doi. 10.1016/j.pain.2008.11.017 |
| Kupers R. *et al.,* | 2009 | 142 | 1 | 89-93 | doi:10.1016/j.pain.2008.12.008 |
| Rezende R. *et al.,* | 2009 | 142 | 1 | 94-100 | doi:10.1016/j.pain.2008.12.005 |
| McGowan N. *et al.,* | 2009 | 142 | 1 | 101-107 | doi:10.1016/j.pain.2008.12.009 |
| Wang N. *et al.,* | 2009 | 142 | 1 | 108-115 | doi. 0.1016/j.pain.2008.12.016 |
| Cottrell J.A. *et al.,* | 2009 | 141 | 1 | 116-126 | doi.10.1016/j.pain.2008.12.019 |
| Maeda Y. *et al.,* | 2009 | 142 | 1 | 127-132 | doi:10.1016/j.pain.2008.12.021 |
| Poh K. *et al.,* | 2009 | 142 | 1 | 133-141 | doi.10.1016/j.pain.2008.12.016 |
| Vervoot T. et al., | 2009 | 142 | 1 | 142-148 | doi:10.1016/j.pain.2008.12.028 |
| Valmunen T. *et al.,* | 2009 | 142 | 1 | 149-158 | doi:10.1016/j.pain.2008.12.031 |
| MacDonald D. *et al.,* | 2009 | 142 | 3 | 183-188 | doi:10.1016/j.pain.2008.12.002 |
| France C.R. *et al.,* | 2009 | 142 | 3 | 189-193 | doi:10.1016/j.pain.2008.12.010 |

| | | | | | |
|---|---|---|---|---|---|
| Volinn E. *et al.,* | 2009 | 142 | 3 | 194-201 | doi:10.1016/j.pain.2008.12.017 |
| Dworkin R.H. *et al.,* | 2009 | 142 | 3 | 209-217 | doi:10.1016/j.pain.2008.12.022 |
| de Mos M. *et al.,* | 2009 | 142 | 3 | 218-224 | doi:10.1016/j.pain.2008.12.032 |
| Martenson M.E. *et al.,* | 2009 | 142 | 3 | 236-244 | doi:10.1016/j.pain.2009.01.011 |
| Lapirot O. *et al.,* | 2009 | 142 | 3 | 245-254 | doi:10.1016/j.pain.2009.01.015 |
| Liossi C. *et al.,* | 2009 | 142 | 3 | 255-263 | doi:10.1016/j.pain.2009.01.017 |
| Russel F.A. *et al.,* | 2009 | 142 | 3 | 264-274 | doi:10.1016/j.pain.2009.01.021 |
| Wang X. *et al.,* | 2009 | 142 | 3 | 275-283 | doi:10.1016/j.pain.2009.02.001 |
| Peng H. *et al.,* | 2009 | 143 | 1 | 12-20 | doi:10.1016/j.pain.2008.12.023 |
| Heneweer H. *et al.,* | 2009 | 143 | 1 | 21-25 | doi:10.1016/j.pain.2008.12.033 |
| Stanchev D. *et al.,* | 2009 | 143 | 1 | 26-36 | doi:10.1016/j.pain.2009.01.006 |
| Li L. *et al.,* | 2009 | 143 | 1 | 37-40 | doi:10.1016/j.pain.2009.01.005 |
| Yang R. *et al.,* | 2009 | 143 | 1 | 48-55 | doi:10.1016/j.pain.2009.01.020 |
| Li J. *et al.,* | 2009 | 143 | 1 | 56-64 | doi:10.1016/j.pain.2009.01.023 |
| Chen L. *et al.,* | 2009 | 143 | 1 | 65-70 | doi:10.1016/j.pain.2009.01.022 |
| Yamada M. *et al.,* | 2009 | 143 | 1 | 71-75 | doi:10.1016/j.pain.2009.01.028 |
| Tuveson B. *et al.,* | 2009 | 143 | 1 | 84-91 | doi:10.1016/j.pain.2009.02.002 |
| Jones G.T. *et al.,* | 2009 | 143 | 1 | 92-96 | doi:10.1016/j.pain.2009.02.003 |
| JantschH.H.F. *et al.,* | 2009 | 143 | 1 | 97-105 | doi:10.1016/j.pain.2009.02.004 |
| Hirsh A.T. *et al.,* | 2009 | 143 | 1 | 106-113 | doi:10.1016/j.pain.2009.02.005 |
| Sullivan M. *et al.,* | 2009 | 143 | 1 | 123-129 | doi:10.1016/j.pain.2009.02.011 |
| Gauthier L.R. *et al.,* | 2009 | 143 | 1 | 147-154 | doi:10.1016/j.pain.2009.02.009 |
| King C. D. *et. al.,* | 2009 | 143 | 3 | 172-178 | doi:10.1016/j.pain.2008.12.027 |
| Kraus A. *et. al.,* | 2009 | 143 | 3 | 179-185 | doi:10.1016/j.pain.2008.12.018 |
| Cruccu G. *et. al.,* | 2009 | 143 | 3 | 186-191 | doi:10.1016/j.pain.2008.12.026 |
| Hu B. *et. al.,* | 2009 | 143 | 3 | 206-212 | doi:10.1016/j.pain.2009.02.018 |
| Holroyd K.A. *et. al.,* | 2009 | 143 | 3 | 213-222 | doi:10.1016/j.pain.2009.02.019 |
| von Baeyer C.L. *et. al.,* | 2009 | 143 | 3 | 223-227 | doi:10.1016/j.pain.2009.03.002 |
| Tambeli C.H. *et. al.*, | 2009 | 143 | 3 | 228-232 | doi:10.1016/j.pain.2009.03.005 |
| Zhang G.H. *et. al.,* | 2009 | 143 | 3 | 233-237 | doi:10.1016/j.pain.2009.03.007 |
| Simon L.S. *et. al.,* | 2009 | 143 | 3 | 238-245 | doi:10.1016/j.pain.2009.03.008 |
| Kohr D. *et. al.,* | 2009 | 143 | 3 | 246-251 | doi:10.1016/j.pain.2009.03.009 |
| Entrena J.M. *et. al.,* | 2009 | 143 | 3 | 252-261 | doi:10.1016/j.pain.2009.03.011 |
| Hsu M.C *et. al.,* | 2009 | 143 | 3 | 262-267 | doi:10.1016/j.pain.2009.03.017 |
| Colloca L. *et al.,* | 2009 | 144 | 1 | 28-34 | doi:10.1016/j.pain.2009.01.033 |
| Taddio A. *et al.,* | 2009 | 144 | 1 | 43-48 | doi:10.1016/j.pain.2009.02.012 |
| Ward S.E. *et al.,* | 2009 | 144 | 1 | 49-56 | doi:10.1016/j.pain.2009.02.021 |
| Leite-Almeida H. *et al.,* | 2009 | 144 | 1 | 57-65 | doi:10.1016/j.pain.2009.02.024 |
| O'Neill S. *et al.,* | 2009 | 144 | 1 | 76-83 | doi:10.1016/j.pain.2009.03.014 |
| Park C. *et al.,* | 2009 | 144 | 1 | 84-94 | doi:10.1016/j.pain.2009.03.016 |
| Jensen K. B. *et al.,* | 2009 | 144 | 1 | 95-100 | doi:10.1016/j.pain.2009.03.018 |
| Zheng Z. *et al.,* | 2009 | 144 | 1 | 101-109 | doi:10.1016/j.pain.2009.03.021 |
| Hathway G.J. *et al.,* | 2009 | 144 | 1 | 110-118 | doi:10.1016/j.pain.2009.03.022 |
| Lotsch J.*et al.,* | 2009 | 144 | 1 | 119-124 | doi:10.1016/j.pain.2009.03.023 |
| Mongini F. *et al.,* | 2009 | 144 | 1 | 125-129 | doi:10.1016/j.pain.2009.03.026 |
| Jungehaenel D.U. *et al.,* | 2009 | 144 | 1 | 130-138 | doi:10.1016/j.pain.2009.03.025 |
| Strom V. *et al.,* | 2009 | 144 | 1 | 147-155 | doi:10.1016/j.pain.2009.04.002 |
| Di C. *et al.,* | 2009 | 144 | 1 | 156-161 | doi:10.1016/j.pain.2009.04.018 |
| Yang X. *et al.,* | 2009 | 144 | 1 | 162-169 | doi:10.1016/j.pain.2009.04.005 |

| | | | | | |
|---|---|---|---|---|---|
| Do C. *et al.,* | 2009 | 144 | 1 | 170-177 | doi:10.1016/j.pain.2009.04.010 |
| Sun Y. *et al.,* | 2009 | 144 | 1 | 178-186 | doi:10.1016/j.pain.2009.04.011 |
| Staaf S. *et al.,* | 2009 | 144 | 1 | 187-199 | doi:10.1016/j.pain.2009.04.013 |
| Vartainen N. *et al.,* | 2009 | 144 | 1 | 200-208 | doi:10.1016/j.pain.2009.04.015 |
| Annahazi A. *et al.,* | 2009 | 144 | 1 | 209-217 | doi:10.1016/j.pain.2009.04.017 |
| Paine P. *et al.,* | 2009 | 144 | 3 | 236-244 | doi:10.1016/j.pain.2009.02.022 |
| Attal N. *et al.,* | 2009 | 144 | 3 | 245-252 | doi:10.1016/j.pain.2009.03.024 |
| Fehrenbacher J.C. *et al.,* | 2009 | 144 | 3 | 253-261 | doi:10.1016/j.pain.2009.03.027 |
| Wei T. *et al.,* | 2009 | 144 | 3 | 278-286 | doi:10.1016/j.pain.2009.04.020 |
| Cao Y. *et al.,* | 2009 | 144 | 3 | 287-293 | doi:10.1016/j.pain.2009.04.029 |
| Moseley L. G. *et al.,* | 2009 | 144 | 3 | 314-319 | doi:10.1016/j.pain.2009.04.030 |
| DeLaTorre S. *et al.,* | 2009 | 144 | 3 | 320-328 | doi:10.1016/j.pain.2009.05.002 |
| Xu J. *et al.,* | 2009 | 144 | 3 | 329-339 | doi:10.1016/j.pain.2009.05.019 |
| Goffaux P. *et al.,* | 2009 | 145 | 1 | 18-23 | doi:10.1016/j.pain.2009.02.008 |
| Watson A. *et al.,* | 2009 | 145 | 1 | 24-30 | doi:10.1016/j.pain.2009.04.003 |
| Wideman T.H. *et al.,* | 2009 | 145 | 1 | 45-51 | doi:10.1016/j.pain.2009.04.022 |
| de Rooij A.M. *et al.,* | 2009 | 145 | 1 | 82-85 | doi:10.1016/j.pain.2009.05.015 |
| Staud R. *et al.,* | 2009 | 145 | 1 | 96-104 | doi:10.1016/j.pain.2009.05.020 |
| Truini A. *et al.,* | 2009 | 145 | 1 | 105-109 | doi:10.1016/j.pain.2009.05.023 |
| Wang Z. *et al.,* | 2009 | 145 | 1 | 120-128 | doi:10.1016/j.pain.2009.05.025 |
| Tiwari V. *et al.,* | 2009 | 145 | 1 | 129-135 | doi:10.1016/j.pain.2009.05.028 |
| Haack M. *et al.,* | 2009 | 145 | 1 | 136-141 | doi:10.1016/j.pain.2009.05.029 |
| Boettger M.K. *et al.,* | 2009 | 145 | 1 | 142-150 | doi:10.1016/j.pain.2009.06.006 |
| Litt M.D. *et al.,* | 2009 | 145 | 1 | 160-168 | doi:10.1016/j.pain.2009.06.003 |
| Elander J. *et al.,* | 2009 | 145 | 1 | 169-175 | doi:10.1016/j.pain.2009.06.004 |
| Centeno M.V. *et al.,* | 2009 | 145 | 1 | 176-183 | doi:10.1016/j.pain.2009.06.014 |
| Messinger R.B. *et al.,* | 2009 | 145 | 1 | 184-195 | doi:10.1016/j.pain.2009.06.012 |
| Hyphantis T. *et al.,* | 2009 | 145 | 1 | 196-203 | doi:10.1016/j.pain.2009.07.005 |
| Gassner M. *et al.,* | 2009 | 145 | 1 | 204-210 | doi:10.1016/j.pain.2009.06.021 |
| Pavone F. *et al.,* | 2009 | 145 | 1 | 219-229 | doi:10.1016/j.pain.2009.06.026 |
| Defrin R. *et al.,* | 2009 | 145 | 1 | 230-236 | doi:10.1016/j.pain.2009.06.028 |
| Torvin M. A. *et al.,* | 2009 | 145 | 1 | 237-245 | doi:10.1016/j.pain.2009.06.032 |
| McClish D. K. *et al.,* | 2009 | 145 | 1 | 246-251 | doi:10.1016/j.pain.2009.06.029 |
| Kvarstein G. *et al.,* | 2009 | 145 | 3 | 279-286 | doi:10.1016/j.pain.2009.05.001 |
| Weisner C.M. *et al.,* | 2009 | 145 | 3 | 287-293 | doi:10.1016/j.pain.2009.05.006 |
| Sigtermans M.J. *et al.,* | 2009 | 145 | 3 | 304-311 | doi:10.1016/j.pain.2009.06.023 |
| Truin M. *et al.,* | 2009 | 145 | 3 | 312-318 | doi:10.1016/j.pain.2009.06.034 |
| Stevens D.L. *et al.,* | 2009 | 145 | 3 | 319-324 | doi:10.1016/j.pain.2009.06.030 |
| Sullivan M.J.L. *et al.,* | 2009 | 145 | 3 | 325-331 | doi:10.1016/j.pain.2009.06.031 |
| Kim H.Y. *et al.,* | 2009 | 145 | 3 | 332-340 | doi:10.1016/j.pain.2009.06.035 |
| Sahbaie P. *et al.,* | 2009 | 145 | 3 | 341-349 | doi:10.1016/j.pain.2009.06.037 |
| Hashmi J.A. *et al.,* | 2009 | 145 | 3 | 350-357 | doi:10.1016/j.pain.2009.07.002 |
| Finch P.M. *et al.,* | 2009 | 146 | 1 | 18-25 | doi:10.1016/j.pain.2009.05.017 |
| Nagakura Y. *et al.,* | 2009 | 146 | 1 | 26-33 | doi:10.1016/j.pain.2009.05.024 |
| Baron R. *et al.,* | 2009 | 146 | 1 | 34-40 | doi:10.1016/j.pain.2009.06.001 |
| Zhou Q. *et al.,* | 2009 | 146 | 1 | 41-46 | doi:10.1016/j.pain.2009.06.017 |
| Tousignant-Laflamme Y. *et al.,* | 2009 | 146 | 1 | 47-55 | doi:10.1016/j.pain.2009.06.018 |
| Heckman B. D. *et al.,* | 2009 | 146 | 1 | 56-64 | doi:10.1016/j.pain.2009.06.019 |

| | | | | | |
|---|---|---|---|---|---|
| Ruscheweyh R. *et al.,* | 2009 | 146 | 1 | 65-74 | doi:10.1016/j.pain.2009.06.020 |
| Egea J. *et al.,* | 2009 | 146 | 1 | 75-83 | doi:10.1016/j.pain.2009.07.007 |
| Smith W.R. *et al.,* | 2009 | 146 | 1 | 91-98 | doi:10.1016/j.pain.2009.07.008 |
| Hong Y. *et al.,* | 2009 | 146 | 1 | 105-113 | doi:10.1016/j.pain.2009.07.015 |
| Ahn D. K. *et al.,* | 2009 | 146 | 1 | 114-120 | doi:10.1016/j.pain.2009.07.012 |
| Martin L. *et al.,* | 2009 | 146 | 1 | 121-129 | doi:10.1016/j.pain.2009.07.016 |
| Zhang Y. *et al.,* | 2009 | 146 | 1 | 130-140 | doi:10.1016/j.pain.2009.07.011 |
| Norcin M. *et al.,* | 2009 | 146 | 1 | 141-147 | doi:10.1016/j.pain.2009.07.017 |
| Vervoot T. *et al.,* | 2009 | 146 | 1 | 170-176 | doi:10.1016/j.pain.2009.07.021 |
| Slater R. *et al.,* | 2009 | 146 | 1 | 177-182 | doi:10.1016/j.pain.2009.07.022 |
| Li T. *et al.,* | 2009 | 146 | 1 | 183-193 | doi:10.1016/j.pain.2009.07.027 |
| Morton D.L. *et al.,* | 2009 | 146 | 1 | 194-198 | doi:10.1016/j.pain.2009.07.026 |
| Munts A. G. *et al.,* | 2009 | 146 | 1 | 199-204 | doi:10.1016/j.pain.2009.07.030 |
| Palermo T. M. *et al.,* | 2009 | 146 | 1 | 205-213 | doi:10.1016/j.pain.2009.07.034 |
| Davies A.J. *et al.,* | 2009 | 146 | 1 | 214-221 | doi:10.1016/j.pain.2009.07.038 |
| Rowbotham M.C. et al., | 2009 | 146 | 3 | 245-252 | doi:10.1016/j.pain.2009.06.013 |
| Chappell A.S. *et al.,* | 2009 | 146 | 3 | 253-260 | doi:10.1016/j.pain.2009.06.024 |
| Burns J.W. *et al.,* | 2009 | 146 | 3 | 276-282 | doi:10.1016/j.pain.2009.07.024 |
| Gemes G. *et al.,* | 2009 | 146 | 3 | 293-300 | doi:10.1016/j.pain.2009.07.040 |
| Jeske N. A. *et al.,* | 2009 | 146 | 3 | 301-307 | doi:10.1016/j.pain.2009.08.002 |
| Arsenault A. *et al.,* | 2009 | 146 | 3 | 308-314 | doi:10.1016/j.pain.2009.08.003 |
| Gau R. *et al.,* | 2009 | 146 | 3 | 315-324 | doi:10.1016/j.pain.2009.09.018 |
| Bertsche T. *et al.,* | 2009 | 147 | 1 | 20-28 | doi:10.1016/j.pain.2009.07.009 |
| Dong X. *et al.,* | 2009 | 147 | 1 | 36-45 | doi:10.1016/j.pain.2009.07.031 |
| Aouad M. *et al.,* | 2009 | 147 | 1 | 54-59 | doi:10.1016/j.pain.2009.08.001 |
| Ahsin S. *et al.,* | 2009 | 147 | 1 | 60-66 | doi:10.1016/j.pain.2009.08.004 |
| Weingarten T.N. *et al.,* | 2009 | 147 | 1 | 67-71 | doi:10.1016/j.pain.2009.08.021 |
| Pfau D.B. *et al.,* | 2009 | 147 | 1 | 72-83 | doi:10.1016/j.pain.2008.08.010 |
| Martins A.O. *et al.,* | 2009 | 147 | 1 | 84-90 | doi:10.1016/j.pain.2008.08.016 |
| Walker S.M.*et. al.,* | 2009 | 147 | 1 | 99-106 | doi:10.1016/j.pain.2009.08.017 |
| Schwartzman R. J. *et al.,* | 2009 | 147 | 1 | 107-115 | doi:10.1016/j.pain.2009.08.015 |
| Altis K. *et al.,* | 2009 | 147 | 1 | 116-121 | doi:10.1016/j.pain.2009.08.018 |
| Bellieni C.V. *et al.,* | 2009 | 147 | 1 | 128-131 | doi:10.1016/j.pain.2009.08.025 |
| Siniatchkin M. *et al.,* | 2009 | 147 | 1 | 132-140 | doi:10.1016/j.pain.2009.08.028 |
| Taddio A. *et al.,* | 2009 | 147 | 1 | 141-146 | doi:10.1016/j.pain.2009.08.027 |
| Richardson E.J. *et al.,* | 2009 | 147 | 1 | 147-152 | doi:10.1016/j.pain.2009.08.030 |
| Sakurai M. *et al.,* | 2009 | 147 | 1 | 165-174 | doi:10.1016/j.pain.2009.08.003 |
| OsikowicZ M. *et al.,* | 2009 | 147 | 1 | 175-186 | doi:10.1016/j.pain.2009.09.002 |
| Treister R. et al., | 2009 | 147 | 1 | 187-193 | doi:10.1016/j.pain.2009.09.001 |
| Fukuda K. *et al.,* | 2009 | 147 | 1 | 194-201 | doi:10.1016/j.pain.2009.09.004 |
| Xiao W.H. *et al.,* | 2009 | 147 | 1 | 202-209 | doi:10.1016/j.pain.2009.09.006 |
| Bobeck E.N. *et al.,* | 2009 | 147 | 1 | 210-216 | doi:10.1016/j.pain.2009.09.008 |
| Leonard G. *et al.,* | 2009 | 147 | 1 | 217-223 | doi:10.1016/j.pain.2009.09.009 |
| Nahmias F. *et al.,* | 2009 | 147 | 1 | 224-232 | doi:10.1016/j.pain.2009.09.016 |
| Ge H. *et al.,* | 2009 | 147 | 1 | 233-240 | doi:10.1016/j.pain.2009.09.019 |
| Song Z. *et al.,* | 2009 | 147 | 1 | 241-248 | doi:10.1016/j.pain.2009.09.020 |
| Cano A. *et al.,* | 2009 | 147 | 1 | 249-254 | doi:10.1016/j.pain.2009.09.023 |
| McAllister S.L. *et al.,* | 2009 | 147 | 1 | 255-264 | doi:10.1016/j.pain.2009.09.022 |
| Carlton S.M. *et al.,* | 2009 | 147 | 1 | 265-276 | doi:10.1016/j.pain.2009.09.030 |

| | | | | | |
|---|---|---|---|---|---|
| Li W. *et al.,* | 2009 | 147 | 1 | 277-286 | doi:10.1016/j.pain.2009.09.032 |
| Riley J.L. *et al.,* | 2019 | 160 | 1 | 95-101 | doi.org/10.1097/j.pain.0000000000001382 |
| Santa-Cecilia F.V. *et al.,* | 2019 | 160 | 1 | 102-116 | doi.org/10.1097/j.pain.0000000000001383 |
| De Gregorino D. *et al.,* | 2019 | 160 | 1 | 136-150 | doi.org/10.1097/j.pain.0000000000001386 |
| Buehlmann D. *et al.,* | 2019 | 160 | 1 | 151-159 | doi.org/10.1097/j.pain.0000000000001388 |
| Barragan-Iglesias P. *et al.,* | 2019 | 160 | 1 | 160-171 | doi.org/10.1097/j.pain.0000000000001387 |
| Annak O. *et al.,* | 2019 | 160 | 1 | 172-186 | doi.org/10.1097/j.pain.0000000000001393 |
| Kim J.A. *et al.,* | 2019 | 160 | 1 | 187-197 | doi.org/10.1097/j.pain.0000000000001391 |
| Shiue S. *et al.,* | 2019 | 160 | 1 | 210-223 | doi.org/10.1097/j.pain.0000000000001395 |
| Fernandez-Zafra T. *et al.,* | 2019 | 160 | 1 | 224-236 | doi.org/10.1097/j.pain.0000000000001394 |
| Mustonen L. *et al.,* | 2019 | 160 | 1 | 246-256 | doi.org/10.1097/j.pain.0000000000001398 |
| Alschuler K.N. *et al.,* | 2019 | 160 | 1 | 257-262 | doi.org/10.1097/j.pain.0000000000001397 |
| Hooten W. M. *et al.,* | 2019 | 160 | 1 | 263-268 | doi.org/10.1097/j.pain.0000000000001400 |
| O'brien M.S. *et al.,* | 2019 | 160 | 1 | 269-278 | doi.org/10.1097/j.pain.0000000000001399 |
| Reaux-Le G. *et al.,* | 2019 | 160 | 2 | 307-321 | doi.org/10.1097/j.pain.0000000000001419 |
| Vervoort T. *et al.,* | 2019 | 160 | 2 | 322-333 | doi.org/10.1097/j.pain.0000000000001402 |
| Liu S. *et al.,* | 2019 | 160 | 2 | 334-344 | doi.org/10.1097/j.pain.0000000000001414 |
| Rosen S. F. *et al.,* | 2019 | 160 | 2 | 358-366 | doi.org/10.1097/j.pain.0000000000001420 |
| Simons L.E. *et al.,* | 2019 | 160 | 2 | 367-374 | doi.org/10.1097/j.pain.0000000000001403 |
| Yuan J. *et al.,* | 2019 | 160 | 2 | 375-384 | doi.org/10.1097/j.pain.0000000000001405 |
| Vila-Pueyo M. *et al.,* | 2019 | 160 | 2 | 385-394 | doi.org/10.1097/j.pain.0000000000001421 |
| Brandow A. M. *et al.,* | 2019 | 160 | 2 | 407-416 | doi.org/10.1097/j.pain.0000000000001407 |
| Burkill S. *et al.,* | 2019 | 160 | 2 | 424-432 | doi.org/10.1097/j.pain.0000000000001429 |
| Dickie A.C. *et al.,* | 2019 | 160 | 2 | 442-462 | doi.org/10.1097/j.pain.0000000000001406 |
| Baskozos G. *et al.,* | 2019 | 160 | 2 | 463-485 | doi.org/10.1097/j.pain.0000000000001416 |
| Petersen K.K. *et al.,* | 2019 | 160 | 2 | 486-492 | doi.org/10.1097/j.pain.0000000000001427 |
| Green G. *et al.,* | 2019 | 160 | 2 | 493-500 | doi.org/10.1097/j.pain.0000000000001425 |
| Lee J. *et al.,* | 2019 | 160 | 3 | 550-560 | doi.org/10.1097/j.pain.0000000000001417 |
| Bortsov A.V. *et al.,* | 2019 | 160 | 3 | 561-568 | doi.org/10.1097/j.pain.0000000000001432 |
| Bernstein C.A. *et al.,* | 2019 | 160 | 3 | 569-578 | doi.org/10.1097/j.pain.0000000000001434 |
| Smith S.B. *et al.,* | 2019 | 160 | 3 | 579-591 | doi.org/10.1097/j.pain.0000000000001438 |
| Petersen J. *et al.,* | 2019 | 160 | 3 | 607-618 | doi.org/10.1097/j.pain.0000000000001440 |
| Appel C.K. *et al.,* | 2019 | 160 | 3 | 619-631 | doi.org/10.1097/j.pain.0000000000001442 |
| Bove G.M. *et al.,* | 2019 | 160 | 3 | 632-644 | doi.org/10.1097/j.pain.0000000000001443 |
| Schappacher K.A. *et al.,* | 2019 | 160 | 3 | 645-657 | doi.org/10.1097/j.pain.0000000000001444 |
| Burston J.J. *et al.,* | 2019 | 160 | 3 | 658-669 | doi.org/10.1097/j.pain.0000000000001445 |
| Mauck M.C. *et al.,* | 2019 | 160 | 3 | 670-675 | doi.org/10.1097/j.pain.0000000000001446 |
| Bruehl S. *et al.,* | 2019 | 160 | 3 | 676-687 | doi.org/10.1097/j.pain.0000000000001447 |
| Khasabova I.A. *et al.,* | 2019 | 160 | 3 | 688-701 | doi.org/10.1097/j.pain.0000000000001448 |
| Warwick C.A. *et al.,* | 2019 | 160 | 3 | 702-711 | doi.org/10.1097/j.pain.0000000000001449 |
| Han K. et al., | 2019 | 160 | 3 | 712-723 | doi.org/10.1097/j.pain.0000000000001450 |
| Northon S. et al., | 2019 | 160 | 3 | 724-733 | doi.org/10.1097/j.pain.0000000000001451 |
| Chen T. *et al.,* | 2019 | 160 | 3 | 734-741 | doi.org/10.1097/j.pain.0000000000001428 |
| Lukas A. *et al.,* | 2019 | 160 | 3 | 742-753 | doi.org/10.1097/j.pain.0000000000001430 |
| Tansley S.N. *et al.,* | 2019 | 160 | 4 | 784-792 | doi.org/10.1097/j.pain.0000000000001454 |
| Grundy L. *et al.,* | 2019 | 160 | 4 | 793-804 | doi.org/10.1097/j.pain.0000000000001453 |
| Cardoso-Cruz H. *et al.,* | 2019 | 160 | 4 | 805-823 | doi.org/10.1097/j.pain.0000000000001457 |
| Navratilova E. *et al.,* | 2019 | 160 | 4 | 824-832 | doi.org/10.1097/j.pain.0000000000001458 |
| Meints S.M. *et al.,* | 2019 | 160 | 4 | 833-843 | doi.org/10.1097/j.pain.0000000000001461 |

| | | | | | |
|---|---|---|---|---|---|
| Ritchie C. *et al.,* | 2019 | 160 | 4 | 844-851 | doi.org/10.1097/j.pain.0000000000001460 |
| van de Donk T. *et al.,* | 2019 | 160 | 4 | 860-869 | doi.org/10.1097/j.pain.0000000000001464 |
| Mifflin K.A. *et al.,* | 2019 | 160 | 4 | 870-881 | doi.org/10.1097/j.pain.0000000000001465 |
| Kramer S. *et al.,* | 2019 | 160 | 4 | 882-894 | doi.org/10.1097/j.pain.0000000000001467 |
| Sakurai Y. *et al.,* | 2019 | 160 | 4 | 895-907 | doi.org/10.1097/j.pain.0000000000001466 |
| Perez-Aranda A. *et al.,* | 2019 | 160 | 4 | 908-921 | doi.org/10.1097/j.pain.0000000000001468 |
| del Rivero T. *et al.,* | 2019 | 160 | 4 | 922-931 | doi.org/10.1097/j.pain.0000000000001470 |
| Kallewaard J.W. *et al.,* | 2019 | 160 | 4 | 945-953 | doi.org/10.1097/j.pain.0000000000001475 |
| Ray S. *et al.,* | 2019 | 160 | 4 | 954-964 | doi.org/10.1097/j.pain.0000000000001474 |
| Fischer S. *et al.,* | 2019 | 160 | 4 | 965-972 | doi.org/10.1097/j.pain.0000000000001473 |
| Kaplan C.M. *et al.,* | 2019 | 160 | 4 | 973-983 | doi.org/10.1097/j.pain.0000000000001480 |
| Nahman-Averbuch H. *et al.,* | 2019 | 160 | 5 | 1019-1028 | doi.org/10.1097/j.pain.0000000000001477 |
| Gewandter J.S. *et al.,* | 2019 | 160 | 5 | 1029-1036 | doi.org/10.1097/j.pain.0000000000001478 |
| Catuneanu A. *et al.,* | 2019 | 160 | 5 | 1037-1049 | doi.org/10.1097/j.pain.0000000000001483 |
| Bree D. *et al.,* | 2019 | 160 | 5 | 1050-1058 | doi.org/10.1097/j.pain.0000000000001481 |
| DaviesA.J. *et al.,* | 2019 | 160 | 5 | 1059-1069 | doi.org/10.1097/j.pain.0000000000001482 |
| Sapio M.R. *et al.,* | 2019 | 160 | 5 | 1070-1081 | doi.org/10.1097/j.pain.0000000000001485 |
| Zhang Z. *et al.,* | 2019 | 160 | 5 | 1082-1092 | doi.org/10.1097/j.pain.0000000000001489 |
| Attridge N. *et al.,* | 2019 | 160 | 5 | 1093-1102 | doi.org/10.1097/j.pain.0000000000001490 |
| Coppi E. *et al.,* | 2019 | 160 | 5 | 1103-1118 | doi.org/10.1097/j.pain.0000000000001488 |
| Cruz-Almeida Y. *et al.,* | 2019 | 160 | 5 | 1119-1130 | doi.org/10.1097/j.pain.0000000000001491 |
| Boada M. *et al.,* | 2019 | 160 | 5 | 1146-1155 | doi.org/10.1097/j.pain.0000000000001495 |
| Fatima M. *et al.,* | 2019 | 160 | 5 | 1166-1174 | doi.org/10.1097/j.pain.0000000000001499 |
| Kato J. *et al.,* | 2019 | 160 | 5 | 1175-1185 | doi.org/10.1097/j.pain.0000000000001501 |
| Pickering G. *et al.,* | 2019 | 160 | 5 | 1186-1195 | doi.org/10.1097/j.pain.0000000000001502 |
| Seminowicz D.A. *et al.,* | 2019 | 160 | 5 | 1196-1207 | doi.org/10.1097/j.pain.0000000000001498 |
| Chang Y. *et al.,* | 2019 | 160 | 5 | 1208-1223 | doi.org/10.1097/j.pain.0000000000001497 |
| Bazika-Gerasch B. *et al.,* | 2019 | 160 | 6 | 1261-1269 | doi.org/10.1097/j.pain.0000000000001520 |
| Harte S.E. *et al.,* | 2019 | 160 | 6 | 1270-1280 | doi.org/10.1097/j.pain.0000000000001505 |
| Barkai O. *et al.,* | 2019 | 160 | 6 | 1281-1296 | doi.org/10.1097/j.pain.0000000000001523 |
| Allen K.D. *et al.,* | 2019 | 160 | 6 | 1297-1307 | doi.org/10.1097/j.pain.0000000000001525 |
| Tu Y. *et al.,* | 2019 | 160 | 6 | 1308-1318 | doi.org/10.1097/j.pain.0000000000001507 |

| | | | | | |
|---|---|---|---|---|---|
| Trentalange M. *et al.,* | 2019 | 160 | 6 | 1319-1326 | doi.org/10.1097/j.pain.0000000000001509 |
| Meents J.E. *et al.,* | 2019 | 160 | 6 | 1327-1341 | doi.org/10.1097/j.pain.0000000000001511 |
| Okada S. *et al.,* | 2019 | 160 | 6 | 1342-1360 | doi.org/10.1097/j.pain.0000000000001521 |
| Freidin M. B. *et al.,* | 2019 | 160 | 6 | 1361-1373 | doi.org/10.1097/j.pain.0000000000001514 |
| Markman J. *et al.,* | 2019 | 160 | 6 | 1374-1382 | doi.org/10.1097/j.pain.0000000000001517 |
| Dagnino A. *et al.,* | 2019 | 160 | 6 | 1383-1401 | doi.org/10.1097/j.pain.0000000000001513 |
| Konig S. *et al.,* | 2019 | 160 | 6 | 1402-1409 | doi.org/10.1097/j.pain.0000000000001503 |
| Rigoard P. *et al.,* | 2019 | 160 | 6 | 1410-1420 | doi.org/10.1097/j.pain.0000000000001510 |
| Payne L. A. *et al.,* | 2019 | 160 | 6 | 1421-1430 | doi.org/10.1097/j.pain.0000000000001516 |
| Hardt S. *et al.,* | 2019 | 160 | 6 | 1431-1447 | doi.org/10.1097/j.pain.0000000000001518 |
| Hernandez-Reyes J.E. *et al.,* | 2019 | 160 | 6 | 1448-1458 | doi.org/10.1097/j.pain.0000000000001515 |
| Laumet G. *et al.,* | 2019 | 160 | 6 | 1459-1468 | doi.org/10.1097/j.pain.0000000000001512 |
| Mischkowski D. *et al.,* | 2019 | 160 | 6 | 1469-1481 | doi.org/10.1097/j.pain.0000000000001573 |
| Beh S.T. *et al.,* | 2019 | 160 | 7 | 1505-1521 | doi.org/10.1097/j.pain.0000000000001547 |
| Treister R. et al., | 2019 | 160 | 7 | 1522-1528 | doi.org/10.1097/j.pain.0000000000001528 |
| Bhatt R.R. *et al.,* | 2019 | 160 | 7 | 1529-1540 | doi.org/10.1097/j.pain.0000000000001532 |
| Farhangi M. *et al.,* | 2019 | 160 | 7 | 1541-1550 | doi.org/10.1097/j.pain.0000000000001552 |
| Benson S. *et al.,* | 2019 | 160 | 7 | 1562-1571 | doi.org/10.1097/j.pain.0000000000001546 |
| Ashton-James C.E. *et al.,* | 2019 | 160 | 7 | 1586-1593 | doi.org/10.1097/j.pain.0000000000001543 |
| Kim J. *et al.,* | 2019 | 160 | 7 | 1594-1605 | doi.org/10.1097/j.pain.0000000000001541 |
| Fujiki E.N. *et al.,* | 2019 | 160 | 7 | 1606-1613 | doi.org/10.1097/j.pain.0000000000001549 |
| Phelps C. E. et al., | 2019 | 160 | 7 | 1614-1621 | doi.org/10.1097/j.pain.0000000000001553 |
| Buch N.S. *et al.,* | 2019 | 160 | 7 | 1622-1633 | doi.org/10.1097/j.pain.0000000000001536 |
| Attridge N. *et al.,* | 2019 | 160 | 7 | 1662-1669 | doi.org/10.1097/j.pain.0000000000001542 |
| Ragachov A. *et al.,* | 2019 | 160 | 7 | 1670-1679 | doi.org/10.1097/j.pain.0000000000001545 |
| Boersma K. *et al.,* | 2019 | 160 | 8 | 1708-1718 | doi.org/10.1097/j.pain.0000000000001575 |
| Benson S. *et al.,* | 2019 | 160 | 8 | 1719-1728 | doi.org/10.1097/j.pain.0000000000001579 |
| Noel M. *et al.,* | 2019 | 160 | 8 | 1729-1739 | doi.org/10.1097/j.pain.0000000000001582 |

| | | | | | |
|---|---|---|---|---|---|
| Fu W. *et al.,* | 2019 | 160 | 8 | 1754-1765 | doi.org/10.1097/j.pain.0000000000001557 |
| Mueller A. et al., | 2019 | 160 | 8 | 1766-1780 | doi.org/10.1097/j.pain.0000000000001567 |
| Pogatzki-Zahn E.M. *et al.,* | 2019 | 160 | 8 | 1781-1793 | doi.org/10.1097/j.pain.0000000000001561 |
| Cowie A.M. *et al.,* | 2019 | 160 | 8 | 1794-1816 | doi.org/10.1097/j.pain.0000000000001555 |
| McBeth J. *et al.,* | 2019 | 160 | 8 | 1817-1823 | doi.org/10.1097/j.pain.0000000000001568 |
| Colloca L. *et al.,* | 2019 | 160 | 8 | 1824-1834 | doi.org/10.1097/j.pain.0000000000001578 |
| Hincker A. *et al.,* | 2019 | 160 | 8 | 1835-1846 | doi.org/10.1097/j.pain.0000000000001577 |
| Xie J.Y. *et al.,* | 2019 | 160 | 8 | 1847-1855 | doi.org/10.1097/j.pain.0000000000001569 |
| Stamer U.M. *et al.,* | 2019 | 160 | 8 | 1856-1865 | doi.org/10.1097/j.pain.0000000000001560 |
| Noel M. *et al.,* | 2019 | 160 | 8 | 1866-1875 | doi.org/10.1097/j.pain.0000000000001565 |
| Alvarez P. *et al.,* | 2019 | 160 | 8 | 1876-1882 | doi.org/10.1097/j.pain.0000000000001570 |
| Pas R. *et al.,* | 2019 | 160 | 8 | 1883-1890 | doi.org/10.1097/j.pain.0000000000001566 |
| Drummond P.D. *et al.,* | 2019 | 160 | 8 | 1891-1900 | doi.org/10.1097/j.pain.0000000000001574 |
| Eather A. *et al.,* | 2019 | 160 | 9 | 1954-1966 | doi.org/10.1097/j.pain.0000000000001581 |
| Leone C. *et al.,* | 2019 | 160 | 9 | 1967-1975 | doi.org/10.1097/j.pain.0000000000001584 |
| Krotov V. *et al.,* | 2019 | 160 | 9 | 1982-1988 | doi.org/10.1097/j.pain.0000000000001586 |
| Williams W.A. *et al.,* | 2019 | 160 | 9 | 1989-2003 | doi.org/10.1097/j.pain.0000000000001587 |
| Bednarska O. *et al.,* | 2019 | 160 | 9 | 2004-2012 | doi.org/10.1097/j.pain.0000000000001589 |
| Guo T. *et al.,* | 2019 | 160 | 9 | 2013-2027 | doi.org/10.1097/j.pain.0000000000001592 |
| Adler-Neal A.L. *et al.,* | 2019 | 160 | 9 | 2028-2035 | doi.org/10.1097/j.pain.0000000000001595 |
| Sannajust S. *et al.,* | 2019 | 160 | 9 | 2036-2049 | doi.org/10.1097/j.pain.0000000000001598 |
| Harada Y. *et al.,* | 2019 | 160 | 9 | 2050-2062 | doi.org/10.1097/j.pain.0000000000001596 |
| Horn-Hofmann C. *et al.,* | 2019 | 160 | 9 | 2063-2071 | doi.org/10.1097/j.pain.0000000000001597 |
| Lopez-Sola M. *et al.,* | 2019 | 160 | 9 | 2072-2085 | doi.org/10.1097/j.pain.0000000000001590 |
| Curtis A.F. *et al.,* | 2019 | 160 | 9 | 2086-2092 | doi.org/10.1097/j.pain.0000000000001600 |
| Gierthmuhlen J. *et al.,* | 2019 | 160 | 9 | 2093-2104 | doi.org/10.1097/j.pain.0000000000001601 |
| Yatziv S. *et al.,* | 2019 | 160 | 9 | 2105-2114 | doi.org/10.1097/j.pain.0000000000001602 |
| Summers S.J. *et al.,* | 2019 | 160 | 9 | 2115-2125 | doi.org/10.1097/j.pain.0000000000001604 |

| | | | | | |
|---|---|---|---|---|---|
| Banerjee G. *et al.,* | 2019 | 160 | 9 | 2126-2135 | doi.org/10.1097/j.pain.0000000000001606 |
| Cropper H. *et al.,* | 2019 | 160 | 9 | 2136-2148 | doi.org/10.1097/j.pain.0000000000001607 |
| Matsuoka Y. *et al.,* | 2019 | 160 | 9 | 2149-2160 | doi.org/10.1097/j.pain.0000000000001611 |
| Jackson T. *et al.,* | 2019 | 160 | 10 | 2221-2228 | doi.org/10.1097/j.pain.0000000000001614 |
| Hirsh A.T. et al., | 2019 | 160 | 10 | 2229-2240 | doi.org/10.1097/j.pain.0000000000001634 |
| Barriere D.A. *et al.,* | 2019 | 160 | 10 | 2241-2254 | doi.org/10.1097/j.pain.0000000000001610 |
| Stokes A. *et al.,* | 2019 | 160 | 10 | 2255-2262 | doi.org/10.1097/j.pain.0000000000001612 |
| Kringel D. *et al.,* | 2019 | 160 | 10 | 2263-2277 | doi.org/10.1097/j.pain.0000000000001616 |
| Dietz C. *et al.,* | 2019 | 160 | 10 | 2278-2289 | doi.org/10.1097/j.pain.0000000000001617 |
| Geers A.L. *et al.,* | 2019 | 160 | 10 | 2290-2297 | doi.org/10.1097/j.pain.0000000000001618 |
| Hodkinson D.L. *et al.,* | 2019 | 160 | 10 | 2298-2304 | doi.org/10.1097/j.pain.0000000000001619 |
| Nocchi L. *et al.,* | 2019 | 160 | 10 | 2305-2315 | doi.org/10.1097/j.pain.0000000000001620 |
| Held M. *et al.,* | 2019 | 160 | 10 | 2316-2327 | doi.org/10.1097/j.pain.0000000000001623 |
| Bruehl S. *et al.,* | 2019 | 160 | 10 | 2328-2337 | doi.org/10.1097/j.pain.0000000000001624 |
| Matthewson G.M. *et al.,* | 2019 | 160 | 10 | 2338-2349 | doi.org/10.1097/j.pain.0000000000001621 |
| Sharpe E.E. *et al.,* | 2019 | 160 | 10 | 2350-2357 | doi.org/10.1097/j.pain.0000000000001628 |
| Wild J. *et al.,* | 2019 | 160 | 10 | 2358-2364 | doi.org/10.1097/j.pain.0000000000001629 |
| Khan J.S. *et al.,* | 2019 | 160 | 10 | 2374-2379 | doi.org/10.1097/j.pain.0000000000001631 |
| Zhang F. *et al.,* | 2019 | 160 | 10 | 2398-2408 | doi.org/10.1097/j.pain.0000000000001633 |
| Chen L. *et al.,* | 2019 | 160 | 11 | 2456-2463 | doi.org/10.1097/j.pain.0000000000001651 |
| Gradinger T. *et al.,* | 2019 | 160 | 11 | 2487-2496 | doi.org/10.1097/j.pain.0000000000001641 |
| Dull M.M. *et al.,* | 2019 | 160 | 11 | 2497-2507 | doi.org/10.1097/j.pain.0000000000001644 |
| Perez-Aranda A. *et al.,* | 2019 | 160 | 11 | 2508-2523 | doi.org/10.1097/j.pain.0000000000001655 |
| Taylor N.E. *et al.,* | 2019 | 160 | 11 | 2524-2534 | doi.org/10.1097/j.pain.0000000000001647 |
| Bernard C. *et al.,* | 2019 | 160 | 11 | 2535-2543 | doi.org/10.1097/j.pain.0000000000001637 |
| Higgins K.S. *et al.,* | 2019 | 160 | 11 | 2544-2553 | doi.org/10.1097/j.pain.0000000000001658 |
| Lee M.C. *et al.,* | 2019 | 160 | 11 | 2554-2565 | doi.org/10.1097/j.pain.0000000000001638 |
| Ge P. *et al.,* | 2019 | 160 | 11 | 2566-2579 | doi.org/10.1097/j.pain.0000000000001657 |

| | | | | | |
|---|---|---|---|---|---|
| Campbel-Yeo M. *et al.,* | 2019 | 160 | 11 | 2580-2588 | doi.org/10.1097/j.pain.0000000000001646 |
| Minerbi A. *et al.,* | 2019 | 160 | 11 | 2589-2602 | doi.org/10.1097/j.pain.0000000000001640 |
| Ericson H. *et al.,* | 2019 | 160 | 11 | 2603-2611 | doi.org/10.1097/j.pain.0000000000001649 |
| Cavaleri R. *et al.,* | 2019 | 160 | 11 | 2624-2633 | doi.org/10.1097/j.pain.0000000000001656 |
| Green-Fulgham S.M. *et al.,* | 2019 | 160 | 11 | 2634-2640 | doi.org/10.1097/j.pain.0000000000001652 |
| Weyer-Menkhoff I. *et al.,* | 2019 | 160 | 12 | 2699-2709 | doi.org/10.1097/j.pain.0000000000001660 |
| Liu S. et al., | 2019 | 160 | 12 | 2710-2723 | doi.org/10.1097/j.pain.0000000000001671 |
| Georgieva M. *et al.,* | 2019 | 160 | 12 | 2724-2742 | doi.org/10.1097/j.pain.0000000000001670 |
| Cracco E. *et al.,* | 2019 | 160 | 12 | 2743-2750 | doi.org/10.1097/j.pain.0000000000001665 |
| Lee M.J. *et al.,* | 2019 | 160 | 12 | 2776-2786 | doi.org/10.1097/j.pain.0000000000001676 |
| Tiemann L. et al., | 2019 | 160 | 12 | 2811-2818 | doi.org/10.1097/j.pain.0000000000001661 |
| Huang S. *et al.,* | 2019 | 160 | 12 | 2829-2840 | doi.org/10.1097/j.pain.0000000000001674 |
| Martire L.M. *et al.,* | 2019 | 160 | 12 | 2841-2847 | doi.org/10.1097/j.pain.0000000000001673 |
| Schreijenberg M. *et al.,* | 2019 | 160 | 12 | 2848-2854 | doi.org/10.1097/j.pain.0000000000001685 |
| Cuhadar U. *et al.,* | 2019 | 160 | 12 | 2855-2865 | doi.org/10.1097/j.pain.0000000000001662 |
| McPhee M.E. *et al.,* | 2019 | 160 | 12 | 2866-2876 | doi.org/10.1097/j.pain.0000000000001679 |
| Ma J. *et al.,* | 2019 | 160 | 12 | 2877-2890 | doi.org/10.1097/j.pain.0000000000001667 |
| Kleine-Borgmann J. *et al.,* | 2019 | 160 | 12 | 2891-2897 | doi.org/10.1097/j.pain.0000000000001683 |
| Turner J.A. *et al.,* | 2010 | 148 | 1 | 14-25 | doi:10.1016/j.pain.2009.08.014 |
| Telleria-Diaz A. *et al.,* | 2010 | 148 | 1 | 26 - 35 | doi:10.1016/j.pain.2009.08.013 |
| Conte A. *et al.,* | 2010 | 148 | 1 | 43-48 | doi:10.1016/j.pain.2009.09.031 |
| Piche M. *et al.,* | 2010 | 148 | 1 | 49-58 | doi:10.1016/j.pain.2009.10.005 |
| DeSantana J. *et al.,* | 2010 | 148 | 1 | 84-93 | doi:10.1016/j.pain.2009.10.011 |
| Radtke C.*et al.,* | 2010 | 148 | 1 | 94-100 | doi:10.1016/j.pain.2009.10.014 |
| Lotsch J.*et al.,* | 2010 | 148 | 1 | 103-106 | doi:10.1016/j.pain.2009.10.021 |
| Hu E. *et al.,* | 2010 | 148 | 1 | 107-113 | doi:10.1016/j.pain.2009.10.026 |
| Platon B. *et al.,* | 2010 | 148 | 1 | 114-119 | doi:10.1016/j.pain.2009.10.023 |
| Schutze R. *et al.,* | 2010 | 148 | 1 | 120-127 | doi:10.1016/j.pain.2009.10.030 |
| McCracken L.M. *et al.,* | 2010 | 148 | 1 | 141-147 | doi:10.1016/j.pain.2009.10.034 |
| Zhou L. *et al.,* | 2010 | 148 | 1 | 148-157 | doi:10.1016/j.pain.2009.10.032 |
| Chen J. *et al.,* | 2010 | 148 | 1 | 158-166 | doi:10.1016/j.pain.2009.11.004 |
| Sluka K.A.*et al.,* | 2010 | 148 | 2 | 188-197 | doi:10.1016/j.pain.2009.07.001 |
| Stosser S. *et al.,* | 2010 | 148 | 2 | 206-214 | doi:10.1016/j.pain.2009.09.024 |
| Schnakers C. *et al.,* | 2010 | 148 | 2 | 215-219 | doi:10.1016/j.pain.2009.09.028 |
| Zanette G. *et al.,* | 2010 | 148 | 2 | 227-236 | doi:10.1016/j.pain.2009.10.025 |

| | | | | | |
|---|---|---|---|---|---|
| Weyerbacher A.R. *et al.,* | 2010 | 148 | 2 | 237-246 | doi:10.1016/j.pain.2009.11.003 |
| Pinto L.G. *et al.,* | 2010 | 148 | 2 | 247-256 | doi:10.1016/j.pain.2009.11.006 |
| Wang F. *et al.,* | 2010 | 148 | 2 | 275-283 | doi:10.1016/j.pain.2009.11.011 |
| Hu Y. *et al.,* | 2010 | 148 | 2 | 284-301 | doi:10.1016/j.pain.2009.11.014 |
| Asiedu M. *et al.,* | 2010 | 148 | 2 | 302-308 | doi:10.1016/j.pain.2009.11.015 |
| Gao Y. *et al.,* | 2010 | 148 | 2 | 309-319 | doi:10.1016/j.pain.2009.11.017 |
| Rhudy J.L. *et al.,* | 2010 | 148 | 2 | 320-327 | doi:10.1016/j.pain.2009.11.018 |
| Hoschouer E.L. et al., | 2010 | 148 | 2 | 328-342 | doi:10.1016/j.pain.2009.11.023 |
| Krummenacher P. *et al.,* | 2010 | 148 | 3 | 368-374 | doi:10.1016/j.pain.2009.09.033 |
| Owen D.G. *et al.,* | 2010 | 148 | 3 | 375-386 | doi:10.1016/j.pain.2009.10.003 |
| Foster N.E. *et al.,* | 2010 | 148 | 3 | 398-406 | doi:10.1016/j.pain.2009.11.002 |
| Rukwied R. *et al.,* | 2010 | 148 | 3 | 407-413 | doi:10.1016/j.pain.2009.11.022 |
| Lu N. *et al.,* | 2010 | 148 | 3 | 414-425 | doi:10.1016/j.pain.2009.11.025 |
| Nishino T. *et al.,* | 2010 | 148 | 3 | 426-430 | doi:10.1016/j.pain.2009.11.024 |
| da Costa D.S.N. *et al.,* | 2010 | 148 | 3 | 431-437 | doi:10.1016/j.pain.2009.12.002 |
| Gustin S.M. *et al.,* | 2010 | 148 | 3 | 438-445 | doi:10.1016/j.pain.2009.12.001 |
| Buccafusco J.T. *et al.,* | 2010 | 148 | 3 | 446-453 | doi:10.1016/j.pain.2009.12.003 |
| Chen X. *et al.,* | 2010 | 148 | 3 | 462-472 | doi:10.1016/j.pain.2009.12.006 |
| Svensson P. *et al.,* | 2010 | 148 | 3 | 473-480 | doi:10.1016/j.pain.2009.12.009 |
| Lundqvist C. *et al.,* | 2010 | 148 | 3 | 487-491 | doi:10.1016/j.pain.2009.12.010 |
| Pozza D.H. *et al.,* | 2010 | 148 | 3 | 492-502 | doi:10.1016/j.pain.2009.12.013 |
| Rennefeld C. et al., | 2010 | 148 | 3 | 503-508 | doi:10.1016/j.pain.2009.12.014 |
| Fu E.S. *et al.,* | 2010 | 148 | 3 | 509-518 | doi:10.1016/j.pain.2010.01.001 |
| Zautra A.J. et al., | 2010 | 149 | 1 | Dec 18 | doi:10.1016/j.pain.2009.10.001 |
| Buhle J. et al., | 2010 | 149 | 1 | 19-26 | doi:10.1016/j.pain.2009.10.027 |
| Gremeau-Richard C. *et al.,* | 2010 | 149 | 1 | 27-32 | doi:10.1016/j.pain.2009.11.016 |
| Bruehl S. *et al.,* | 2010 | 149 | 1 | 57-63 | doi:10.1016/j.pain.2010.01.009 |
| Linnman C. *et al.,* | 2010 | 149 | 1 | 64-70 | doi:10.1016/j.pain.2010.01.008 |
| Saade N. E. *et al.,* | 2010 | 149 | 1 | 89-99 | doi:10.1016/j.pain.2010.01.014 |
| Goldman R.H. *et al.,* | 2010 | 149 | 1 | 117-123 | doi:10.1016/j.pain.2010.01.016 |
| Peng B. *et al.,* | 2010 | 149 | 1 | 124-129 | doi:10.1016/j.pain.2010.01.021 |
| Boettger M.K. *et al.,* | 2010 | 149 | 1 | 130-134 | doi:10.1016/j.pain.2010.01.022 |
| Sorkin L.S. *et al.,* | 2010 | 149 | 1 | 135-142 | doi:10.1016/j.pain.2010.01.024 |
| Miletic G. *et al.,* | 2010 | 149 | 1 | 152-159 | doi:10.1016/j.pain.2010.01.001 |
| Martin A. L. *et al.,* | 2010 | 149 | 2 | 208-215 | doi:10.1016/j.pain.2009.12.004 |
| Bailey B. *et al.,* | 2010 | 149 | 2 | 216-221 | doi:10.1016/j.pain.2009.12.008 |
| Younger J. W. *et al.,* | 2010 | 149 | 2 | 222-228 | doi:10.1016/j.pain.2010.01.006 |
| Verhoeven K. *et al.,* | 2010 | 149 | 2 | 229-234 | doi:10.1016/j.pain.2010.01.019 |
| Okamoto K. *et al.,* | 2010 | 149 | 2 | 235-242 | doi:10.1016/j.pain.2010.02.004 |
| Quang P. N. *et al.,* | 2010 | 149 | 2 | 254-262 | doi:10.1016/j.pain.2010.02.009 |
| Lam D. K. et al., | 2010 | 149 | 2 | 263-272 | doi:10.1016/j.pain.2010.02.010 |
| Hosoi M. *et al.,* | 2010 | 149 | 2 | 273-277 | doi:10.1016/j.pain.2010.02.012 |
| McWilliams L.A. *et al.,* | 2010 | 149 | 2 | 278-283 | doi:10.1016/j.pain.2010.02.019 |
| Dessem D. *et al.,* | 2010 | 149 | 2 | 284-295 | doi:10.1016/j.pain.2010.02.022 |
| Diers M. *et al.,* | 2010 | 149 | 2 | 296-304 | doi:10.1016/j.pain.2010.02.020 |
| Kiguchi N. *et al.,* | 2010 | 149 | 2 | 305-315 | doi:10.1016/j.pain.2010.02.025 |
| Carmichael N. M. E. *et al.,* | 2010 | 149 | 2 | 316-324 | doi:10.1016/j.pain.2010.02.024 |
| Quartana P.J. *et al.,* | 2010 | 149 | 2 | 325-331 | doi:10.1016/j.pain.2010.02.029 |
| van Laarhoven A.I.M. *et al.,* | 2010 | 149 | 2 | 332-337 | doi:10.1016/j.pain.2010.02.026 |

| | | | | | |
|---|---|---|---|---|---|
| Aggarwal V.R. *et al.,* | 2010 | 149 | 2 | 354-359 | doi:10.1016/j.pain.2010.02.040 |
| Moore A.R. *et al.,* | 2010 | 149 | 2 | 360-364 | doi:10.1016/j.pain.2010.02.039 |
| Rhudy J.L. *et al.,* | 2010 | 149 | 2 | 365-372 | doi:10.1016/j.pain.2010.02.041 |
| Geisser M.E. *et al.,* | 2010 | 149 | 2 | 373-378 | doi:10.1016/j.pain.2010.02.043 |
| Pazzaglia C. *et al.,* | 2010 | 149 | 2 | 379-385 | doi:10.1016/j.pain.2010.03.001 |
| McNamee K.E. *et al.,* | 2010 | 149 | 2 | 386-392 | doi:10.1016/j.pain.2010.03.002 |
| Hori K. *et al.,* | 2010 | 149 | 2 | 393-405 | doi:10.1016/j.pain.2010.03.005 |
| Fish R.A. *et al.,* | 2010 | 149 | 3 | 435-443 | doi:10.1016/j.pain.2009.12.016 |
| Christiansen S. *et al.,* | 2010 | 149 | 3 | 444-452 | doi:10.1016/j.pain.2009.12.015 |
| Lewis J. S. et al., | 2010 | 149 | 3 | 463-469 | doi:10.1016/j.pain.2010.02.007 |
| Brotz D. *et al.,* | 2010 | 149 | 3 | 470-475 | doi:10.1016/j.pain.2010.02.015 |
| Brenchat A. *et al.,* | 2010 | 149 | 3 | 483-494 | doi:10.1016/j.pain.2010.03.007 |
| Mhalla A. *et al.,* | 2010 | 149 | 3 | 495-500 | doi:10.1016/j.pain.2010.03.009 |
| Sacharczuk M. *et al.,* | 2010 | 149 | 3 | 506-513 | doi:10.1016/j.pain.2010.03.017 |
| Shibasaki M. *et al.,* | 2010 | 149 | 3 | 514-521 | doi:10.1016/j.pain.2010.03.023 |
| Bates B.D. *et al.,* | 2010 | 149 | 3 | 522-528 | doi:10.1016/j.pain.2010.03.024 |
| Weigelt A. *et al.,* | 2010 | 149 | 3 | 529-538 | doi:10.1016/j.pain.2010.03.027 |
| Kuehl L.K. *et al.,* | 2010 | 149 | 3 | 539-546 | doi:10.1016/j.pain.2010.03.026 |
| Tang N.K.Y. *et al.,* | 2010 | 149 | 3 | 547-554 | doi:10.1016/j.pain.2010.03.028 |
| Mickle A. *et al.,* | 2010 | 149 | 3 | 555-564 | doi:10.1016/j.pain.2010.03.030 |
| Olechowski C.J. *et al.,* | 2010 | 149 | 3 | 565-572 | doi:10.1016/j.pain.2010.03.037 |
| Arendt-Nielsen L. *et al.,* | 2010 | 149 | 3 | 573-581 | doi:10.1016/j.pain.2010.04.003 |
| Joseph E.K. *et al.,* | 2010 | 150 | 1 | 17-21 | doi:10.1016/j.pain.2010.02.011 |
| Kim H.Y. *et al.,* | 2010 | 150 | 1 | 29-40 | doi:10.1016/j.pain.2010.02.016 |
| Walton K.D. *et al.,* | 2010 | 150 | 1 | 41-51 | doi:10.1016/j.pain.2010.02.023 |
| McMurtry M.C. *et al.,* | 2010 | 150 | 1 | 52-58 | doi:10.1016/j.pain.2010.02.021 |
| Petrovic P. *et al.,* | 2010 | 150 | 1 | 59-65 | doi:10.1016/j.pain.2010.03.011 |
| Klinger R. *et al.,* | 2010 | 150 | 1 | 66-74 | doi:10.1016/j.pain.2010.03.036 |
| Petersen K.L. *et al.,* | 2010 | 150 | 1 | 75-82 | doi:10.1016/j.pain.2010.04.002 |
| Petersen K.L. et al., | 2010 | 150 | 1 | 83-92 | doi:10.1016/j.pain.2010.04.005 |
| Tran T.D. *et al.,* | 2010 | 150 | 1 | 93-102 | doi:10.1016/j.pain.2010.04.001 |
| Santha P. *et al.,* | 2010 | 150 | 1 | 103-112 | doi:10.1016/j.pain.2010.04.006 |
| Moont R. *et al.,* | 2010 | 150 | 1 | 113-120 | doi:10.1016/j.pain.2010.04.009 |
| Ferrini F. *et al.,* | 2010 | 150 | 1 | 128-140 | doi:10.1016/j.pain.2010.04.016 |
| Bishop T. *et al.,* | 2010 | 150 | 1 | 141-152 | doi:10.1016/j.pain.2010.04.018 |
| Riley J.L. *et al.,* | 2010 | 150 | 1 | 153-160 | doi:10.1016/j.pain.2010.04.020 |
| Huijnen I.P.J. *et al.,* | 2010 | 150 | 1 | 161-166 | doi:10.1016/j.pain.2010.04.021 |
| Tsao H. *et al.,* | 2010 | 150 | 1 | 167-172 | doi:10.1016/j.pain.2010.04.023 |
| Takahashi T. *et al.,* | 2010 | 150 | 1 | 183-191 | doi:10.1016/j.pain.2010.04.022 |
| Soler M.D. *et al.,* | 2010 | 150 | 1 | 192-198 | doi:10.1016/j.pain.2010.04.027 |
| Mouraux A. *et al.,* | 2010 | 150 | 1 | 199-207 | doi:10.1016/j.pain.2010.04.026 |
| Sumitani M. *et al.,* | 2010 | 150 | 2 | 243-249 | doi:10.1016/j.pain.2010.03.032 |
| Perrot S. *et al.,* | 2010 | 150 | 2 | 250-256 | doi:10.1016/j.pain.2010.03.034 |
| Hohmeister J. *et al.,* | 2010 | 150 | 2 | 257-267 | doi:10.1016/j.pain.2010.04.004 |
| Briggs A.M. *et al.,* | 2010 | 150 | 2 | 275-283 | doi:10.1016/j.pain.2010.04.031 |
| Braz J.M. *et al.,* | 2010 | 150 | 2 | 290-301 | doi:10.1016/j.pain.2010.05.005 |
| Tiemann L. *et al.,* | 2010 | 150 | 2 | 302-308 | doi:10.1016/j.pain.2010.05.014 |
| Puttfarcken P.S. *et al.,* | 2010 | 150 | 2 | 319-326 | doi:10.1016/j.pain.2010.05.015 |
| Matthew L. *et al.,* | 2010 | 150 | 2 | 327-331 | doi:10.1016/j.pain.2010.05.018 |

| | | | | | |
|---|---|---|---|---|---|
| Sullivan M.D. *et al.,* | 2010 | 150 | 2 | 332-339 | doi:10.1016/j.pain.2010.05.020 |
| Knowlton W.M. *et al.,* | 2010 | 150 | 2 | 340-350 | doi:10.1016/j.pain.2010.05.021 |
| Nijrolder I. *et al.,* | 2010 | 150 | 2 | 351-357 | doi:10.1016/j.pain.2010.05.030 |
| Rivat C. *et al.,* | 2010 | 150 | 2 | 358-368 | doi:10.1016/j.pain.2010.05.031 |
| Jamison R.N. *et al.,* | 2010 | 150 | 3 | 390-400 | doi:10.1016/j.pain.2010.02.033 |
| Horvath R.J. *et al.,* | 2010 | 150 | 3 | 401-413 | doi:10.1016/j.pain.2010.02.042 |
| Baron R. *et al.,* | 2010 | 150 | 3 | 420-427 | doi:10.1016/j.pain.2010.04.013 |
| Brown C.A. *et al.,* | 2010 | 150 | 3 | 428-438 | doi:10.1016/j.pain.2010.04.017 |
| Maier C. *et al.,* | 2010 | 150 | 3 | 439-450 | doi:10.1016/j.pain.2010.05.002 |
| Gulbrandsen P. *et al.,* | 2010 | 150 | 3 | 458-461 | doi:10.1016/j.pain.2010.05.024 |
| Tu C. *et al.,* | 2010 | 150 | 3 | 462-468 | doi:10.1016/j.pain.2010.05.026 |
| Sharpe L. *et al.,* | 2010 | 150 | 3 | 469-476 | doi:10.1016/j.pain.2010.05.027 |
| Ilfeld B.M. *et al.,* | 2010 | 150 | 3 | 477-484 | doi:10.1016/j.pain.2010.05.028 |
| de Andrade D.C. *et al.,* | 2010 | 150 | 3 | 485-491 | doi:10.1016/j.pain.2010.06.001 |
| Visscher C.M. *et al.,* | 2010 | 150 | 3 | 492-500 | doi:10.1016/j.pain.2010.06.002 |
| Sterling M. | 2010 | 150 | 3 | 501-506 | doi:10.1016/j.pain.2010.06.003 |
| Seery M.D. *et al.,* | 2010 | 150 | 3 | 507-515 | doi:10.1016/j.pain.2010.06.007 |
| Truin A. *et al.,* | 2010 | 150 | 3 | 516-521 | doi:10.1016/j.pain.2010.06.006 |
| Waehrens E.E. *et al.,* | 2010 | 150 | 3 | 535-541 | doi:10.1016/j.pain.2010.06.008 |
| Pacini A. *et al.,* | 2010 | 150 | 3 | 542-549 | doi:10.1016/j.pain.2010.06.014 |
| Willemen H.L.D.M. *et al.,* | 2010 | 150 | 3 | 550-560 | doi:10.1016/j.pain.2010.06.015 |
| Walker L.S. *et al.,* | 2010 | 150 | 3 | 568-572 | doi:10.1016/j.pain.2010.06.018 |
| Lee D.M. *et al.,* | 2010 | 151 | 1 | 30-36 | doi:10.1016/j.pain.2010.04.024 |
| Nes L.S. *et al.,* | 2010 | 151 | 1 | 37-44 | doi:10.1016/j.pain.2010.05.009 |
| Morch C.D. | 2010 | 151 | 1 | 45-52 | doi:10.1016/j.pain.2010.05.016 |
| Martinez V. *et al.,* | 2010 | 151 | 1 | 53-60 | doi:10.1016/j.pain.2010.05.017 |
| Gustin S.M. *et al.,* | 2010 | 151 | 1 | 69-76 | doi:10.1016/j.pain.2010.06.022 |
| Lannersten L. *et al.,* | 2010 | 151 | 1 | 77-86 | doi:10.1016/j.pain.2010.06.021 |
| Takasu K. *et al.,* | 2010 | 151 | 1 | 87-96 | doi:10.1016/j.pain.2010.06.020 |
| Aderjan D. *et al.,* | 2010 | 151 | 1 | 97-103 | doi:10.1016/j.pain.2010.06.024 |
| Eisenberg E. *et al.,* | 2010 | 151 | 1 | 104-109 | doi:10.1016/j.pain.2010.06.025 |
| Litt M.D. *et al.,* | 2010 | 151 | 1 | 110-116 | doi:10.1016/j.pain.2010.06.030 |
| Qiao L. *et al.,* | 2010 | 151 | 1 | 117-127 | doi:10.1016/j.pain.2010.06.029 |
| Akerman S. *et al.,* | 2010 | 151 | 1 | 128-136 | doi:10.1016/j.pain.2010.06.028 |
| Martins I. *et al.,* | 2010 | 151 | 1 | 137-145 | doi:10.1016/j.pain.2010.06.027 |
| Da Silver L.F.S. *et al.,* | 2010 | 151 | 1 | 155-161 | doi:10.1016/j.pain.2010.06.037 |
| Millecamps M. et al., | 2010 | 151 | 1 | 174-183 | doi:10.1016/j.pain.2010.07.003 |
| Densmore V.S. *et al.,* | 2010 | 151 | 1 | 184-193 | doi:10.1016/j.pain.2010.07.007 |
| Wang Z. *et al.,* | 2010 | 151 | 1 | 194-205 | doi:10.1016/j.pain.2010.07.006 |
| Leonard G. *et al.,* | 2010 | 151 | 1 | 215-219 | doi:10.1016/j.pain.2010.07.012 |
| Lewandowski A.S. *et al.,* | 2010 | 151 | 1 | 220-225 | doi:10.1016/j.pain.2010.07.016 |
| Atianjoh F.E. *et al.,* | 2010 | 151 | 1 | 226-234 | doi:10.1016/j.pain.2010.07.017 |
| Vardanyan M. *et al.,* | 2010 | 151 | 2 | 257-265 | doi:10.1016/j.pain.2010.05.022 |
| Giamberardino M.A. *et al.,* | 2010 | 151 | 2 | 307-322 | doi:10.1016/j.pain.2010.06.023 |
| Hashmi J.A. *et al.,* | 2010 | 151 | 2 | 323-329 | doi:10.1016/j.pain.2010.06.026 |
| Martel M.O. *et al.,* | 2010 | 151 | 2 | 330-336 | doi:10.1016/j.pain.2010.07.004 |
| Yu X.H. *et al.,* | 2010 | 151 | 2 | 337-344 | doi:10.1016/j.pain.2010.07.019 |
| Manjavachi M.N. *et al.,* | 2010 | 151 | 2 | 345-355 | doi:10.1016/j.pain.2010.07.018 |
| Hossaini M. *et al.,* | 2010 | 151 | 2 | 356-365 | doi:10.1016/j.pain.2010.07.023 |

| | | | | | |
|---|---|---|---|---|---|
| Borsook T.K. *et al.,* | 2010 | 151 | 2 | 372-377 | doi:10.1016/j.pain.2010.07.022 |
| Akiyama T. *et al.,* | 2010 | 151 | 2 | 378-383 | doi:10.1016/j.pain.2010.07.024 |
| Nash P.G. *et al.,* | 2010 | 151 | 2 | 384-393 | doi:10.1016/j.pain.2010.07.027 |
| Christianson C.A. *et al.,* | 2010 | 151 | 2 | 394-403 | doi:10.1016/j.pain.2010.07.030 |
| Fritsche G. *et al.,* | 2010 | 151 | 2 | 404-413 | doi:10.1016/j.pain.2010.07.032 |
| Quinn K.P. *et al.,* | 2010 | 151 | 2 | 414-421 | doi:10.1016/j.pain.2010.07.034 |
| Sheilds S.D. *et al.,* | 2010 | 151 | 2 | 422-429 | doi:10.1016/j.pain.2010.08.001 |
| Colloca l. *et al.,* | 2010 | 151 | 2 | 430-439 | doi:10.1016/j.pain.2010.08.007 |
| Sullivan M.J.L. *et al.,* | 2010 | 151 | 2 | 440-446 | doi:10.1016/j.pain.2010.08.005 |
| Chen X. et al., | 2010 | 151 | 2 | 460-466 | doi:10.1016/j.pain.2010.08.004 |
| Defrin R. *et al.,* | 2010 | 151 | 2 | 489-495 | doi:10.1016/j.pain.2010.08.009 |
| Fuchs D. *et al.,* | 2010 | 151 | 2 | 496-505 | doi:10.1016/j.pain.2010.08.010 |
| Lima F.O. *et al.,* | 2010 | 151 | 2 | 506-515 | doi:10.1016/j.pain.2010.08.014 |
| Meusel T. *et al.,* | 2010 | 151 | 2 | 516-521 | doi:10.1016/j.pain.2010.08.013 |
| Carson J.W. *et al.,* | 2010 | 151 | 2 | 530-539 | doi:10.1016/j.pain.2010.08.020 |
| Christianson J.A. *et al.,* | 2010 | 151 | 2 | 540-549 | doi:10.1016/j.pain.2010.08.029 |
| Taylor K.S. *et al.,* | 2010 | 151 | 3 | 582-591 | doi:10.1016/j.pain.2010.06.032 |
| Gazerani P. *et al.,* | 2010 | 151 | 3 | 606-616 | doi:10.1016/j.pain.2010.07.029 |
| Portenoy R.K. *et al.,* | 2010 | 151 | 3 | 617-624 | doi:10.1016/j.pain.2010.07.028 |
| Lu S. *et al.,* | 2010 | 151 | 3 | 633-643 | doi:10.1016/j.pain.2010.08.030 |
| Badr H. *et al.,* | 2010 | 151 | 3 | 644-645 | doi:10.1016/j.pain.2010.08.022 |
| Bosshard S.C *et al.,* | 2010 | 151 | 3 | 655-663 | doi:10.1016/j.pain.2010.08.025 |
| Amris K. *et al.,* | 2010 | 151 | 3 | 664-669 | doi:10.1016/j.pain.2010.08.023 |
| Baastrup C. *et al.,* | 2010 | 151 | 3 | 670-679 | doi:10.1016/j.pain.2010.08.024 |
| Sambo C. F. *et al.,* | 2010 | 151 | 3 | 687-693 | doi:10.1016/j.pain.2010.08.035 |
| Williams D. A. *et al.,* | 2010 | 151 | 3 | 694-702 | doi:10.1016/j.pain.2010.08.034 |
| Lautenbacher S. *et al.,* | 2010 | 151 | 3 | 722-731 | doi:10.1016/j.pain.2010.08.041 |
| Hashmi J.A. *et al.,* | 2010 | 151 | 3 | 737-743 | doi:10.1016/j.pain.2010.08.045 |
| Xu J. *et al.,* | 2010 | 151 | 3 | 744-755 | doi:10.1016/j.pain.2010.09.003 |
| Pitcher M.H. *et al.,* | 2010 | 151 | 3 | 756-762 | doi:10.1016/j.pain.2010.09.008 |
| Taguchi T. *et al.,* | 2010 | 151 | 3 | 771-782 | doi:10.1016/j.pain.2010.09.011 |
| Carlino E. *et al.,* | 2010 | 151 | 3 | 783-789 | doi:10.1016/j.pain.2010.09.013 |
| Neziri A.Y. *et al.,* | 2010 | 151 | 3 | 798-805 | doi:10.1016/j.pain.2010.09.07 |
| Lawson K.P. *et al.,* | 2010 | 151 | 3 | 806-815 | doi:10.1016/j.pain.2010.09.018 |
| Lui F. *et al.,* | 2010 | 151 | 3 | 816-824 | doi:10.1016/j.pain.2010.09.021 |
| Abrahamsen R. *et al.,* | 2010 | 151 | 3 | 825-833 | doi:10.1016/j.pain.2010.09.020 |
| Li W. *et al.,* | 2010 | 151 | 3 | 843-852 | doi:10.1016/j.pain.2010.09.026 |
| Gray M. A *et al.,* | 2010 | 151 | 3 | 853-861 | doi:10.1016/j.pain.2010.09.028 |
| Stokvis A. *et al.,* | 2010 | 151 | 3 | 862-869 | doi:10.1016/j.pain.2010.09.032 |
| Harden N.R. *et al.,* | 2010 | 151 | 3 | 870-876 | doi:10.1016/j.pain.2010.09.031 |
| Linnstaedt | 2020 | 161 | 1 | 47-60 | doi.org/10.1097/j.pain.0000000000001709 |
| Pomares F.B. *et al.,* | 2020 | 161 | 1 | 74-82 | doi.org/10.1097/j.pain.0000000000001707 |
| Liu L. *et al.,* | 2020 | 161 | 1 | 96-108 | doi.org/10.1097/j.pain.0000000000001681 |
| Zhou X. *et al.,* | 2020 | 161 | 1 | 127-134 | doi.org/10.1097/j.pain.0000000000001702 |
| Li J. *et al.,* | 2020 | 161 | 1 | 166-176 | doi.org/10.1097/j.pain.0000000000001691 |
| Stockstill K. et al., | 2020 | 161 | 1 | 177-184 | doi.org/10.1097/j.pain.0000000000001697 |
| Brierley S.M. *et al.,* | 2020 | 161 | 1 | 202-210 | doi.org/10.1097/j.pain.0000000000001692 |
| Wu P.Y. *et al.,* | 2020 | 161 | 1 | 211-219 | doi.org/10.1097/j.pain.0000000000001703 |
| Waller R. *et al.,* | 2020 | 161 | 1 | 220-229 | doi.org/10.1097/j.pain.0000000000001704 |

| | | | | | |
|---|---|---|---|---|---|
| Niedermayer S. *et al.,* | 2020 | 161 | 2 | 266-273 | doi.org/10.1097/j.pain.0000000000001713 |
| Mauck M.C. *et al.,* | 2020 | 161 | 2 | 274-280 | doi.org/10.1097/j.pain.0000000000001728 |
| Schlaeger J.M. *et al.,* | 2020 | 161 | 2 | 281-287 | doi.org/10.1097/j.pain.0000000000001716 |
| Li Y. *et al.,* | 2020 | 161 | 2 | 288-299 | doi.org/10.1097/j.pain.0000000000001725 |
| Lau J.Y.F. *et al.,* | 2020 | 161 | 2 | 319-327 | doi.org/10.1097/j.pain.0000000000001723 |
| Vanderclausen C. *et al.,* | 2020 | 161 | 2 | 328-337 | doi.org/10.1097/j.pain.0000000000001721 |
| Iordanova S. *et al.,* | 2020 | 161 | 2 | 338-350 | doi.org/10.1097/j.pain.0000000000001731 |
| Hosomi K. *et al.,* | 2020 | 161 | 2 | 351-360 | doi.org/10.1097/j.pain.0000000000001712 |
| Brown C.A. *et al.,* | 2020 | 161 | 2 | 369-378 | doi.org/10.1097/j.pain.0000000000001730 |
| Manion J. *et al.,* | 2020 | 161 | 2 | 379-387 | doi.org/10.1097/j.pain.0000000000001733 |
| Rhudy J.L. et al., | 2020 | 161 | 2 | 388-404 | doi.org/10.1097/j.pain.0000000000001715 |
| Potenzieri A. *et al.,* | 2020 | 161 | 2 | 405-415 | doi.org/10.1097/j.pain.0000000000001722 |
| Jin Y. *et al.,* | 2020 | 161 | 2 | 416-428 | doi.org/10.1097/j.pain.0000000000001724 |
| De Icco R. *et al.,* | 2020 | 161 | 2 | 429-438 | doi.org/10.1097/j.pain.0000000000001726 |
| Zhang X. *et al.,* | 2020 | 161 | 2 | 446-458 | doi.org/10.1097/j.pain.0000000000001734 |
| Simons L.E. *et al.,* | 2020 | 161 | 3 | 520-531 | doi.org/10.1097/j.pain.0000000000001735 |
| Wu M. *et al.,* | 2020 | 161 | 3 | 532-544 | doi.org/10.1097/j.pain.0000000000001738 |
| Gruener H. *et al.,* | 2020 | 161 | 3 | 545-556 | doi.org/10.1097/j.pain.0000000000001740 |
| Szikszay T.M. *et al.,* | 2020 | 161 | 3 | 557-564 | doi.org/10.1097/j.pain.0000000000001739 |
| Carleton R.N. *et al.,* | 2020 | 161 | 3 | 584-594 | doi.org/10.1097/j.pain.0000000000001746 |
| Azqueta-Gavaldon M. *et al.,* | 2020 | 161 | 3 | 595-608 | doi.org/10.1097/j.pain.0000000000001745 |
| Navratilova E. et al., | 2020 | 161 | 3 | 609-618 | doi.org/10.1097/j.pain.0000000000001748 |
| Heathcote L.C. *et al.,* | 2020 | 161 | 3 | 630-640 | doi.org/10.1097/j.pain.0000000000001753 |
| Soros P. *et al.,* | 2020 | 161 | 3 | 641-650 | doi.org/10.1097/j.pain.0000000000001756 |
| Quesada C. *et al.,* | 2020 | 161 | 4 | 718-728 | doi.org/10.1097/j.pain.0000000000001760 |
| De Pauw R. *et al.,* | 2020 | 161 | 4 | 729-741 | doi.org/10.1097/j.pain.0000000000001762 |
| Roquelaure Y. *et al.,* | 2020 | 161 | 4 | 749-757 | doi.org/10.1097/j.pain.0000000000001766 |
| Xu J. *et al.,* | 2020 | 161 | 4 | 758-772 | doi.org/10.1097/j.pain.0000000000001761 |
| Cibert-Goton V. *et al.,* | 2020 | 161 | 4 | 773-786 | doi.org/10.1097/j.pain.0000000000001764 |
| Heitmann H. *et al.,* | 2020 | 161 | 4 | 787-796 | doi.org/10.1097/j.pain.0000000000001767 |
| Guo T. et al., | 2020 | 161 | 4 | 797-809 | doi.org/10.1097/j.pain.0000000000001765 |
| Letzen J.E. *et al.,* | 2020 | 161 | 4 | 810-820 | doi.org/10.1097/j.pain.0000000000001759 |
| Tiwari V. *et al.,* | 2020 | 161 | 4 | 842-855 | doi.org/10.1097/j.pain.0000000000001768 |
| Ferrari L.F. *et al.,* | 2020 | 161 | 4 | 865-874 | doi:org/10.1097/j.pain.0000000000001798 |
| Tohyama S. *et al.,* | 2020 | 161 | 5 | 916-925 | doi.org/10.1097/j.pain.0000000000001777 |
| Suttoin K.S. *et al.,* | 2020 | 161 | 5 | 926-937 | doi.org/10.1097/j.pain.0000000000001778 |
| Zorina-Lichtenwalter K. *et al.,* | 2020 | 161 | 5 | 938-948 | doi.org/10.1097/j.pain.0000000000001780 |
| Boada M. et al., | 2020 | 161 | 5 | 949-959 | doi.org/10.1097/j.pain.0000000000001781 |
| Boada M. *et al.,* | 2020 | 161 | 5 | 960-969 | doi.org/10.1097/j.pain.0000000000001782 |
| Hayes C.J. *et al.,* | 2020 | 161 | 5 | 979-988 | doi.org/10.1097/j.pain.0000000000001784 |
| Wang H. *et al.,* | 2020 | 161 | 5 | 989-1004 | doi.org/10.1097/j.pain.0000000000001785 |
| Schain A.J. *et al.,* | 2020 | 161 | 5 | 1019-1026 | doi.org/10.1097/j.pain.0000000000001789 |
| Brenner D.M. *et al.,* | 2020 | 161 | 5 | 1027-1036 | doi.org/10.1097/j.pain.0000000000001754 |
| Melo-Carrillo A. *et al.,* | 2020 | 161 | 5 | 1037-1043 | doi.org/10.1097/j.pain.0000000000001791 |
| Quiton R.L.*et al.,* | 2020 | 161 | 5 | 1044-1051 | doi.org/10.1097/j.pain.0000000000001793 |

| | | | | | |
|---|---|---|---|---|---|
| Verma V. *et al.,* | 2020 | 161 | 5 | 1052-1064 | doi.org/10.1097/j.pain.0000000000001792 |
| Holden S. *et al.,* | 2020 | 161 | 5 | 1065-1071 | doi.org/10.1097/j.pain.0000000000001796 |
| Neville A. *et al.,* | 2020 | 161 | 5 | 1072-1082 | doi.org/10.1097/j.pain.0000000000001799 |
| Labastida-Ramirez A. *et al.,* | 2020 | 151 | 5 | 1092-1099 | doi.org/10.1097/j.pain.0000000000001801 |
| Chaumette T. *et al.,* | 2020 | 161 | 5 | 1109-1123 | doi.org/10.1097/j.pain.0000000000001808 |
| GniB S. *et al.,* | 2020 | 161 | 6 | 1191-1201 | doi.org/10.1097/j.pain.0000000000001811 |
| Schouppe S. *et al.,* | 2020 | 161 | 6 | 1212-1226 | doi.org/10.1097/j.pain.0000000000001813 |
| Hellman K.M. *et al.,* | 2020 | 161 | 6 | 1227-1236 | doi.org/10.1097/j.pain.0000000000001826 |
| Chen L. et al., | 2020 | 161 | 6 | 1237-1254 | doi.org/10.1097/j.pain.0000000000001831 |
| Iwabuchi S. *et al.,* | 2020 | 161 | 6 | 1255-1263 | doi.org/10.1097/j.pain.0000000000001829 |
| Grant A.D. *et al.,* | 2020 | 161 | 6 | 1264-1269 | doi.org/10.1097/j.pain.0000000000001806 |
| van den Heuvel M. *et al.,* | 2020 | 161 | 6 | 1278-1285 | doi.org/10.1097/j.pain.0000000000001797 |
| Pilch M. *et al.,* | 2020 | 161 | 6 | 1286-1296 | doi.org/10.1097/j.pain.0000000000001827 |
| Kim J.Y. *et al.,* | 2020 | 161 | 6 | 1311-1320 | doi.org/10.1097/j.pain.0000000000001816 |
| Fattori V. *et al.,* | 2020 | 161 | 6 | 1321-1331 | doi.org/10.1097/j.pain.0000000000001832 |
| Morasco B. J. *et al.,* | 2020 | 161 | 6 | 1332-1340 | doi.org/10.1097/j.pain.0000000000001817 |
| Gould H.M. *et al.,* | 2020 | 161 | 6 | 1341-1349 | doi.org/10.1097/j.pain.0000000000001834 |
| Chen L. et al., | 2020 | 161 | 6 | 1350-1360 | doi.org/10.1097/j.pain.0000000000001807 |
| Ten B.A.F. *et al.,* | 2020 | 161 | 6 | 1361-1370 | doi.org/10.1097/j.pain.0000000000001830 |
| Da Silver J.T. *et al.,* | 2020 | 161 | 6 | 1371-1380 | doi.org/10.1097/j.pain.0000000000001810 |
| Zhang J. *et al.,* | 2020 | 161 | 6 | 1381-1398 | doi.org/10.1097/j.pain.0000000000001818 |
| La Porta C. *et al.,* | 2020 | 161 | 7 | 1442-1458 | doi.org/10.1097/j.pain.0000000000001850 |
| Torta D.M. *et al.,* | 2020 | 161 | 7 | 1459-1469 | doi.org/10.1097/j.pain.0000000000001841 |
| Locke S. *et al.,* | 2020 | 161 | 7 | 1483-1496 | doi.org/10.1097/j.pain.0000000000001837 |
| Waxman J.A. *et al.,* | 2020 | 161 | 7 | 1518-1531 | doi.org/10.1097/j.pain.0000000000001840 |
| Heber S. *et al.,* | 2020 | 161 | 7 | 1532-1541 | doi.org/10.1097/j.pain.0000000000001848 |
| Jung C. *et al.,* | 2020 | 161 | 7 | 1555-1564 | doi.org/10.1097/j.pain.0000000000001815 |
| Albers K.I. *et al.,* | 2020 | 161 | 7 | 1565-1570 | doi.org/10.1097/j.pain.0000000000001842 |

| | | | | | |
|---|---|---|---|---|---|
| Leuchtweis J. *et al.,* | 2020 | 161 | 7 | 1571-1583 | doi.org/10.1097/j.pain.0000000000001852 |
| Ma P. *et al.,* | 2020 | 161 | 7 | 1584-1596 | doi.org/10.1097/j.pain.0000000000001855 |
| Thomaidou M. A. *et al.,* | 2020 | 161 | 7 | 1597-1608 | doi.org/10.1097/j.pain.0000000000001861 |
| Vitterso A.D. *et al.,* | 2020 | 161 | 7 | 1608-1628 | doi.org/10.1097/j.pain.0000000000001845 |
| Moy J.K. *et al.,* | 2020 | 161 | 7 | 1636-1649 | doi.org/10.1097/j.pain.0000000000001846 |
| Fan W. *et al.,* | 2020 | 161 | 7 | 1650-1660 | doi.org/10.1097/j.pain.0000000000001836 |
| Morais A. *et al.,* | 2020 | 161 | 7 | 1661-1669 | doi.org/10.1097/j.pain.0000000000001856 |
| Hoffmann J. *et al.,* | 2020 | 161 | 7 | 1670-1681 | doi.org/10.1097/j.pain.0000000000001858 |
| Davoli-Ferreira M. *et al.,* | 2020 | 161 | 8 | 1730-1743 | doi.org/10.1097/j.pain.0000000000001879 |
| Tchivileva I.E. *et al.,* | 2020 | 161 | 8 | 1755-1767 | doi.org/10.1097/j.pain.0000000000001882 |
| Walsh J. *et al.,* | 2020 | 161 | 8 | 1776-1786 | doi.org/10.1097/j.pain.0000000000001873 |
| Moayedi M. *et al.,* | 2020 | 161 | 8 | 1787-1797 | doi.org/10.1097/j.pain.0000000000001864 |
| Stine C. *et al.,* | 2020 | 161 | 8 | 1798-1807 | doi.org/10.1097/j.pain.0000000000001886 |
| Ohmichi Y. *et al.,* | 2020 | 161 | 8 | 1808-1823 | doi.org/10.1097/j.pain.0000000000001867 |
| Xiong B. *et al.,* | 2020 | 161 | 8 | 1824-1836 | doi.org/10.1097/j.pain.0000000000001878 |
| Seminowicz D.A. *et al.,* | 2020 | 161 | 8 | 1837-1846 | doi.org/10.1097/j.pain.0000000000001860 |
| Firouzian S. *et al.,* | 2020 | 161 | 8 | 1847-1860 | doi:org.10.1097/j.pain.0000000000001863 |
| Cardenas A. *et al.,* | 2020 | 161 | 8 | 1861-1871 | doi.org/10.1097/j.pain.0000000000001875 |
| Okusogu C. *et al.,* | 2020 | 161 | 8 | 1872-1883 | doi.org/10.1097/j.pain.0000000000001876 |
| Wells R. E. *et al.,* | 2020 | 161 | 8 | 1884-1893 | doi.org/10.1097/j.pain.0000000000001865 |
| Kucharczyk M. W. *et al.,* | 2020 | 161 | 8 | 1894-1905 | doi.org/10.1097/j.pain.0000000000001880 |
| Brewer C.L. *et al.,* | 2020 | 161 | 8 | 1906-1924 | doi.org/10.1097/j.pain.0000000000001884 |
| Xie W. *et al.,* | 2020 | 161 | 8 | 1925-1936 | doi.org/10.1097/j.pain.0000000000001887 |
| Lin T. *et al.,* | 2020 | 161 | 9 | 1995-2009 | doi.org/10.1097/j.pain.0000000000001898 |
| Colloca L. *et al.,* | 2020 | 161 | 9 | 2010-2021 | doi.org/10.1097/j.pain.0000000000001900 |
| Hadschieff V. *et al.,* | 2020 | 161 | 9 | 2022-2034 | doi.org/10.1097/j.pain.0000000000001901 |
| Chen K. *et al.,* | 2020 | 161 | 9 | 2035-2040 | doi.org/10.1097/j.pain.0000000000001906 |
| Speltz R. *et al.,* | 2020 | 161 | 9 | 2041-2057 | doi.org/10.1097/j.pain.0000000000001902 |

| | | | | | |
|---|---|---|---|---|---|
| Markman J.D. *et al.,* | 2020 | 161 | 9 | 2068-2078 | doi.org/10.1097/j.pain.0000000000001928 |
| Sandstrom A. *et al.,* | 2020 | 161 | 9 | 2079-2088 | doi.org/10.1097/j.pain.0000000000001907 |
| Liang Z. *et al.,* | 2020 | 161 | 9 | 2089-2106 | doi.org/10.1097/j.pain.0000000000001914 |
| Chakrabarti C. *et al.,* | 2020 | 161 | 9 | 2129-2141 | doi.org/10.1097/j.pain.0000000000001890 |
| Flores B. *et al.,* | 2020 | 161 | 9 | 2142-2154 | doi.org/10.1097/j.pain.0000000000001903 |
| Allen B.I. *et al.,* | 2020 | 161 | 9 | 2155-2166 | doi.org/10.1097/j.pain.0000000000001908 |
| Barroso J. *et al.,* | 2020 | 161 | 9 | 2167-2178 | doi.org/10.1097/j.pain.0000000000001904 |
| Lucarin E. *et al.,* | 2020 | 161 | 9 | 2179-2190 | doi.org/10.1097/j.pain.0000000000001905 |
| Verrico C,D. *et al.,* | 2020 | 161 | 9 | 2191-2202 | doi.org/10.1097/j.pain.0000000000001896 |
| Campos-Jurado Y. *et al.,* | 2020 | 161 | 9 | 2203-2211 | doi.org/10.1097/j.pain.0000000000001915 |
| Soltani S. *et al.,* | 2020 | 161 | 10 | 2263-2273 | doi.org/10.1097/j.pain.0000000000001916 |
| Syx D. *et al.,* | 2020 | 161 | 10 | 2274-2283 | doi.org/10.1097/j.pain.0000000000001935 |
| Jensen M.P. *et al.,* | 2020 | 161 | 10 | 2284-2298 | doi.org/10.1097/j.pain.0000000000001943 |
| Olopoenia A. *et al.,* | 2020 | 161 | 10 | 2299-2307 | doi.org/10.1097/j.pain.0000000000001934 |
| Alaiti R. *et al.,* | 2020 | 161 | 10 | 2321-2329 | doi.org/10.1097/j.pain.0000000000001919 |
| Scheff N. N. *et al.,* | 2020 | 161 | 10 | 2330-2343 | doi.org/10.1097/j.pain.0000000000001926 |
| Laumet G. et al., | 2020 | 161 | 10 | 2344-2352 | doi.org/10.1097/j.pain.0000000000001921 |
| Duron D.I. *et al.,* | 2020 | 161 | 10 | 2353-2363 | doi.org/10.1097/j.pain.0000000000001918 |
| Hamity M.V. *et al.,* | 2020 | 161 | 10 | 2364-2375 | doi.org/10.1097/j.pain.0000000000001924 |
| Gee T.A. *et al.,* | 2020 | 161 | 10 | 2376-2384 | doi.org/10.1097/j.pain.0000000000001925 |
| Fasolino A. *et al.,* | 2020 | 161 | 10 | 2385-2393 | doi.org/10.1097/j.pain.0000000000001920 |
| Topham L. et al., | 2020 | 161 | 10 | 2394-2409 | doi.org/10.1097/j.pain.0000000000001917 |
| Shiers S. *et al.,* | 2020 | 161 | 10 | 2410-2424 | doi.org/10.1097/j.pain.0000000000001973 |
| Ma K. *et al.,* | 2020 | 161 | 11 | 2502-2510 | doi.org/10.1097/j.pain.0000000000001957 |
| Burns J.W. et al., | 2020 | 161 | 11 | 2511-2519 | doi.org/10.1097/j.pain.0000000000001960 |
| Arai H. *et al.,* | 2020 | 161 | 11 | 2520-2531 | doi.org/10.1097/j.pain.0000000000001946 |
| Lin L. *et al.,* | 2020 | 161 | 11 | 2532-2538 | doi.org/10.1097/j.pain.0000000000001940 |
| Avona A. *et al.,* | 2020 | 161 | 11 | 2539-2550 | doi.org/10.1097/j.pain.0000000000001953 |

| | | | | | |
|---|---|---|---|---|---|
| Cai S. *et al.,* | 2020 | 161 | 11 | 2551-2570 | doi.org/10.1097/j.pain.0000000000001955 |
| Merrer J. *et al.,* | 2020 | 161 | 11 | 2571-2580 | doi.org/10.1097/j.pain.0000000000001956 |
| Liu Z. *et al.,* | 2020 | 161 | 1 | 2581-2591 | doi.org/10.1097/j.pain.0000000000001954 |
| Salvo E. *et al.,* | 2020 | 161 | 11 | 2592-2602 | doi.org/10.1097/j.pain.0000000000001986 |
| Zhaoyang R. *et al.,* | 2020 | 161 | 11 | 2603-2610 | doi.org/10.1097/j.pain.0000000000001959 |
| Montes A. *et al.,* | 2020 | 161 | 11 | 2611-2618 | doi.org/10.1097/j.pain.0000000000001945 |
| Moutal A. *et al.,* | 2020 | 161 | 11 | 2629-2651 | doi.org/10.1097/j.pain.0000000000001951 |
| Power A. *et al.,* | 2020 | 161 | 12 | 2720-2730 | doi.org/10.1097/j.pain.0000000000001966 |
| Gillving M. *et al.,* | 2020 | 161 | 12 | 2731-2736 | doi.org/10.1097/j.pain.0000000000001964 |
| Hegde S. *et al.,* | 2020 | 161 | 12 | 2737-2749 | doi.org/10.1097/j.pain.0000000000001967 |
| Zhang Z. *et al.,* | 2020 | 161 | 12 | 2750-2762 | doi.org/10.1097/j.pain.0000000000001962 |
| Palermo T. M. *et al.,* | 2020 | 161 | 12 | 2763-2774 | doi.org/10.1097/j.pain.0000000000001994 |
| Domenichiello A.F. *et al.,* | 2020 | 161 | 12 | 2775-2785 | doi.org/10.1097/j.pain.0000000000001983 |
| Lagos-Rodriguez V. *et al.,* | 2020 | 161 | 12 | 2786-2797 | doi.org/10.1097/j.pain.0000000000001992 |
| Massaly N. *et al.,* | 2020 | 161 | 12 | 2798-2804 | doi.org/10.1097/j.pain.0000000000001968 |
| Elias G.J.B. *et al.,* | 2020 | 161 | 12 | 2805-2819 | doi.org/10.1097/j.pain.0000000000001998 |
| Eldable S. *et al.,* | 2020 | 161 | 12 | 2820-2829 | doi.org/10.1097/j.pain.0000000000001977 |
| Araya E.I. *et al.,* | 2020 | 161 | 12 | 2830-2840 | doi.org/10.1097/j.pain.0000000000001970 |
| Shulman J. *et al.,* | 2020 | 161 | 12 | 2852-2859 | doi.org/10.1097/j.pain.0000000000001984 |
| van Tilburg M.A.L. et al., | 2020 | 161 | 12 | 2860-2871 | doi.org/10.1097/j.pain.0000000000001996 |
| Chao D. *et al.,* | 2020 | 161 | 12 | 2872-2886 | doi.org/10.1097/j.pain.0000000000001982 |
| Bruehl S. et al., | 2020 | 161 | 12 | 2887-2897 | doi.org/10.1097/j.pain.0000000000001969 |
| Lesnak J.B. *et al.,* | 2020 | 161 | 12 | 2898-2908 | doi.org/10.1097/j.pain.0000000000001985 |
| Kajiwara I. *et al.,* | 2020 | 161 | 12 | 2909-2919 | doi.org/10.1097/j.pain.0000000000001997 |