Changes of pain intensity results with different follow-up times in randomized controlled trials of osteoarthritis

Siebers, Nys

Master's thesis / Diplomski rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: University of Split, School of Medicine / Sveučilište u Splitu, Medicinski fakultet

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:171:229121

Rights / Prava: In copyright

Download date / Datum preuzimanja: 2020-11-08

Repository / Repozitorij:

MEFST Repository
Nys Willem Siebers

CHANGES OF PAIN INTENSITY RESULTS WITH DIFFERENT FOLLOW-UP TIMES IN RANDOMIZED CONTROLLED TRIALS OF OSTEOARTHRITIS: A CASE STUDY OF CELECOXIB

Diploma Thesis

Academic Year:

2018/2019

Mentor:
Prof. Livia Puljak, MD, PhD

Split, July 2019
Table of Contents

1. INTRODUCTION .............................................................................................................. 1
   1.1. Osteoarthritis ............................................................................................................. 2
   1.4. Cochrane systematic review ....................................................................................... 4
   1.5. Cochrane review about celecoxib for osteoarthritis ..................................................... 5
2. OBJECTIVES ....................................................................................................................... 7
3. METHODS ........................................................................................................................... 9
   3.1. Study Design ................................................................................................................ 10
   3.2. Inclusion of studies ...................................................................................................... 10
   3.3. Types of Intervention .................................................................................................. 10
   3.4. Types of outcome measures ....................................................................................... 10
   3.5. Extraction of data ....................................................................................................... 10
   3.6. Data analysis ............................................................................................................... 10
   3.7. Data imputations ........................................................................................................ 11
4. RESULTS ............................................................................................................................ 12
   Included studies .................................................................................................................. 13
   Effect sizes .......................................................................................................................... 15
5. DISCUSSION ....................................................................................................................... 29
6. CONCLUSION ...................................................................................................................... 34
7. SUMMARY .......................................................................................................................... 36
8. CROATIAN SUMMARY ..................................................................................................... 38
9. CURRICULUM VITAE ........................................................................................................ 40
10. REFERENCES ..................................................................................................................... 42
ACKNOWLEDGEMENT

First of all, I would like to thank my mentor Prof. Livia Puljak, MD, PhD. Without her help this Thesis would not have been possible. Additionally, I want to thank Assist. Prof. Filipa Markotic MD, PhD who has helped me collect data and was always available when I was in doubt about anything.

Lastly, I would also like to thank my parents, Dr. Derk Siebers and Elke Siebers as well as my sister Merle Siebers who have always supported me throughout my professional as well as private life. Without them I would not be the person I am today.
**List of abbreviations**

BMI – Body Mass Index

CDSR – Cochrane Database of Systematic Reviews

CFB – Change from Baseline

COX1/COX2 – Cyclooxygenase 1/Cyclooxygenase 2

DALY’s – Disability Adjusted Life Years

ICTRP – International Clinical Trials Registry Platform

IPR – Inadequate pain relieve

NSAID’s – Non-steroidal anti-inflammatory drugs

OA – Osteoarthritis

RCT’s – Randomized Controlled Trials

SD – Standard deviation

SMD – Standardized mean difference

VAS – Visual Analog Scale

WOMAC – Western Ontario and Mc Master Universities Osteoarthritis Index
1. INTRODUCTION
1.1. Osteoarthritis

Osteoarthritis is a chronic degenerative disease predominantly affecting weight bearing joints in the human body. It is currently suggested to be a heterogeneous disease caused by a combination of excessive wear and tear as well as abnormal joint mechanics and inflammation. The concept of its pathophysiology is still unfolding. Through progress in molecular biology it evolved from being viewed as a cartilage-limited disorder to a multifactorial disease affecting the whole joint. This intricate relationship between local and systemic factors modulates its structural features and clinical presentation leading to a common final pathway of joint destruction [1].

The pain underlying Osteoarthritis is of heterogeneous nature. An interplay between local pathologic changes, neuroplastic changes as well as general factors like adipositas, diabetes mellitus and psychosocial factors have been identified to be responsible for the development of chronic joint pain [2]. Presumably arising from mechanical sensitization of joint nociceptors through inflammation, pain perception progresses in response to a complex series of neurophysiologic events. They are comprised of sensitization of peripheral and central pathways as well as reduction of descending conditioning pain modulation and atrophy of cortical areas involved in pain processing [2]. Moreover, a subset of patients pain phenotype indicates a neuropathic component [3]. All those mechanisms combined likely skew the relationship between the extend of tissue injury and perceived pain in any situation but the acute one. Current evidence suggests osteoarthritic damage predisposes to pain but there is little correlation between the severity of pain and the extend of joint damage [4].

Osteoarthritis most commonly results from a combination of modifiable and non-modifiable risk factors including obesity, trauma, increasing age, genetic predisposition and gender. Those affected classically suffer from pain, stiffness and limited range of motion ultimately leading to joint destruction and the necessity to perform joint replacement surgery [5].

Osteoarthritis is the single most common cause of chronic disability in older adults [5]. A report from the Global Burden of Disease 2010 study indicated that of the 291 conditions listed, hip and knee osteoarthritis was ranked globally as the 11th highest contributor to global disability and the 38th highest in disability-adjusted life years (DALYs) [6]. It is estimated that 10-15% of all adults aged over 60 have some degree of osteoarthritis with a prevalence that is higher among women than men [7] The increasing lifespan of the general population combined with an expected rise in obese patients can potentially aggravate the global impact of this disease.
According to the United Nations, by 2050, 130 million people will suffer from OA worldwide of whom 40 million will be severely disabled by the disease [8].

Treatment of osteoarthritis is directed at pain alleviation, improvement of physical function and the delay of joint replacement surgery. Treatment modalities are generally divided into non-pharmacological, pharmacological and surgical options. Non-pharmacologic options include: patient education, application of heat and cold, weight loss, low to moderate intensity exercise, physical therapy and mechanical joint unloading through braces or foot wear [5].

Pharmacologic options include medicines such as acetaminophen, topical and oral non-steroidal anti-inflammatory drugs (NSAIDS), tramadol and intra-articular injections like corticosteroids or autologous Plasma. Additionally, there are certain nutritional supplements, foremost glucosamine, which has shown some beneficial results in osteoarthritis clinical trials [9]. New potential targets for analgesic therapy have been identified. The antibody tanezumab targeting nerve growth factor; sensory proteins at the nociceptive nerve endings such as the activating TRPV and ASIC channel. Additionally, axonal channels such as voltage-gated Sodium channels, various potassium channels as well as inhibitory opioid and cannabinoid receptors [10].

Surgical procedures that are used for symptomatic treatment of osteoarthritis include arthroscopy for debridement, osteotomy or fusion [11]. Unfortunately, the only definite treatment is joint replacement which, due to limited durability of modern implants, is often preceded by years of chronic analgesic use and significant disability. After successful arthroplasty, as defined by prosthesis-related outcomes, still a proportion of about 9% of patients with hip and about 20% of patients with knee replacements have unfavorable long-term results with patient-centered pain outcomes over a follow up from 3 month to 5 years after surgery [12]. In order to address the time period from onset of disease to joint replacement, patients need therapies that provide adequate pain relief over an extended period of time.
1.3. Celecoxib

Celecoxib is a drug belonging to the class of coxibs, recommended for symptomatic treatment of osteoarthritis. Like conventional NSAIDS, they work by inhibiting the cyclooxygenase enzyme which converts arachidonic acid into prostaglandins which further mediate pain and inflammation amongst other functions. The cyclooxygenase enzyme has two isoforms active in humans. The constitutive COX-1 is present, for example in the endothelium, stomach and kidney, whereas COX-2 is induced by pro-inflammatory cytokines and endotoxin in cells in vitro and at inflammatory sites in vivo [13]. In contrast to conventional NSAIDS which inhibit both isoforms of the enzyme non-selectively, coxibs are relatively more selective for the COX-2 enzyme. Whilst equally efficacious their specificity presumably gives them a more favorable side effect profile. Unfortunately, this has not been confirmed in the Cochrane review “Celecoxib for osteoarthritis” [14].

1.4. Cochrane systematic review

Systematic reviews are secondary research projects in which researchers attempts to gather all the existing empirical evidence that meets pre-specified eligibility criteria in order to answer a specific research question. Methods used for conducting such reviews should be explicit and systematic, striving towards minimization of bias in order to facilitate the production of reliable findings for further decision making [15]. Those reviews are complicated and their results heavily depend on the availability, and even more importantly, the quality of existing clinical trials. The strength of the evidence of the systematic review directly correlates with the quality of included studies. If possible, in a systematic review authors may pool numerical data about treatment effects through the process of meta-analysis. Through this process systematic reviews are able to summarize all the existing clinical research concerning a particular research question [16].

Cochrane systematic reviews are considered gold standard in evidence synthesis field. They are produced by Cochrane, a global independent network of health practitioners, researchers, patient advocates and people interested in health from over 130 countries. Cochrane has more than 37000 contributors that collectively respond to the challenge of making the vast amount of evidence available through research applicable for consumers. Cochrane is a not-for profit organization whose mission is to produce high quality evidence that is free from commercial
sponsorship and other conflicts of interests in order to facilitate evidence-based decision making in the health-care setting [17].

Cochrane systematic reviews are systematic reviews in the field of health-care published in the Cochrane Database of Systematic Reviews (CDSR), the leading journal for systematic reviews in health care [18]. Cochrane has developed a meticulous methodological approach for producing systematic reviews. There are five types of systematic reviews in CDSR: reviews of the effects of interventions, reviews of diagnostic test accuracy, methodology reviews, qualitative reviews, and methodology reviews [19]. Additionally, CDSR publishes overviews of systematic reviews, i.e. systematic reviews that summarize systematic reviews [20]. All those reviews follow a clear structured review model which is provided in the Cochrane Handbook for Systematic Reviews of Interventions [21]. Guidance available in the Cochrane Handbook should guarantee consistency of methods used in Cochrane reviews.

1.5. Cochrane review about celecoxib for osteoarthritis
Cochrane review “Celecoxib for osteoarthritis” was published in the CDSR in 2017 [14]. The review showed no statistically significant difference between celecoxib and placebo for serious adverse effects, gastro-intestinal events (perforations, ulcers bleeds) and cardiovascular events (myocardial infarction, stroke). Due to high risk of bias and imprecision it is to be noted that evidence level was downgraded to very low quality [14].

The same review reached several other conclusions relevant for clinical decision making: Firstly, they noted that benefits of celecoxib were not much different than placebo or other NSAIDS. Furthermore, they noticed decreasing efficacy of celecoxib for pain with longer duration of included studies, expressed as decreasing standardized mean difference [14].

On the contrary, previous data from the research group of prof. Andrew Moore indicated that there is high correlation of pain scores measured with a visual-analog-scale (VAS) after 2 and 6 weeks of treatment with VAS pain scores at 12 weeks in randomized controlled trials (RCTs) about rheumatoid arthritis and osteoarthritis [22]. This group of authors concluded that early analgesic response measuring pain scores with VAS beyond 2 weeks of treatment with a particular NSAID is likely to be predictive of pain VAS response at 12 weeks, and that these results have implications for future study design of randomized controlled trials RCTs).
Namely, the authors suggest that appropriate treatment duration for studies of efficacy in this setting could be shorter, for example 6 weeks instead of 12 weeks [22].

Since efficacy data for pain from the Cochrane review “Celecoxib for Osteoarthritis” would imply different conclusion compared to conclusions of Moore and colleagues, the aim of this study was to conduct more comprehensive analysis of efficacy data for pain in RCTs about celecoxib for osteoarthritis over different follow-up times.
2. OBJECTIVES
The aim of this Thesis was to conduct more comprehensive analysis of efficacy data for pain in RCTs about celecoxib in osteoarthritis. The purpose of this is to improve long-term management of pain for patients suffering from osteoarthritis by guiding clinical decision making, and to create evidence that will inform design of future RCTs about osteoarthritis.
3. METHODS
3.1. Study Design
This was a retrospective primary methodological study, in which publicly available data from published RCTs were analyzed. Therefore, permission of the ethics committee for data collection was not necessary.

3.2. Inclusion of studies
We included RCTs analyzing the effects of celecoxib on pain intensity measured with the Visual Analog Scale (VAS) and/or the Western Ontario and McMaster University Osteoarthritis Index (WOMAC), and comparing celecoxib with placebo. We did not limit studies based on duration, and we had no limits regarding language. Search strategy used for retrieving eligible studies was described in the Cochrane review “Celecoxib for Osteoarthritis”, and we used for this analysis all eligible RCTs that were found in the literature while conducting our Cochrane review [14].

3.3. Types of Intervention
Oral celecoxib 200 mg daily (either as 200 mg once daily or 100 mg twice daily) versus Placebo. Dosage of 200 mg was used because it is the recommended dosage.

3.4. Types of outcome measures
The outcome measure was pain. Pain scales used were the VAS scale and the WOMAC osteoarthritis index pain sub score.

3.5. Extraction of data
We extracted the following data from eligible RCTs: study ID (first author, year), study duration in weeks, follow-up times used in the study for measuring pain intensity, efficacy data for pain measured with VAS and/or WOMAC for all reported follow-up times (mean, standard deviation, number of participants). If the study reported data only in figures, we extracted data from figures using the Plot Digitizer software [23]. We extracted data in the way they were presented, including baseline data, final end-of-study data and change from baseline.

3.6. Data analysis
Since the majority of data were reported as change from baseline, for the studies that reported baseline data and absolute values at different time points, we calculated change from baseline
using baseline data and time point data using methods recommended in the Cochrane Handbook for Systematic Reviews of Interventions [24].

We used random-effects meta-analyses for synthesis of pain scores for different pain outcome measures and different follow-up time points that were reported in included studies. Standardized mean differences (SMDs) were used to report the data. We used Review Manager (RevMan) for data analyses [25].

3.7. Data imputations
Missing standard deviations (SDs) were imputed only from baseline data or other follow-up data of the same manuscript. We did not do any imputations for missing SDs from other manuscripts. For studies that have shown only absolute results, we calculated change from baseline.
4. RESULTS
Included studies
We included 35 RCTs in this analysis. All included RCTs were published as full-text manuscripts. We did not find any eligible RCTs that were published as conference abstracts, or that were unpublished. The list of included studies is shown in Table 1.

Table 1. List of included studies

<table>
<thead>
<tr>
<th>No</th>
<th>Included studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Asmus 2014 Study 1 [26]</td>
</tr>
<tr>
<td>2.</td>
<td>Asmus 2014 Study 2 [26]</td>
</tr>
<tr>
<td>3.</td>
<td>Bensen 1999 [27]</td>
</tr>
<tr>
<td>4.</td>
<td>Bingham 2007 Study 1 [28]</td>
</tr>
<tr>
<td>5.</td>
<td>Bingham 2007 Study 2 [28]</td>
</tr>
<tr>
<td>6.</td>
<td>Birbara 2006 Study 1 [29]</td>
</tr>
<tr>
<td>7.</td>
<td>Birbara 2006 Study 2 [29]</td>
</tr>
<tr>
<td>10.</td>
<td>Clegg 2006 [31]</td>
</tr>
<tr>
<td>11.</td>
<td>Conaghan 2013 [32]</td>
</tr>
<tr>
<td>12.</td>
<td>De Lemos 2011 [33]</td>
</tr>
<tr>
<td>13.</td>
<td>Essex 2016 [34]</td>
</tr>
<tr>
<td>15.</td>
<td>Gibofsky 2003 [36]</td>
</tr>
<tr>
<td>16.</td>
<td>Gordo 2017 [37]</td>
</tr>
<tr>
<td>17.</td>
<td>Hochberg 2011 Study 307 [38]</td>
</tr>
<tr>
<td>18.</td>
<td>Hochberg 2011 Study 309 [38]</td>
</tr>
<tr>
<td>20.</td>
<td>Lee M 2017 [40]</td>
</tr>
<tr>
<td>21.</td>
<td>Lehman 2005 [41]</td>
</tr>
<tr>
<td>22.</td>
<td>Mc Kenna 2001a [42]</td>
</tr>
<tr>
<td>23.</td>
<td>Mc Kenna 2001b [42]</td>
</tr>
<tr>
<td>24.</td>
<td>Pincus 2004 PACES-a [43]</td>
</tr>
<tr>
<td>25.</td>
<td>Pincus 2004 PACES-b [43]</td>
</tr>
<tr>
<td>26.</td>
<td>Reginster 2017 [44]</td>
</tr>
</tbody>
</table>
We excluded 14 studies due to reasons listed in the Table 2.

Table 2. List of excluded studies

<table>
<thead>
<tr>
<th>No</th>
<th>Excluded studies</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bianchi 2003 [53]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>2</td>
<td>Bianchi 2007 [54]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>3</td>
<td>Detrembleur 2005 [55]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>4</td>
<td>EUCTR2005-002772-14-GB</td>
<td>Outcome data is the same as in Schnitzer 2011</td>
</tr>
<tr>
<td>5</td>
<td>EUCTR2011-005398-22-ES</td>
<td>Results are not available</td>
</tr>
<tr>
<td>6</td>
<td>Gallelli 2013 [56]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>7</td>
<td>Leeb 2004 [57]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>8</td>
<td>Mastbergen 2010 [58]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>9</td>
<td>NCT01768520</td>
<td>Results of study could not be found. Stated to use Korean WOMAC</td>
</tr>
<tr>
<td>10</td>
<td>Ozgocmen 2005 [59]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>11</td>
<td>Sampalis 2012 [60]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>12</td>
<td>Simon 1998 [61]</td>
<td>SD not reported, and could not be imputed from other results reported in this manuscript</td>
</tr>
<tr>
<td>13</td>
<td>Tascioglu 2004 [62]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
<tr>
<td>14</td>
<td>Trudeau 2015 [63]</td>
<td>fewer than 50 participants in each arm</td>
</tr>
</tbody>
</table>
Results from the Schnitzer 2011 study and results posted for the study EUCTR2005-002772-14-GB registered at the International Clinical Trials Registry Platform (ICTRP) had exactly the same results, up to two decimals. Even though the Schnitzer 2011 study reported in the manuscript that the study was registered only on ClinicalTrials.gov (NCT00154219), details in these two registrations on ICTRP and on ClinicalTrials.gov are identical, and therefore we considered that these are the same studies, and we did not include these data two times in our analysis.

**Effect sizes**

Time points for results available for data analysis in included studies that have reported pain using VAS are shown in Table 3, while the time points for results in studies that reported pain using WOMAC is shown in Table 4.

**Table 3.** Time points with results for pain measured with visual analogue scale (VAS) in included studies

<table>
<thead>
<tr>
<th>Time point</th>
<th>Study name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 weeks</td>
<td>Bensen 1999</td>
</tr>
<tr>
<td></td>
<td>Bingham 2007 study 1</td>
</tr>
<tr>
<td></td>
<td>Bingham 2007 study 2</td>
</tr>
<tr>
<td></td>
<td>Fleischmann 2005</td>
</tr>
<tr>
<td></td>
<td>Kivitz 2001</td>
</tr>
<tr>
<td></td>
<td>Lehman 2005</td>
</tr>
<tr>
<td></td>
<td>McKenna 2001b</td>
</tr>
<tr>
<td></td>
<td>Sheldon 2005</td>
</tr>
<tr>
<td></td>
<td>Simon 1998</td>
</tr>
<tr>
<td></td>
<td>Tannenbaum 2004</td>
</tr>
<tr>
<td></td>
<td>Williams 2000</td>
</tr>
<tr>
<td></td>
<td>Williams 2001</td>
</tr>
<tr>
<td>3 weeks</td>
<td>Gibofsky 2003</td>
</tr>
<tr>
<td></td>
<td>McKenna 2001a</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Bingham 2007 study 1</td>
</tr>
<tr>
<td></td>
<td>Bingham 2007 study 2</td>
</tr>
<tr>
<td></td>
<td>Lehman 2005</td>
</tr>
<tr>
<td>Time Frame</td>
<td>Studies</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>30 days</td>
<td>Sheldon 2005, Schnitzer 2011, Tannenbaum 2004, Reginster 2017</td>
</tr>
<tr>
<td>9 weeks</td>
<td>Conaghan 2013, DeLemos 2011</td>
</tr>
<tr>
<td>Time point</td>
<td>Study name</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>1 week</td>
<td>Boswell 2008 Study A</td>
</tr>
<tr>
<td></td>
<td>De Lemos 2011</td>
</tr>
<tr>
<td></td>
<td>Wittenberg 2006</td>
</tr>
<tr>
<td>2 weeks</td>
<td>Bensen WG 1999</td>
</tr>
<tr>
<td></td>
<td>Birbara 2006 Study 1</td>
</tr>
<tr>
<td></td>
<td>Birbara 2006 Study 2</td>
</tr>
<tr>
<td></td>
<td>Boswell 2008 Study A</td>
</tr>
<tr>
<td></td>
<td>Boswell 2008 Study B</td>
</tr>
<tr>
<td></td>
<td>Conaghan 2013</td>
</tr>
<tr>
<td></td>
<td>De Lemos 2011</td>
</tr>
<tr>
<td></td>
<td>Fleischmann 2005</td>
</tr>
<tr>
<td></td>
<td>Kivitz 2001</td>
</tr>
<tr>
<td></td>
<td>Lehman 2005</td>
</tr>
<tr>
<td></td>
<td>Smugars 2006 Study 1</td>
</tr>
<tr>
<td></td>
<td>Smugars 2006 Study 2</td>
</tr>
<tr>
<td></td>
<td>Tannenbaum 2004</td>
</tr>
<tr>
<td>3 weeks</td>
<td>De Lemos 2011</td>
</tr>
<tr>
<td></td>
<td>Lee M 2017</td>
</tr>
<tr>
<td>4 weeks</td>
<td>Birbara 2006 Study 1</td>
</tr>
<tr>
<td></td>
<td>Birbara 2006 Study 2</td>
</tr>
<tr>
<td></td>
<td>Boswell 2008 Study A</td>
</tr>
<tr>
<td></td>
<td>Boswell 2008 Study B</td>
</tr>
<tr>
<td></td>
<td>Mastbergen 2010</td>
</tr>
<tr>
<td>14 weeks</td>
<td>Pincus 2004 PACES-a</td>
</tr>
<tr>
<td></td>
<td>Pincus 2004 PACES-b</td>
</tr>
<tr>
<td>15 weeks</td>
<td>Fleischmann 2005</td>
</tr>
<tr>
<td>26 weeks</td>
<td>Reginster 2017</td>
</tr>
</tbody>
</table>

**Table 4.** Time points with results for pain measured with Western Ontario and McMaster Universities Osteoarthritis index (WOMAC) in included studies.
Schnitzer 2011
Smugar 2006 Study 1
Smugar 2006 Study 2

6 weeks
Asmus 2014 Study 1
Asmus 2014 Study 2
Birbara 2006 Study 1
Birbara 2006 Study 2
Boswell 2008 Study A
Conaghan 2013
De Lemos 2011
Essex 2016
Gibofsky 2003
Gordo 2017
Hochberg 2011 Study 307
Hochberg 2011 Study 309
Lee M 2017
Rother 2007
Williams 2000
Williams 2001

8 weeks
Boswell 2008 Study B
Schnitzer 2011

9 weeks
Conaghan 2013
De Lemos 2011

12 weeks
Bensen WG 1999
Bingham 2007 Study 1
Bingham 2007 Study 2
Boswell 2008 Study B
Conaghan 2013
De Lemos 2011
Hochberg 2011 Study 307
Hochberg 2011 Study 309
Kivitz 2001

13 weeks
Fleischmann 2005
We imputed SD from other parts of the manuscript in 7 studies (Bensen 1999, Kivitz 2001, McKenna 2001a, Lehman 2005, Schnitzer 2011, Tannenbaum 2004, De Lemos 2001). There were 5 studies that showed only absolute values, and for which we calculated change from baseline (Bensen 1999, Gibofsky 2003, Reginster 2007, Williams 2000, Williams 2001).

We made 20 meta-analyses based on the included studies. There were 2 meta-analyses with only one study included. Other meta-analyses had from 2 to 18 included studies. The list of meta-analyses conducted is shown in Table 5 and 6 for pain VAS and pain WOMAC respectively.

As shown in Tables 5 and 6, SMDs had decreasing trend from earliest to latest analyzed time points, both for VAS and for WOMAC, indicating that the effect of celecoxib as an intervention was decreasing with time. However, all these studies were of short duration – the longest follow-up time used in studies that reported pain using VAS was 13 weeks, while the longest follow-up time reported for WOMAC pain was 24 weeks.

**Table 5.** Standardized mean difference (SMD) for pain in studies that compared celecoxib versus placebo, measured with visual analog scale

<table>
<thead>
<tr>
<th>Time point</th>
<th>SMD and 95% CI</th>
<th>Heterogeneity</th>
<th>Number of studies</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 weeks</td>
<td>-0.50 (-0.63 to -0.38)</td>
<td>81%</td>
<td>12</td>
<td>6047</td>
</tr>
<tr>
<td>3 weeks</td>
<td>-0.53 (-0.73 to -0.32)</td>
<td>0%</td>
<td>2</td>
<td>408</td>
</tr>
<tr>
<td>4 weeks</td>
<td>-0.45 (-0.55 to -0.35)</td>
<td>54%</td>
<td>6</td>
<td>3910</td>
</tr>
<tr>
<td>30 days</td>
<td>-0.13 (-0.32 to 0.07)</td>
<td>Not applicable</td>
<td>1</td>
<td>399</td>
</tr>
<tr>
<td>6 weeks</td>
<td>-0.44 (-0.55 to -0.32)</td>
<td>71%</td>
<td>12</td>
<td>4141</td>
</tr>
<tr>
<td>8 weeks</td>
<td>-0.41 (-0.52 to -0.30)</td>
<td>64%</td>
<td>6</td>
<td>3910</td>
</tr>
<tr>
<td>12 weeks</td>
<td>-0.48 (-0.64 to -0.31)</td>
<td>66%</td>
<td>5</td>
<td>1822</td>
</tr>
<tr>
<td>13 weeks</td>
<td>-0.23 (-0.30 to -0.17)</td>
<td>0%</td>
<td>7</td>
<td>3763</td>
</tr>
</tbody>
</table>
Table 6. Standardized mean difference (SMD) for pain in studies that compared celecoxib versus placebo, measured with WOMAC scale

<table>
<thead>
<tr>
<th>Time point</th>
<th>SMD and 95% CI</th>
<th>Heterogeneity</th>
<th>Number of studies</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 week</td>
<td>-0.32 (-0.46 to -0.18)</td>
<td>0%</td>
<td>3</td>
<td>828</td>
</tr>
<tr>
<td>2 weeks</td>
<td>-0.37 (-0.44 to -0.29)</td>
<td>53%</td>
<td>13</td>
<td>6146</td>
</tr>
<tr>
<td>3 weeks</td>
<td>-0.37 (-0.53 to -0.20)</td>
<td>0%</td>
<td>2</td>
<td>618</td>
</tr>
<tr>
<td>4 weeks</td>
<td>-0.30 (-0.39 to -0.21)</td>
<td>23%</td>
<td>7</td>
<td>3052</td>
</tr>
<tr>
<td>6 weeks</td>
<td>-0.35 (-0.43 to -0.27)</td>
<td>45%</td>
<td>16</td>
<td>5128</td>
</tr>
<tr>
<td>8 weeks</td>
<td>-0.29 (-0.49 to -0.08)</td>
<td>63%</td>
<td>2</td>
<td>1165</td>
</tr>
<tr>
<td>9 weeks</td>
<td>-0.24 (-0.37 to -0.11)</td>
<td>0%</td>
<td>2</td>
<td>862</td>
</tr>
<tr>
<td>12 weeks</td>
<td>-0.32 (-0.40 to -0.25)</td>
<td>13%</td>
<td>9</td>
<td>3468</td>
</tr>
<tr>
<td>13 weeks</td>
<td>-0.27 (-0.33 to -0.20)</td>
<td>0%</td>
<td>5</td>
<td>3393</td>
</tr>
<tr>
<td>24 weeks</td>
<td>-0.13 (-0.28 to -0.03)</td>
<td>Not applicable</td>
<td>1</td>
<td>631</td>
</tr>
</tbody>
</table>

Forest plots for individual meta-analyses are shown in Figures 1-19.

Figure 1. Comparison celecoxib versus placebo, outcome: pain VAS, 2 weeks
Figure 2. Comparison celecoxib versus placebo, outcome: pain VAS, 3 weeks

Figure 3. Comparison celecoxib versus placebo, outcome: pain VAS, 4 weeks

Figure 4. Comparison celecoxib versus placebo, outcome: pain VAS, 30 days
### Table 1: Comparison of Celecoxib versus Placebo for Pain VAS, 6 Weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asmus 2014 study 1</td>
<td>-27.3</td>
<td>29</td>
<td>186</td>
<td>-14.9</td>
<td>27.12</td>
<td>184</td>
<td>8.7%</td>
<td>-0.44 [-0.65, -0.23]</td>
<td></td>
</tr>
<tr>
<td>Asmus 2014 study 2</td>
<td>-28</td>
<td>29.24</td>
<td>194</td>
<td>-24.6</td>
<td>27.27</td>
<td>186</td>
<td>8.9%</td>
<td>-0.12 [-0.32, 0.08]</td>
<td></td>
</tr>
<tr>
<td>Bensen WC 1999</td>
<td>-29.16</td>
<td>9.2635</td>
<td>197</td>
<td>-18.11</td>
<td>38.1341</td>
<td>203</td>
<td>8.9%</td>
<td>-0.39 [-0.59, -0.20]</td>
<td></td>
</tr>
<tr>
<td>Essex 2016</td>
<td>-37.1</td>
<td>22</td>
<td>121</td>
<td>-33.6</td>
<td>19.801</td>
<td>58</td>
<td>6.4%</td>
<td>-0.16 [-0.48, 0.15]</td>
<td></td>
</tr>
<tr>
<td>Glibofsky 2003</td>
<td>-34</td>
<td>27.4955</td>
<td>189</td>
<td>-21.2</td>
<td>26.4545</td>
<td>96</td>
<td>7.8%</td>
<td>-0.47 [-0.72, -0.22]</td>
<td></td>
</tr>
<tr>
<td>Corco 2017</td>
<td>-34.5</td>
<td>24.6312</td>
<td>122</td>
<td>-28.4</td>
<td>25.5181</td>
<td>56</td>
<td>6.4%</td>
<td>-0.24 [-0.56, 0.07]</td>
<td></td>
</tr>
<tr>
<td>Kivitz 2001</td>
<td>-25.1</td>
<td>17</td>
<td>207</td>
<td>-13.2</td>
<td>14.9</td>
<td>217</td>
<td>9.0%</td>
<td>-0.74 [-0.94, -0.55]</td>
<td></td>
</tr>
<tr>
<td>McKenna 2001b</td>
<td>-34.9</td>
<td>28.1</td>
<td>199</td>
<td>-23.1</td>
<td>14.9</td>
<td>217</td>
<td>9.0%</td>
<td>-0.42 [-0.62, -0.22]</td>
<td></td>
</tr>
<tr>
<td>Pincus 2004 PACES-a</td>
<td>-19</td>
<td>25.6964</td>
<td>181</td>
<td>-10.5</td>
<td>25.1806</td>
<td>172</td>
<td>8.7%</td>
<td>-0.33 [-0.54, -0.12]</td>
<td></td>
</tr>
<tr>
<td>Pincus 2004 PACES-b</td>
<td>-21.8</td>
<td>25.96</td>
<td>189</td>
<td>-7.6</td>
<td>26.09</td>
<td>182</td>
<td>8.7%</td>
<td>-0.54 [-0.75, -0.34]</td>
<td></td>
</tr>
<tr>
<td>Williams 2000</td>
<td>-26.4</td>
<td>16.6</td>
<td>182</td>
<td>-13</td>
<td>15.1</td>
<td>146</td>
<td>8.3%</td>
<td>-0.84 [-1.07, -0.61]</td>
<td></td>
</tr>
<tr>
<td>Williams 2001</td>
<td>-24.2</td>
<td>16.5</td>
<td>231</td>
<td>-17.4</td>
<td>16.5</td>
<td>243</td>
<td>9.3%</td>
<td>-0.41 [-0.59, -0.23]</td>
<td></td>
</tr>
</tbody>
</table>

Total (95% CI)  2198       1943       100.0%  -0.44 [-0.55, -0.32]  

Heterogeneity: $\tau^2 = 0.03, \chi^2 = 37.49, df = 11 (P < 0.0001); I^2 = 71$

Test for overall effect: $Z = 7.30 (P < 0.00001)$

**Figure 6.** Comparison celecoxib versus placebo, outcome: pain VAS, 6 weeks
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo Mean</th>
<th>SD</th>
<th>Total</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bingham 2007 study 1</td>
<td>-24.65</td>
<td>24.94</td>
<td>241</td>
<td>-13.05</td>
<td>20.28</td>
<td>127</td>
<td>-0.49</td>
<td>[-0.71, -0.28]</td>
</tr>
<tr>
<td>Bingham 2007 study 2</td>
<td>-26.95</td>
<td>23.73</td>
<td>247</td>
<td>-16.21</td>
<td>18.90</td>
<td>117</td>
<td>-0.48</td>
<td>[-0.70, -0.26]</td>
</tr>
<tr>
<td>Lehman 2005</td>
<td>-26.54</td>
<td>23.65</td>
<td>420</td>
<td>-20.34</td>
<td>23.97</td>
<td>424</td>
<td>0.19</td>
<td>[-0.40, -0.12]</td>
</tr>
<tr>
<td>Schnitzer 2011</td>
<td>41.7</td>
<td>15.46</td>
<td>419</td>
<td>49.9</td>
<td>15.37</td>
<td>416</td>
<td>-0.53</td>
<td>[-0.67, -0.39]</td>
</tr>
<tr>
<td>Sheldon 2005</td>
<td>-25.6</td>
<td>14.17</td>
<td>393</td>
<td>-18.8</td>
<td>13.18</td>
<td>382</td>
<td>0.50</td>
<td>[-0.64, -0.35]</td>
</tr>
<tr>
<td>Tannenbaum 2004</td>
<td>-24.8</td>
<td>24.7</td>
<td>481</td>
<td>-18.4</td>
<td>26.1</td>
<td>243</td>
<td>0.25</td>
<td>[-0.41, -0.10]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>2201</td>
<td></td>
<td></td>
<td>1709</td>
<td></td>
<td></td>
<td>0.41 [-0.52, -0.30]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.01; Chi² = 13.91, df = 5 (P = 0.02); I² = 64%
Test for overall effect: Z = 7.30 (P < 0.00001)

**Figure 7.** Comparison celecoxib versus placebo, outcome: pain VAS, 8 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo Mean</th>
<th>SD</th>
<th>Total</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bensen WC 1999</td>
<td>-27.07</td>
<td>9.26</td>
<td>197</td>
<td>-16.74</td>
<td>38.18</td>
<td>203</td>
<td>-0.37</td>
<td>[-0.57, -0.17]</td>
</tr>
<tr>
<td>Bingham 2007 study 1</td>
<td>-24.83</td>
<td>23.14</td>
<td>192</td>
<td>-15.47</td>
<td>20.28</td>
<td>85</td>
<td>-0.40</td>
<td>[-0.68, -0.16]</td>
</tr>
<tr>
<td>Bingham 2007 study 2</td>
<td>-27.84</td>
<td>23.87</td>
<td>202</td>
<td>-16.27</td>
<td>19.68</td>
<td>117</td>
<td>-0.51</td>
<td>[-0.75, -0.28]</td>
</tr>
<tr>
<td>DeLemos 2011</td>
<td>-33.40</td>
<td>29.86</td>
<td>202</td>
<td>-23.88</td>
<td>29.69</td>
<td>200</td>
<td>0.32</td>
<td>[-0.52, -0.13]</td>
</tr>
<tr>
<td>Kivitz 2001</td>
<td>-23.30</td>
<td>17.20</td>
<td>207</td>
<td>-11.11</td>
<td>14.93</td>
<td>217</td>
<td>0.76</td>
<td>[-0.96, -0.57]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>1000</td>
<td></td>
<td></td>
<td>822</td>
<td></td>
<td></td>
<td>-0.48</td>
<td>[-0.64, -0.31]</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.02; Chi² = 11.91, df = 4 (P = 0.02); I² = 66%
Test for overall effect: Z = 5.68 (P < 0.00001)

**Figure 8.** Comparison celecoxib versus placebo, outcome: pain VAS, 12 weeks
Figure 9. Comparison celecoxib versus placebo, outcome: pain VAS, 13 weeks

Figure 10. Comparison celecoxib versus placebo, outcome: pain WOMAC, 1 week
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib: Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo: Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference (IV, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bensen WG 1999</td>
<td>-2.92</td>
<td>3.0878</td>
<td>197</td>
<td>-0.83</td>
<td>3.562</td>
<td>203</td>
<td>7.6%</td>
<td>-0.63 [-0.83, -0.42]</td>
</tr>
<tr>
<td>Birbara 2006 Study 1</td>
<td>-25.3</td>
<td>53.707</td>
<td>156</td>
<td>-17.72</td>
<td>51.3125</td>
<td>78</td>
<td>5.4%</td>
<td>-0.14 [-0.41, 0.13]</td>
</tr>
<tr>
<td>Birbara 2006 Study 2</td>
<td>-24.71</td>
<td>55.0669</td>
<td>164</td>
<td>-19.18</td>
<td>52.56</td>
<td>81</td>
<td>5.5%</td>
<td>-0.10 [-0.37, 0.16]</td>
</tr>
<tr>
<td>Boswell 2008 Study A</td>
<td>-14.0</td>
<td>18.0346</td>
<td>105</td>
<td>-11.9</td>
<td>20.0998</td>
<td>101</td>
<td>5.3%</td>
<td>-0.11 [-0.38, 0.16]</td>
</tr>
<tr>
<td>Boswell 2008 Study B</td>
<td>-23.2</td>
<td>19.6614</td>
<td>163</td>
<td>-14.9</td>
<td>18.8674</td>
<td>167</td>
<td>7.0%</td>
<td>-0.43 [-0.65, -0.21]</td>
</tr>
<tr>
<td>Conaghan 2013</td>
<td>-0.99</td>
<td>1.2211</td>
<td>233</td>
<td>-0.737</td>
<td>1.2053</td>
<td>227</td>
<td>8.3%</td>
<td>-0.21 [-0.39, 0.02]</td>
</tr>
<tr>
<td>DeLemos 2011</td>
<td>-122.39</td>
<td>96.1</td>
<td>202</td>
<td>-87.23</td>
<td>103.5</td>
<td>200</td>
<td>7.7%</td>
<td>-0.35 [-0.55, -0.15]</td>
</tr>
<tr>
<td>Fleischmann 2005</td>
<td>-2.9</td>
<td>3.6</td>
<td>444</td>
<td>-1.6</td>
<td>3.1</td>
<td>231</td>
<td>9.3%</td>
<td>-0.38 [-0.54, -0.22]</td>
</tr>
<tr>
<td>Kivitz 2001</td>
<td>-2.52</td>
<td>5.1795</td>
<td>207</td>
<td>-0.6</td>
<td>3.6827</td>
<td>217</td>
<td>7.9%</td>
<td>-0.43 [-0.62, -0.24]</td>
</tr>
<tr>
<td>Lehman 2005</td>
<td>-2.5</td>
<td>3.17</td>
<td>420</td>
<td>-1.1</td>
<td>3.08</td>
<td>424</td>
<td>10.4%</td>
<td>-0.45 [-0.58, -0.31]</td>
</tr>
<tr>
<td>Smugar 2006 Study 1</td>
<td>-33.7</td>
<td>25.1594</td>
<td>447</td>
<td>-22.7</td>
<td>25.8577</td>
<td>146</td>
<td>8.1%</td>
<td>-0.43 [-0.62, -0.25]</td>
</tr>
<tr>
<td>Smugar 2006 Study 2</td>
<td>-32.2</td>
<td>25.7091</td>
<td>459</td>
<td>-18.4</td>
<td>24.4949</td>
<td>150</td>
<td>8.1%</td>
<td>-0.54 [-0.73, -0.36]</td>
</tr>
<tr>
<td>Tannenbaum 2004</td>
<td>-2.4</td>
<td>3.1</td>
<td>481</td>
<td>-1.4</td>
<td>3.3</td>
<td>243</td>
<td>9.5%</td>
<td>-0.32 [-0.47, 0.16]</td>
</tr>
</tbody>
</table>

Total (95% CI) | 3678 | 2468 | 100.0% | -0.37 [-0.44, -0.29] |

Heterogeneity: $\tau^2 = 0.01$, $\chi^2 = 25.30$, df = 12 ($P = 0.01$); $I^2 = 53$
Test for overall effect: $Z = 9.13$ ($P < 0.00001$)

Figure 11. Comparison celecoxib versus placebo, outcome: pain WOMAC, 2 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib: Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo: Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference (IV, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeLemos 2011</td>
<td>-141.85</td>
<td>96.1</td>
<td>202</td>
<td>-102.09</td>
<td>103.5</td>
<td>200</td>
<td>67.6%</td>
<td>-0.40 [-0.59, -0.20]</td>
</tr>
<tr>
<td>Lee M 2017</td>
<td>-4.4</td>
<td>8.1883</td>
<td>145</td>
<td>-2.1</td>
<td>7.1622</td>
<td>71</td>
<td>32.4%</td>
<td>-0.30 [-0.59, -0.02]</td>
</tr>
</tbody>
</table>

Total (95% CI) | 347 | 271 | 100.0% | -0.37 [-0.53, -0.20] |

Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 0.28$, df = 1 ($P = 0.60$); $I^2 = 0$
Test for overall effect: $Z = 4.43$ ($P < 0.00001$)

Figure 12. Comparison celecoxib versus placebo, outcome: pain WOMAC, 3 weeks
### Figure 13. Comparison celecoxib versus placebo, outcome: pain WOMAC, 4 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib</th>
<th>Placebo</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>Birbara 2006 Study 1</td>
<td>-27.2</td>
<td>55.3307</td>
<td>156</td>
<td>-19.17</td>
</tr>
<tr>
<td>Birbara 2006 Study 2</td>
<td>-29.58</td>
<td>56.3475</td>
<td>164</td>
<td>-21.03</td>
</tr>
<tr>
<td>Boswell 2008 Study A</td>
<td>-19.52</td>
<td>20.5964</td>
<td>105</td>
<td>-15.36</td>
</tr>
<tr>
<td>Boswell 2008 Study B</td>
<td>-25.61</td>
<td>20.6828</td>
<td>163</td>
<td>-21.11</td>
</tr>
<tr>
<td>Schnitzer 2011</td>
<td>-3.08</td>
<td>4.28</td>
<td>419</td>
<td>-1.76</td>
</tr>
<tr>
<td>Smugar 2006 Study 1</td>
<td>-35.08</td>
<td>25.1594</td>
<td>447</td>
<td>-25.02</td>
</tr>
<tr>
<td>Smugar 2006 Study 2</td>
<td>-32.7</td>
<td>26.7804</td>
<td>459</td>
<td>-20.09</td>
</tr>
</tbody>
</table>

Total (95% CI)  1913  1139  100.0%  -0.30 [-0.39, -0.21]

Heterogeneity: Tau² = 0.00; Chi² = 7.83, df = 6 (P = 0.25); I² = 23%
Test for overall effect: Z = 6.57 (P < 0.00001)

### Figure 14. Comparison celecoxib versus placebo, outcome: pain WOMAC, 6 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib</th>
<th>Placebo</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>Birbara 2006 Study 1</td>
<td>-19.24</td>
<td>39.6546</td>
<td>78</td>
<td>5.4%</td>
</tr>
<tr>
<td>Birbara 2006 Study 2</td>
<td>-20.95</td>
<td>39.69</td>
<td>81</td>
<td>5.5%</td>
</tr>
<tr>
<td>Boswell 2008 Study A</td>
<td>-16.9</td>
<td>36.883</td>
<td>101</td>
<td>5.4%</td>
</tr>
<tr>
<td>Conaghan 2013</td>
<td>-1.18</td>
<td>1.5067</td>
<td>227</td>
<td>8.2%</td>
</tr>
<tr>
<td>DeLemos 2011</td>
<td>-2.98</td>
<td>125.885</td>
<td>200</td>
<td>7.7%</td>
</tr>
<tr>
<td>Essex 2016</td>
<td>-1.58</td>
<td>3.6845</td>
<td>59</td>
<td>4.5%</td>
</tr>
<tr>
<td>Gabinetsky 2003</td>
<td>-2.6</td>
<td>3.9392</td>
<td>64</td>
<td>4.9%</td>
</tr>
<tr>
<td>Gordo 2017</td>
<td>-1.7</td>
<td>4.1509</td>
<td>78</td>
<td>5.3%</td>
</tr>
<tr>
<td>Hochberg 2011 Study 307</td>
<td>-31.95</td>
<td>26.5</td>
<td>124</td>
<td>7.0%</td>
</tr>
<tr>
<td>Hochberg 2011 Study 309</td>
<td>-35.26</td>
<td>26.5</td>
<td>122</td>
<td>7.0%</td>
</tr>
<tr>
<td>Lee M 2017</td>
<td>-2.6</td>
<td>7.5835</td>
<td>71</td>
<td>5.1%</td>
</tr>
<tr>
<td>Rother 2007</td>
<td>-1.4</td>
<td>9.4</td>
<td>102</td>
<td>5.4%</td>
</tr>
<tr>
<td>Smugar 2006 Study 1</td>
<td>-22.17</td>
<td>17.3586</td>
<td>93</td>
<td>6.6%</td>
</tr>
<tr>
<td>Smugar 2006 Study 2</td>
<td>-16.7</td>
<td>17.265</td>
<td>92</td>
<td>6.5%</td>
</tr>
<tr>
<td>Williams 2000</td>
<td>-3.4</td>
<td>3.6</td>
<td>231</td>
<td>6.4%</td>
</tr>
<tr>
<td>Williams 2001</td>
<td>-1.8</td>
<td>3.3</td>
<td>243</td>
<td>8.3%</td>
</tr>
</tbody>
</table>

Total (95% CI)  3211  1917  100.0%  -0.35 [-0.43, -0.27]

Heterogeneity: Tau² = 0.01; Chi² = 27.16, df = 15 (P = 0.03); I² = 45%
Test for overall effect: Z = 8.52 (P < 0.00001)
Figure 15. Comparison celecoxib versus placebo, outcome: pain WOMAC, 8 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib</th>
<th>Placebo</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
</tr>
<tr>
<td>Boswell 2008 Study B</td>
<td>-28.25</td>
<td>23.3639</td>
<td>163</td>
</tr>
<tr>
<td>Schnitzer 2011</td>
<td>-3.64</td>
<td>4.28</td>
<td>419</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>582</td>
<td>583</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.01; Chi² = 2.73, df = 1 (P = 0.10); I² = 63%
Test for overall effect: Z = 2.68 (P = 0.007)

Figure 16. Comparison celecoxib versus placebo, outcome: pain WOMAC, 9 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib</th>
<th>Placebo</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
</tr>
<tr>
<td>Conaghan 2013</td>
<td>-1.71</td>
<td>1.5264</td>
<td>233</td>
</tr>
<tr>
<td>DeLemos 2011</td>
<td>-160.46</td>
<td>127.914</td>
<td>202</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>435</td>
<td>427</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%
Test for overall effect: Z = 3.50 (P = 0.0005)

Figure 17. Comparison celecoxib versus placebo, outcome: pain WOMAC, 12 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib</th>
<th>Placebo</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
</tr>
<tr>
<td>Bensen WC 1999</td>
<td>-3.06</td>
<td>4.4914</td>
<td>197</td>
</tr>
<tr>
<td>Bingham 2007 study 1</td>
<td>-24.7</td>
<td>22.9</td>
<td>236</td>
</tr>
<tr>
<td>Bingham 2007 study 2</td>
<td>-26.7</td>
<td>24.1</td>
<td>246</td>
</tr>
<tr>
<td>Boswell 2008 Study B</td>
<td>-28</td>
<td>23.8746</td>
<td>163</td>
</tr>
<tr>
<td>Conaghan 2013</td>
<td>-1.9</td>
<td>1.62</td>
<td>233</td>
</tr>
<tr>
<td>DeLemos 2011</td>
<td>-130</td>
<td>127.914</td>
<td>202</td>
</tr>
<tr>
<td>Hochberg 2011 Study 307</td>
<td>-41.8</td>
<td>26.5</td>
<td>242</td>
</tr>
<tr>
<td>Hochberg 2011 Study 309</td>
<td>-42.9</td>
<td>26.5</td>
<td>244</td>
</tr>
<tr>
<td>Kivitz 2001</td>
<td>-2</td>
<td>4.32</td>
<td>207</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>1970</td>
<td>1498</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 9.21, df = 8 (P = 0.32); I² = 13%
Test for overall effect: Z = 8.58 (P < 0.00001)
### Figure 18. Comparison celecoxib versus placebo, outcome: pain WOMAC, 13 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fleischmann 2005</td>
<td>-3.5</td>
<td>4.11</td>
<td>444</td>
<td>-2.3</td>
<td>3.9</td>
<td>231</td>
<td>18.6%</td>
<td>-0.30 [-0.46, -0.14]</td>
<td></td>
</tr>
<tr>
<td>Lehman 2005</td>
<td>-3.4</td>
<td>3.67</td>
<td>368</td>
<td>-2.5</td>
<td>4.12</td>
<td>360</td>
<td>22.4%</td>
<td>-0.23 [-0.38, -0.08]</td>
<td></td>
</tr>
<tr>
<td>Schnitzer 2011</td>
<td>-3.6</td>
<td>4.28</td>
<td>327</td>
<td>-2.2</td>
<td>3.94</td>
<td>287</td>
<td>18.7%</td>
<td>-0.34 [-0.50, -0.18]</td>
<td></td>
</tr>
<tr>
<td>Sheldon 2005</td>
<td>-3.4</td>
<td>4.21</td>
<td>393</td>
<td>-2.3</td>
<td>3.84</td>
<td>382</td>
<td>23.8%</td>
<td>-0.27 [-0.41, -0.13]</td>
<td></td>
</tr>
<tr>
<td>Tannenbaum 2004</td>
<td>-3.1</td>
<td>3.8</td>
<td>401</td>
<td>-2.4</td>
<td>3.8</td>
<td>200</td>
<td>16.5%</td>
<td>-0.18 [-0.35, -0.01]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>1933</strong></td>
<td></td>
<td></td>
<td><strong>1460</strong></td>
<td></td>
<td></td>
<td><strong>100.0%</strong></td>
<td></td>
<td><strong>-0.27 [-0.33, -0.20]</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 2.07, df = 4 (P = 0.72); I² = 0%
Test for overall effect: Z = 7.54 (P < 0.00001)

### Figure 19. Comparison celecoxib versus placebo, outcome: pain WOMAC, 24 weeks

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Celecoxib Mean</th>
<th>SD</th>
<th>Total</th>
<th>Placebo Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clegg 2006</td>
<td>-100</td>
<td>102.9</td>
<td>318</td>
<td>-86.1</td>
<td>114.2</td>
<td>313</td>
<td>100.0%</td>
<td>-0.13 [-0.28, 0.03]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>318</strong></td>
<td></td>
<td></td>
<td><strong>313</strong></td>
<td></td>
<td></td>
<td><strong>100.0%</strong></td>
<td></td>
<td><strong>-0.13 [-0.28, 0.03]</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: Not applicable
Test for overall effect: Z = 1.60 (P = 0.11)
5. DISCUSSION
In this study we found a decreasing trend of a numerical indicator for efficacy of celecoxib for treatment of pain in studies that compared celecoxib and placebo, and reported pain results with VAS and WOMAC scales. It has to be emphasized that these studies were relatively short, considering the chronic nature of osteoarthritis; the longest follow-up in the group of studies that reported VAS pain was 13 weeks and those that reported WOMAC pain was 24 weeks. There was only one study for data analysis for the domain WOMAC at 24 weeks. Additionally, SMD remained fairly constant with VAS over most follow up times (SMD at 2 weeks: -0.50; SMD at 12 weeks: -0.48) as well as with WOMAC (SMD at 2 weeks: -0.32; SMD at 12 weeks: -0.32). The later follow-up times showed a decreased SMD for VAS at 13 weeks (SMD: -0.23) and for WOMAC at 24 weeks (SMD: -0.13).

This study was conducted because we observed discrepancies between the 2017 Cochrane review on celecoxib for osteoarthritis and results published by Moore et al. [22]. Moore et al. used regression models to assess correlation between efficacy comparing diclofenac, ibuprofen, naproxen, celecoxib or etoricoxib with each other or with placebo at 2, 6 and 12 weeks. Their evidence base consisted of 50 RCTs used for analysis. The results suggested that average change from baseline (CFB) of VAS pain at all time-points were highly associated. Therefore, pain VAS at 2 weeks was predictive of pain VAS at 6 and 12 weeks [22]. Similar predictive effects of early response to NSAIDs in predicting late response where demonstrated earlier by Bingham et al. in a pooled analysis of 2 identical 26-week studies testing etoricoxib, celecoxib and placebo in patients with OA of the hip and knee. With active treatment 75% of patients who were responders at 2 weeks were also responders at 12 weeks [64]. Both groups analyzed follow-up times up to 12 weeks.

Moore et al. concluded that clinical trials for efficacy of NSAIDs can be shorter as early response is likely associated with late response and early treatment failure is likely to be associated with treatment failure in general [22].

In line with studies of Moore et al. and Bingham et al. this study showed that SMDs observed at 2, 6 and 12 weeks remained relatively constant, for pain measured with both VAS and WOMAC.

However, our data also indicate that efficacy of celecoxib measured with pain VAS decreased from week 2 (SMD: -0.50) to weeks 13 (SMD: -0.23), as well as with pain WOMAC from week 2 (SMD: -0.32) to week 24 (SMD: -0.13).

Based on our findings, we can provide suggestions concerning future designs of RCTs as well as clinical decision-making that are contrary to conclusions of Moore et al. and Bingham et al.
Our results indicate that trialists should conduct studies with longer follow-up times in order to adequately assess efficacy of celecoxib over a prolonged period. In a chronic condition such as OA, for which patient will require adequate analgesic treatment over extended periods, it could be misleading to conduct studies with short follow-up times. The overwhelming majority of the studies included in our analysis was conducted within 13 weeks, with only one that had 24-week follow-up and none longer than that. Most patients with OA are treated with analgesics for far longer periods than this and a potential decrease in efficacy of celecoxib with longer treatment duration has to be addressed due to the multiple reasons:

Firstly, inadequate pain relief (IPR) could lead to amplification of pain response through maladaptive neuroplastic mechanisms [65]. Secondly, pain is not merely a symptom but also a disease on its own which can manifest with chronic pain leading to significant morbidity and health care related costs [66]. Additionally, almost one half of all patients are dissatisfied with the control of OA pain provided by NSAID therapy, according to a study by Taylor-Stokes et al. from 2013 [67]. In a 2014 prospective multinational longitudinal study about real-world therapies for OA, it was shown that inadequate pain relief is a highly relevant problem among patients with OA. Predictors for IPR included, female sex, higher body mass index (BMI), longer OA duration, bilateral knee OA, depression and diabetes. IPR was associated with functional loss and impaired quality of life [68]. Furthermore, patients with OA presumably prefer medication with a longer treatment effect. This was shown in a study conducted by Oxford University and published in April 2019. Researchers conducted a discrete choice experiment with 300 residents of the United Kingdom with hip and/or knee OA to quantify patients’ preferences for the duration of treatment effect relative to treatment benefits and risks. Results showed that pain, severity and duration of treatment effect had the greatest influence on medication preference. This suggests that patients would be willing to take medication which is less effective in relieving pain if the effect of the medication lasts longer. Moreover, participants were willing to accept an increase in the risk of heart attack of 2.6% to increase the duration of the treatment effect from one to 12 month [69]. Similar to results of our study, the authors appealed to future trialists to conduct clinical trials with longer follow-up for investigating treatment effect to evaluate if significant benefit is sustained over time. Duration of treatment effect seems to be an important factor in the medication choices of people with osteoarthritis and therefore should not be dismissed by researchers and physicians.
All in all, it remains unclear how effective NSAIDs are in the treatment of chronic conditions such as OA. It was shown in the 2017 review about celecoxib for osteoarthritis that celecoxib proved only to be negligibly better than placebo. Results from this review suggest that the effectiveness of celecoxib could decrease with treatment duration above 12 weeks. We appeal to researchers to conduct additional RCTs with longer follow-up times to address this issue. Stratification of patients according to known risk factors for IPR could be helpful in further studies.

Joint pain is complex and yet to be understood thoroughly, especially when chronic. A multidisciplinary approach in management is crucial in order to provide adequate treatment and improved quality of life.

There remains a compelling need for effective, well-tolerated analgesic drugs in order to limit inadequate therapy for patients with conditions that can lead to the development of chronic pain and all its associated sequelae. There have been significant advances in our understanding of the neurobiology of joint pain in OA. Potentially new targets for novel analgesics have been identified [10].

In addition to clarifying the effectiveness of traditionally used drugs like NSAIDs and coxibs, future research should focus on novel analgesics in order to bridge the gap between our understanding of pain and clinical practice.

Our study had several limitations. We used studies identified in the 2017 Cochrane review, as well as additional studies, but we did not systematically search for all potentially newly published studies. However, despite this lack of additional systematic search, we are not aware of any new studies that have studied efficacy of celecoxib vs. placebo with longer follow-up times than reported in this study. As already indicated in the 2017 Cochrane review about celecoxib for osteoarthritis, included studies had major limitations and evidence quality was poor. Therefore, results should be interpreted with caution. None of the studies in that Cochrane review had low risk of bias for all seven assessed domains. Selection bias was poorly reported in most trials and attrition bias was high in most trials. Additionally, there was selective reporting in about one third of trials.
We did not assess risk of bias in additional studies used in this analysis. Another limitation is that there was only one study available for the analysis of WOMAC at 24 weeks as well as for VAS at 30 days.

Due to involvement of industry sponsors in most of the analyzed trials there is a reason to be reserved because it has been shown that such sponsorship may lead to more favorable results of the intervention. Cochrane review of Lundh et al. included 75 studies that have analyzed whether industry sponsored drug and medical device studies have more favorable results compared to studies without such sponsorship and they found that there is an industry bias which cannot be explained by standard risk of bias assessment. [70]

Patients in most studies included in the Cochrane review on celecoxib for OA were allowed to use rescue medication in case the study medication did not provide adequate pain relieve. This is of course necessary from an ethical standpoint, but concerning adequacy of comparative results this represents a confounding factor. Trialists did not measure amount of rescue medication used by patients and did not include this factor in their analyses of drug efficacy.

Some studies had to be excluded from this analysis because they did not report standard deviation or standard error with their main effect. Additionally, we had trouble obtaining complete data from certain studies as it was not provided in the published study and further requests to study sponsors were not successful. Study authors and sponsors should provide open access to their full data sets in order to make use of complete data sets for future analysis. Furthermore, in trials with multiple follow-up time points, the trialists should not report only results for the final follow-up, but also for all measured follow-ups. Lastly, pain experience in OA patients is complex and potentially not measured adequately by existing measures which are used in current analysis.
6. CONCLUSION
1. It remains unclear how effective celecoxib 200 mg is in comparison to placebo for the treatment of osteoarthritis.

2. Our data indicates that efficacy of celecoxib 200 mg could decrease over longer follow-up times due to decreasing SMDs found at 13 weeks for VAS and at 24 weeks for WOMAC pain scales.

3. Previous research showed a similar trend of efficacy for celecoxib over 12 weeks follow-up. Data from later time points used in our study suggest a decrease of efficacy with longer follow-up times.

4. Current research about use of celecoxib in osteoarthritis is potentially insufficient for patient groups taking Celecoxib for a prolonged period of time.

5. Osteoarthritis is a complex disease with significant socio-economic burden. In order to optimize treatment and reduce disease related negative health outcomes new treatment modalities are needed that bridge the gap between our understanding of Osteoarthritis pain and clinical practice.
7. SUMMARY
**Objectives:** The aim of this Thesis was to conduct comprehensive analysis of efficacy data for pain in randomized controlled trial RCTs about Celecoxib in osteoarthritis (OA). The ultimate purpose of this study is to improve long-term management of pain for patients suffering from OA by guiding clinical decision making, and to create evidence that will inform design of future RCTs about OA.

**Material and Methods:** This was a methodological study in which publicly available data from RCTs were analyzed. RCTs analyzing the effects of 200 mg celecoxib vs. placebo on pain intensity with the Visual Analog Scale (VAS) and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score were included. Random effect meta-analysis was used for different pain outcome measures and different follow-up times. Standardized mean differences were used to report the data.

**Results:** We found a decreasing trend of a numerical indicator for efficacy of celecoxib for treatment of pain in RCTs comparing Celecoxib 200 mg to Placebo and reported pain results with the VAS and WOMAC scale. Standardized mean differences remained relatively constant with VAS and WOMAC over most follow-up times. The later follow-up times showed a decreased SMD for VAS at 13 weeks as well as for WOMAC with 24 weeks.

**Conclusion:** Our data indicates that efficacy of celecoxib 200 mg could decrease over longer follow-up times. Future trials should include assessment at longer follow-up times for adequate assessment of efficacy and safety of celecoxib.
8. CROATIAN SUMMARY
Naslov na hrvatskom jeziku: Promjene intenziteta boli u različitim vremenima praćenja u randomiziranim kontroliranim pokusima o celekoksibu za osteoartritis

Ciljevi: Cilj ove disertacije bio je provesti detaljnu analizu o ishodima koji opisuju intenzitet boli u randomiziranim kontroliranim pokusima (engl. randomized controlled trials; RCTs) o celekoksibu za osteoartritis (OA). Konačni cilj je dati nove informacije za praksu i omogućiti ustroj boljih RCT-ova u budućnosti.

Metode: Provedeno je metodološko istraživanje u kojem su analizirani javno dostupni podatci iz RCT-ova. Uključeni su RCT-ovi koji su analizirali djelotvornost i sigurnost celekoksiba 200 mg u usporedbi s placebom. Intenzitet boli je analiziran vizualno-analognom ljestvicom (engl. Visual-Analog Scale; VAS) i WOMAC ljestvicom (engl. Western Ontario and McMaster Universities Osteoarthritis Index). Provedena je meta-analiza nasumičnih učinaka (engl. random effect meta-analysis) kako bi se analizirao zbirni učinak za različite mjere ishoda u različitim vremenima praćenja. Standardizirane srednje razlike (engl. standardized mean differences; SMD) su korištene za prikaz podataka.

Rezultati: Uočen je trend smanjenja SMD za djelotvornost celekoksiba za liječenje OA prema mjernim instrumentima VAS i WOMAC u kasnijim vremenima praćenja. Pokusi koji su koristili ljestvicu VAS trajali su najviše 13 tjedana, a pokusi koji su koristili ljestvicu WOMAC najviše 24 tjedna.

Zaključak: Dobiveni podatci ukazuju da bi djelotvornost celekoksiba za liječenje boli u OA mogla biti manja s duljim vremenom primjene lijeka. Novi klinički pokusi trebali bi uključiti dulje vrijeme praćenja kako bi se dobili odgovarajući podatci iz istraživanja o dugoročnoj djelotvornosti i sigurnosti celekoksiba u OA koji je kronična bolest.
### Personal Data

<table>
<thead>
<tr>
<th>Personal Data</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name</strong></td>
<td>Nys Siebers</td>
</tr>
<tr>
<td><strong>Address</strong></td>
<td>Spinčićeva 25c 21000 Split</td>
</tr>
<tr>
<td><strong>Date of Birth</strong></td>
<td>19/08/1990</td>
</tr>
<tr>
<td><strong>Place of Birth</strong></td>
<td>Berlin / Germany</td>
</tr>
</tbody>
</table>

### Education

<table>
<thead>
<tr>
<th>Year</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013 - 2019</td>
<td>University of Split, School of Medicine</td>
</tr>
<tr>
<td>09/2016</td>
<td>United States Medical Licensing Exam</td>
</tr>
<tr>
<td>01/2011-12/2011</td>
<td>Paramedic Training at NAW Berlin</td>
</tr>
<tr>
<td>2002-2010</td>
<td>A-levels at Humboldt Gymnasium Berlin</td>
</tr>
</tbody>
</table>

### Work experience

<table>
<thead>
<tr>
<th>Year</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/2016-06/2016</td>
<td>Tutor for Anatomy at Split School of Medicine</td>
</tr>
<tr>
<td>01/2013-07/2013</td>
<td>Paramedic at Rettungsdienst Oberhavel GmbH</td>
</tr>
<tr>
<td>03/2013-08/2013</td>
<td>First Aid Instructor at NAW Berlin</td>
</tr>
<tr>
<td>12/2012-08/2013</td>
<td>Event Paramedic at Medics Berlin</td>
</tr>
<tr>
<td>05/2012-11/2012</td>
<td>Paramedic at DRK Freital</td>
</tr>
<tr>
<td>06/2011-08/2011</td>
<td>Various internships as part of paramedic training</td>
</tr>
</tbody>
</table>
10. REFERENCES
15. What is a Systematic Review? [09/03/2019]. Available at: https://consumers.cochrane.org/what-systematic-review
19. About Cochrane Reviews. [05/03/2019]. Available at: https://www.cochranelibrary.com/about/about-cochrane-reviews.
20. Overview of reviews. [07/03/2019]. Available at: https://training.cochrane.org/resource/overviews-reviews.
25. Introduction to RevMan. [06/03/2019]. Available at: https://training.cochrane.org/resource/introduction-revman
38. Hochberg MC, Fort JG, Svensson O, Hwang C, Sostek M. Fixed-dose combination of enteric-coated naproxen and immediate-release esomeprazole has comparable efficacy to


Wirksamkeit und Verträglichkeit von retardiertem Acemetacin im Vergleich zu Celecoxib.


