Duration of Symptom Relief and Early Trajectory of Adverse Events for Oral NSAIDs in Knee Osteoarthritis: A Systematic Review and Meta-analysis

Mikala C. Osani, B.A.[1], **Elizaveta E. Vaysbrot, M.D., M.S.**[1], **Mengyu Zhou, M.D.**[1], **Timothy E. McAlindon, M.D., M.P.H.**[1], **Raveendhara R. Bannuru, M.D., Ph.D.**[1]

[1.]Center for Treatment Comparison and Integrative Analysis (CTCIA), Division of Rheumatology, Allergy, and Immunology, Tufts Medical Center, Boston, MA

Abstract

Objective: Despite an extensive body of research on NSAIDs in osteoarthritis, the duration of their efficacy and timeline of adverse event (AE) onset have been understudied. We conducted a systematic review and meta-analyses from 2 to 26 weeks to characterize the efficacy and AE trajectories of oral NSAIDs in knee osteoarthritis.

Methods: We searched MEDLINE, EMBASE, Web of Science, Google Scholar, and the Cochrane Database from inception to May 2018. RCTs assessing the efficacy and/or safety of FDA-approved NSAIDs in knee osteoarthritis patients were included. Two independent reviewers assessed quality and extracted data. We calculated standardized mean differences and risk ratios with 95% confidence intervals.

Results: We included 72 RCTs (26,424 participants). NSAIDs demonstrated moderate, statistically significant effects on pain that peaked at 2 weeks (SMD −0.43 [−0.48, −0.38]), but the magnitude of the effects decreased over time. The results for function were similar. The incidence of GI AEs was significantly higher in NSAID users than placebo users as early as 4 weeks (RR 1.38 [1.21, 1.57]). The incidence of CV AEs in NSAID users was not significantly different from placebo. Most GI and CV AEs were transient and of minor severity.

Conclusion: NSAIDs produced significant pain and function improvements that peaked at 2 weeks but decreased over time. The incidence of minor GI and CV AEs consistently rose, reaching significance as early as 4 weeks. Clinicians should weigh the durability of efficacy with the early

onset of minor AEs along with patient tolerability and preferences when formulating an NSAID regimen.

INTRODUCTION:

Osteoarthritis (OA) is a leading cause of pain and disability among adults in the United States[1], with involvement of the knee joint accounting for more than 80% of the disease's disability burden. The prevalence of the disease is rising, and approximately 14 million adults in the U.S. are now suffering from symptomatic knee OA[2–4]. Since the natural history of OA is long, patients may need therapy for many years, even after arthroplasty.

Oral non-steroidal anti-inflammatory drugs (NSAIDs) are the pharmaceuticals used most frequently for pain control and are routinely recommended in OA Clinical Practice Guidelines[5,6]. In the U.S., 65% of patients with osteoarthritis are prescribed NSAIDs[7]. Given the current need to limit utilization of opioid medications, NSAIDs can be expected to play an even larger role in clinical practice[8]. Despite widespread use, a gap currently exists in our knowledge regarding the consistency and duration of the beneficial effects of NSAIDs on pain and functional outcomes in patients with knee OA, and the efficacy of these drugs has largely been tested in RCTs of short-term duration. For clinicians treating chronic conditions like OA that involve long-term management, the degree of superiority of a treatment over placebo is often of equal importance to its duration of efficacy.

In addition to the uncertainty in efficacy trajectory, there is a lack of research on the timing and evolution of adverse reactions to NSAIDs used in OA. NSAIDs are associated with cardiovascular (CV) adverse events, kidney injury, and gastrointestinal (GI) toxicity; the latter is shown to be likelier for non-selective NSAIDs than for selective COX-2 inhibitors[9–12]. Furthermore, the general OA population, which is characterized by older age and a more frequent use of concomitant medications, could be at a higher risk for NSAIDs-associated complications. Though serious GI and CV risks such as GI bleed or myocardial infarction are associated with prolonged NSAID use, minor adverse events contributing to patient discomfort may begin to manifest even when the treatment duration is relatively short[10,13].

Given the chronic nature of knee OA symptoms and the resulting need for long-term therapeutic solutions, it is important to assess the benefits and risks of any drug- including NSAIDs- in its temporal context. Understanding the durability of the efficacy of oral NSAIDs, as well as the time course of onset of minor adverse effects, is key to this decision-making process. Therefore, we conducted a systematic review and meta-analysis to comprehensively and quantitatively characterize the efficacy trajectory of oral NSAIDs on pain and functional improvement and to summarize the timing of onset and the subsequent progression of minor gastrointestinal and cardiovascular adverse events in patients with knee OA.

METHODS:

Data Sources/Searches:

We searched Medline, EMBASE, Web of Science, Google Scholar, and the Cochrane Central Register of Controlled Trials from inception to May 3, 2018 (eTable 1). We hand-searched reference lists of relevant systematic reviews and meta-analyses and within supplements of conference proceedings that had been published up to May 2018. We limited our search to randomized placebo-controlled trials (RCTs) involving non-steroidal anti-inflammatory drugs (NSAIDs) in human subjects with knee osteoarthritis (OA). No restrictions were placed on publication date, status, or language. We adhered to the PRISMA guidelines, but we elected not to register our study protocol in the PROSPERO database (eTable 5).

Study Selection:

We included RCTs that assessed the efficacy and/or safety of FDA-approved NSAIDs versus placebo in patients with knee OA. We included studies that involved multiple treatment arms, as well as studies that involved treatment modalities other than NSAIDs, as long as the study compared at least one FDA-approved NSAID at an approved dosage against a placebo arm. We included combined knee and hip studies if they either reported separate results for the knee or if they had included more than 70% knee OA patients. Non-randomized studies and studies in which the location of OA was undefined were excluded. Each abstract recovered by the search was screened by two independent reviewers (MO, RB), in line with the pre-established inclusion and exclusion criteria. Full manuscripts of abstracts that were included after initial screening were subsequently gathered and assessed for eligibility in further detail by the same two reviewers (MO, RB). Discordant results in inclusion or exclusion that resulted during either screening stage were adjudicated by a third reviewer (EV).

Data Extraction and Quality Assessment:

Data from each RCT were independently extracted by two reviewers (MO, RB). We drafted a data extraction form in Microsoft Excel to gather information on study and population characteristics, NSAID classification, dosage and frequency, rescue medication protocol, pain and functional outcomes, discontinuation rates and reasons, and relevant safety outcomes. We collected pain and functional outcomes that were reported by any validated scale; in the event that more than one scale was reported, results for all scales were collected. Based on their mechanisms of action, we determined three overarching NSAID classes: "Traditional (non-selective) NSAIDs"– an older group of NSAIDs without a strong COX-2 selectivity (e.g., Diclofenac, Ibuprofen, Indomethacin, Naproxen, and Piroxicam); "Selective COX-2 inhibitors"– a newer class of coxib NSAIDs developed specifically for COX-2 selectivity (Celecoxib was the only representative treatment); and "Intermediate COX inhibitors"– those from the traditional cohort of NSAIDs demonstrating relative COX-2 selectivity but with a chemical structure different from coxibs (e.g., Etodolac, Meloxicam, and Nabumetone)[14–16]. In studies assessing multiple doses of NSAIDs, we only collected data on the dose that most closely matched the recommended dosing range for osteoarthritis; information regarding the recommended dosing for this indication was

obtained directly from the FDA website and/or from package inserts. In studies assessing multiple interventions against a common placebo group, we evenly divided the 'shared' group into two or more smaller groups and included them as independent comparators, as referenced in the Cochrane Handbook, section 16.5.4[17]. In order to comprehensively assess the efficacy trajectories of NSAIDs while maintaining the robustness of our analyses, we collected pain and functional data at all reported time points and grouped the data into the following time point categories: 2 weeks (0-2 weeks), 4 weeks (3-6 weeks), 8 weeks (7-10 weeks), 12 weeks (11-16 weeks), and 26 weeks (17-29 weeks). In all circumstances, we prioritized data that were presented in manuscript text or tables over graphical data. Data that were only presented in figures or graphs were recovered using Engauge Digitizer and double-checked by a second reviewer (EV, MO, or RB)[18]. We transferred the outcome data from Excel into RevMan software, and study quality was assessed in RevMan using the Cochrane Risk of Bias Tool[19,20]. Data extraction and quality ratings were reviewed in their entirety for consistency. Discrepancies were arbitrated by a third reviewer (EV).

Outcome Definitions:

We selected the following outcomes of interest: pain, function, rate of discontinuation due to lack of efficacy, rate of discontinuation due to adverse events (AEs), incidence of treatment-related AEs and serious AEs, and incidence of gastrointestinal and cardiovascular AEs. Pain and functional outcomes were reported as the mean change from baseline to follow-up; in our primary analyses of "All Pain" and "All Function", WOMAC scales were prioritized[21]. If no other scales were available, non-standard Likert scales were included in analyses of "All Pain" and "All Function", but were not used in any other analyses. Rates of discontinuation were reported as the number of participants who discontinued treatment or withdrew from the study due to lack of efficacy or due to any AE. Discontinuation rates were collected for active treatment periods only; we did not collect discontinuation data that were reported after treatment had been stopped or changed, or after randomization had been broken. Serious AEs (SAEs) were defined as those explicitly designated by the outcome assessor(s) as "Serious Adverse Events" within the study period. The criteria for Serious Adverse Events have been delineated by the FDA and include adverse events that are potentially life threatening or result in hospitalization, disability or permanent damage, congenital anomalies or birth defects, or death, or events that may jeopardize the patient and may require medical intervention to prevent one of the above outcomes. Treatment-related AEs were specifically collected so as to better highlight the differences between treatment and placebo groups and were defined as any AEs (or "side effects") that were described by the study investigator(s) as "treatment-related" or "drug-related", or were determined to be of "probable", "possible", and/or "certain" relationship to the study treatment. We excluded studies that only reported incidence of "any adverse event" or "treatment-emergent adverse events" from our analysis of treatment-related AEs [22]. We collected incidence of gastrointestinal (GI) and cardiovascular (CV) AEs as the sum total of the respective AEs at the study endpoints, and at separate time points falling within the pre-established follow-up categories of 2, 4, 8, 12, and 26 weeks, as available. Though we anticipated that the majority of adverse events would be minor due to limitations of follow-up time, all GI and CV events were counted, regardless of severity. The AEs that were most commonly observed were

summarized. All safety data were reported as the number of patients experiencing at least one event.

Statistical analysis:

For continuous outcomes, we calculated standardized mean differences (SMDs) and 95% Confidence Intervals (CIs) using the mean change from baseline to follow-up. We conducted meta-analyses using random effects models to account for methodological and clinical heterogeneity. To allow for direct comparability of effect sizes across different outcomes and subgroups, standardized mean differences were utilized for all analyses of continuous outcome measures, regardless of the variation in their scales. We analyzed dichotomous outcomes using the Mantel-Haenszel method and reported the effects as risk ratios (RR) and 95% CI [23]. Heterogeneity was assessed using the I^2 statistic[24]. Analyses were conducted using RevMan software[20]. Funnel plots were visually inspected for asymmetry as a means of assessing publication bias. To aid in the clinical interpretation of SMDs, we utilized the benchmark of 0.37 units for clinical significance (or importance) per the definition published by Wandel, et al.[25].

We planned the following *a priori* subgroup and sensitivity analyses, all of which were contingent upon the availability of data: analyses based on NSAID classification (Selective COX-2 inhibitor vs. Intermediate COX inhibitor vs. Traditional NSAID), analyses limiting by pain (WOMAC vs. VAS) or functional (WOMAC vs. Any other functional scale) scale, analyses limited to knee OA patients, analyses with potential outliers removed (conducted in the event that I^2 was 75%, as referenced in the Cochrane Handbook, sections 9.5.2 and 9.5.3), and analyses limiting by study quality[17]. In sensitivity analyses limiting by study quality, we chose to eliminate studies of "Very Low Quality". "Very Low Quality" studies were defined *a priori* as those which received 2 High Risk of Bias ratings or one specific High Risk Rating in the "Other" category in addition to 2 Unclear Risk ratings or 3 Unclear Risk of Bias ratings in dimensions other than the "Other" category using the Cochrane Risk of Bias tool[19].

RESULTS:

The initial systematic search returned 1,607 potentially relevant abstracts, of which 191 were eligible for full text review. Of the 191 articles that underwent full text review, 72 randomized controlled trials (RCTs) were eligible for our analyses (eFigure 1). The efficacy and/or safety of the following oral NSAIDs were assessed by the included RCTs: Celecoxib (35 RCTs), Naproxen (18 RCTs), Diclofenac (11 RCTs), Nabumetone (7 RCTs), Ibuprofen (6 RCTs), Meloxicam (3 RCTs), Etodolac (2 RCTs), Indomethacin (1 RCT), and Piroxicam (1 RCT).

Baseline characteristics of the included RCTs are reported in eTable 2; a supplementary table of studies that were excluded due to inappropriate population characteristics can be found in eTable 3. The publication dates of included RCTs ranged from 1976 to 2017, and the sample sizes consisting of eligible treatment arms in included RCTs ranged from 47 to 844 (median: 323). The follow-up duration ranged from one week to two years, but 96% of the trials had a duration of 13 weeks or less (median: 6 weeks). Mean age of included

participants ranged from 53 to 69 years (median: 62 years), and the mean body mass index (BMI) of patients ranged from 27 kg/m^2 to 34 kg/m^2 (median: 31.5 kg/m^2). The percentage of females ranged from 49% to 85% (median: 68%). Limited use of acetaminophen as rescue medication was permitted in 69% of the included RCTs.

A summary of study quality assessment is shown in eTable 4, and eFigure 2 depicts the overall risk of bias distribution. The majority of studies were of moderate quality; potential attrition bias and reporting bias were the most common reasons for High Risk of Bias ratings. The majority of RCTs (80%) reported industry sponsorship and/or direct industry involvement of one or more investigator(s).

Overall effects of NSAIDs on Pain & Function

Our primary analyses of pain and function combined all oral NSAIDs, regardless of classification. Results of all analyses of pain and functional outcomes are displayed in Table 1.

NSAIDs showed statistically significant, clinically important effects on pain as early as two weeks from baseline, with a SMD of −0.43 (95% CI −0.48, −0.38). This treatment effect remained statistically significant up to 26 weeks (SMD −0.21 [95% CI −0.39, −0.03]), though the effects attenuated progressively over time and lost clinical significance (Figure 1A). The analysis of pain at 4 weeks demonstrated high heterogeneity ($I^2 = 76\%$), prompting a sensitivity analysis excluding outliers[26,27]. This analysis reduced the I^2 value to 29%, and the treatment effect decreased (SMD −0.36 [−0.40, −0.33]). Sensitivity analyses restricted to knee OA only populations were not notably different from those observed in the primary analysis at any time point. Sensitivity analyses by study quality also demonstrated results similar to the primary analysis, but there was a trend for effect sizes to increase slightly with the removal of Very Low Quality studies (Table 1). There was no notable asymmetry in our visual inspection of the funnel plots at any time point.

With respect to functional improvement, NSAIDs again showed consistent statistically significant benefits compared with placebo, from 2 weeks (SMD −0.45 [−0.52, −0.38]) to 26 weeks (SMD −0.19 [−0.32, −0.07]) (Figure 1B). None of the analyses of functional improvement demonstrated I^2 values necessitating sensitivity analysis. Sensitivity analyses restricted to knee OA only populations demonstrated beneficial effects on functional outcomes that were not notably different from the primary analysis at any time point. Sensitivity analyses limiting by study quality produced results similar to the main analysis, with a tendency for effect sizes to increase (Table 1). Again, there was no notable asymmetry in our visual inspection of the funnel plots at any time point.

Sensitivity analyses restricting by pain or functional assessment scale demonstrated that functional measurements obtained using scales other than the WOMAC (primarily, the Lequesne Algofunctional Index) tended to result in slightly smaller effect sizes (eFigures 3A, 3B).

Overall Safety of NSAIDs

Patients receiving oral NSAIDs were more likely to withdraw due to an adverse event (AE) during a study's treatment period (RR 1.16 [1.02, 1.32]) but were less likely to withdraw due to a lack of efficacy (0.38 [0.34, 0.43]) (Table 2). Patients receiving oral NSAIDs experienced higher incidence of treatment-related AEs (1.21 [1.04, 1.40]), cardiovascular (CV) AEs (1.37 [1.05, 1.77]), and gastrointestinal (GI) AEs (1.36 [1.25, 1.49]) during the study follow-up period. As we had expected, the most commonly reported GI AEs were transient and mild, and included upper abdominal pain, diarrhea, dyspepsia, and nausea. Edema and hypertension were the most commonly reported CV AEs, and they were mild in severity and duration. The incidence of serious AEs (SAEs) over the duration of study did not differ between groups.

Gastrointestinal & Cardiovascular Safety Trajectory of NSAIDs (all classes combined)

We assessed the likelihood of experiencing GI AEs at 2, 4, 12, and 26 weeks and CV AEs at 2, 4, and 12 weeks (Figures 2A, 2B). No RCT reported data on GI or CV safety at 8 weeks. Patients receiving NSAIDs were more likely to experience a minor GI AE as early as 4 weeks after initiating treatment (N= 31 studies; RR 1.38 [95% CI 1.21, 1.57]), at 12 weeks (N= 22 studies; RR 1.36 [1.18, 1.56]), and at 26 weeks (N= 1 study; RR 1.55 [0.52, 4.62]) (Figure 2A). Though the overall risk of developing a minor CV AE was higher in patients using NSAIDs vs. placebo, it did not reach statistical significance at any individual time point in the analysis of all NSAIDs classes combined.

Comparing the Efficacy and Safety of Different Classes of NSAIDs

Traditional NSAIDs performed consistently better than the other classes (Table 1, Figure 3A). At 2 weeks, Traditional NSAIDs demonstrated effects on pain that were 24% and 64% greater than those of Celecoxib and Intermediate COX inhibitors, respectively; at 12 weeks, the effects of Traditional NSAIDs on pain were 33% and 44% greater than those of Celecoxib and Intermediate COX inhibitors, respectively. Only studies assessing the efficacy of Celecoxib extended to 26 weeks, so a comparison of the different classes could not be undertaken at this time point.

Traditional NSAIDs also outperformed the other classes with regard to functional improvement, demonstrating effects that ranged from 14% to 42% better than those of Celecoxib (Table 1, Figure 3B). Interestingly, for both pain and functional efficacy outcomes, Celecoxib outperformed Intermediate COX inhibitors at most time points. Once again, due to a lack of data, a comparison of the different NSAID classes was not possible at 26 weeks.

Traditional NSAIDs demonstrated the largest effects with regard to efficacy outcomes, and also demonstrated the least favorable safety profile of all the classes (Table 2). Patients receiving Traditional NSAIDs were significantly more likely to withdraw due to an AE during the study period (median: 6 weeks; range: 2-104). These patients were also 42% more likely than patients receiving placebo to report treatment-related AEs (RR 1.42 [1.13, 1.78]), ~50% more likely to report GI AEs (RR 1.49 [1.31, 1.68]), and 92% more likely to report CV AEs (RR 1.92 [1.17, 3.16]), over a median follow-up time of only 6 weeks.

Patients receiving Celecoxib had a statistically significantly higher risk of experiencing a GI AE (RR 1.14 [1.03, 1.27]) than patients receiving placebo over the course of the study period (median: 12 weeks; range: 1-26) but the effect on CV risk was not significant (RR 1.24 [0.86, 1.80]).

Patients receiving Intermediate COX inhibitors were also significantly more likely than those receiving placebo to report GI AEs (RR 1.40 [1.06, 1.86]) over the course of the treatment period (median: 6 weeks; range: 4-12) but not CV AEs (RR 1.29 [0.63, 2.63]) .

Comparing the Safety Trajectories of Different Classes of NSAIDs

The trajectory of GI AEs differed based on NSAID classification. Patients receiving Traditional NSAIDs experienced the highest likelihood of GI AEs at most time points, and the relative risks were statistically significant at 4 (1.54 [1.23, 1.93]) and 12 weeks (1.52 [1.31, 1.77]). For patients using Intermediate COX inhibitors, the likelihood of developing GI AEs was higher than the placebo group at most time points, but it was statistically significant only at 4 weeks (RR 1.37 [1.02, 1.84]). Patients receiving Celecoxib showed the lowest likelihood of GI AEs among NSAID classes, not reaching statistical significance at any individual time point.

The rates of CV AEs observed in patients receiving Celecoxib and Intermediate COX Inhibitors were not statistically significantly different from their respective placebo groups at any time point. Patients receiving Traditional NSAIDs were statistically significantly more likely to report a CV AE at 4 weeks. There were no significant differences between the Traditional NSAIDs and placebo groups at 2 or 12 weeks. The majority of GI and CV AEs collected at any time point for all NSAID classifications were transient and of minor severity.

DISCUSSION

Our meta-analysis showed that, while NSAIDs demonstrated rapid benefits on pain and functional outcomes, the effects attenuated and lost clinical significance by 8 weeks. Although the magnitude of the effect differed between the three different classes of NSAIDs, the effect consistently waned over time across all classes. Meanwhile, the incidence of minor GI and CV AEs in NSAID users rose as early as 2 weeks after treatment initiation and remained elevated thereafter. The use of Traditional NSAIDs was associated with the least favorable safety profile.

Our study expands upon the oral NSAIDs-related findings from another meta-analysis that was conducted over a decade ago[28]. At the time, the study assembled pain outcomes from 25 RCTs of oral NSAIDs, which included some treatments we considered ineligible for our review (e.g. Valdecoxib and Lumiracoxib) and double-counted the results from one RCT published as two reports[29,30]. Our study included a much larger pool of 72 RCTs and evaluated both pain and functional outcomes in addition to the timeline of safety measurements. Nevertheless, the results of the prior meta-analysis were consistent with ours with regard to pain relief that peaked at around 2 weeks and steadily declined thereafter. Based on our findings coupled with the current literature, clinicians should weigh the

likelihood of a decline in symptomatic benefit against the risk of early-onset minor adverse events, along with the patient perceptions, tolerability and preferences when extending an NSAID treatment regimen beyond 12 weeks[31–33].

The results of our subgroup analyses contingent upon NSAID classification revealed that although all NSAIDs shared a similar trend of efficacy, traditional NSAIDs as a group demonstrated the largest effects on pain and function. This could be explained by the fact that the group of selective COX-2 inhibitors in our review was populated by only one drug – Celecoxib- since other coxibs that might belong to the group did not satisfy our selection criteria (i.e. FDA-approval for use in the U.S.). Three recent network meta-analyses ranking the efficacy of individual NSAIDs of all classes demonstrated very modest clinical effects of Celecoxib, even at its maximum approved daily dose, relative to other coxibs or traditional NSAIDs[34–36]; our findings thus corroborate those results. An important takeaway from the aforementioned network meta-analyses was that the effects of NSAIDs were dose-dependent and varied among individual drugs within and even between NSAID classes. While these meta-analyses provided detailed treatment rankings, a major strength of our study is that we have examined the therapeutic trajectory instead of single time points targeted by network meta-analyses and answered a broader question regarding the expected duration of beneficial effects of NSAIDs. The information from our study and from previous network meta-analyses provides clinicians with a strong background of evidence by which to establish the optimal treatment regimen.

The distinction between NSAIDs based upon COX-2 selectivity was made primarily to examine their safety profile, with the focus on GI and CV AE. In our study, all classes of NSAIDs demonstrated a greater probability of GI AEs. The incidence of minor GI AEs rose with the decline of COX-2 selectivity among NSAID classes, reaching the highest point with the Traditional NSAIDs group. This group (and, to a lesser extent, the Intermediate NSAIDs group) demonstrated statistically significantly more minor GI AEs at 4 weeks; the relative risk of minor GI events remained elevated at 12 weeks in patients taking Traditional NSAIDs. We did not observe the similarity in GI tolerability between the Selective and Intermediate COX inhibitors noted in a 2015 network meta-analysis on NSAID-induced GI injury, possibly owing to the difference in the selection of coxibs under review. Our analyses of CV AE risks indicated that the incidence of minor CV events also rose with the reduction in COX-2 selectivity: they were the lowest for Celecoxib and the highest for the Traditional NSAIDs[37].

Though the follow-up time of our study limited our safety analyses to the observation of minor adverse events, the overall trends we observed in our results align with the findings from a 2004 study by Richy, et al. This study assessed a more heterogeneous population of NSAID users, but demonstrated an early development of GI complications after initiation of treatment with NSAIDs, with timing that varied from one week for Indomethacin to over one month for other NSAIDs[13].

Current Clinical Practice Guidelines recommend use of Traditional NSAIDs with a proton-pump inhibitor (PPI), or use of Celecoxib with or without PPI, to minimize risk of GI toxicity in patients with moderate or high comorbidity risk[5,6]. For patients with CV

comorbidities, Naproxen or Celecoxib have been suggested to minimize the risk of a CV AE[38]. Clinical Practice Guidelines have also indicated that NSAIDs should be used at the lowest effective dose and for the shortest duration [5]. The results of our study support these recommendations, demonstrating the rapidity with which minor negative reactions can occur during the use of NSAID treatment.

The results of our study should be interpreted in light of certain limitations. First, we did not perform separate analyses of the studies utilizing the individual drugs' highest recommended dose because the data would be too scarce to derive meaningful trajectories of their effects on pain and function. Thus, our results may have not pinpointed the absolute measure of potency for the included NSAIDs. Second, the quality of our study was limited by the quality of the underlying data. One of the primary concerns among the included studies was the potential for attrition bias. There was a tendency for attrition rates in both treatment and placebo study arms to be high, but the reasons for discontinuation were unbalanced: a larger share of patients withdrew from the placebo group due to lack of efficacy, whereas more patients from treatment groups withdrew due to AEs. In the context of our results, a higher withdrawal rate due to AEs in the intervention group could skew treatment effects toward the null because patients who discontinued may have been experiencing pain relief or functional improvement despite any adverse experience; conversely, a higher withdrawal rate due to lack of efficacy in the placebo group could inflate the placebo effects because the participants who experienced the least benefit have discontinued the study. In a majority of the studies, withdrawals were tallied at the end of the study and details were not provided for each time point. Therefore, it was not possible for us to quantify the effect of attrition on the treatment effect at specific time points. Overall, due to the attrition imbalance we observed across many of the RCTs, our results may have ultimately understated the overall treatment effects of NSAIDs. Another limitation of our study is the lack of data at and beyond the 26-week time point. Only two studies reported efficacy results at 26 weeks, and Celecoxib was the only treatment represented at this time point. Consequently, the overall estimates for treatment effects at 26 weeks are less precise due to a lower number of participants, and we were unable to conduct subgroup analyses restricting by NSAID classification at this time point. Given that Celecoxib was found to be less effective than Traditional NSAIDs overall, our efficacy estimates for pain and functional outcomes at 26 weeks could be an underestimation of the longer term effects of NSAIDs overall. The scarcity of longer-term follow-up data means that our results may not be generalizable beyond 12 weeks.

Our analyses of safety outcomes were limited by several factors. First, the risk estimates from our study might be smaller than those observed in clinical practice because the knee OA population that was included is more restricted and less representative of general OA population, and because patients with previous GI or cardiovascular issues were most likely excluded from the enrollment. However, our risk estimates are less biased compared with the observational studies because the randomized nature of our data more accurately controls for confounding factors and other biases that limit the interpretation of non-RCT data. Second, we were unable to evaluate the risks for major vascular events (such as myocardial infarction, stroke, or coronary death) or serious GI complications (e.g. GI bleed, perforation, or obstruction) because very few of these events were observed during the study periods in the included pool of studies. Therefore, our safety analyses incorporated more commonly

reported minor events, such as the symptoms of GI upset or edema or hypertension, and the resulting risk ratios may inadequately reflect the risks of major GI and CV AEs. Furthermore, it is possible some of the AEs assembled within the CV AEs group, such as edema and hypertension, were at least partly effected by prostaglandin-mediated effects on renal physiology[39]. Third, the median follow-up time for included studies was 6 weeks, which prevented us from analyzing data on AEs that may arise from extended NSAID use. The shorter follow-up times of the included studies may have also introduced bias with regard to the types of AEs we collected. For example, minor GI events are common and may manifest in numbers sufficient for an analysis in short-term usage periods of an oral treatment, but very few major CV events may be detected within such a brief follow-up period. Finally, in order to maximize the use of available data, we collected composite rates, and in situations where individual events were reported separately (e.g. "diarrhea", "dyspepsia", and "nausea") we summed the number of participants experiencing each event to mimic a composite rate for the organ system. We considered this approach to be justified by the fact that summation of event rates occurred in both treatment and placebo groups; however, the raw event rates may be a slight overestimation of the actual number of patients who experienced GI and/or CV AEs. Despite the above limitations, we detected a statistically significantly heightened risk of minor GI AEs and (in the case of Traditional NSAIDs) minor CV AEs as early as 4 weeks after treatment initiation in the knee OA RCT population. Considering that this estimate is coming from a relatively low risk population, it is reasonable to assume that these values may be a conservative estimate of those observed in the general OA population.

Our results should be interpreted with caution as they focus on the trajectory of response to single regimens in contrast to the dose adjustments and switching that happen in clinical practice[40]. Repeated cycles of continuous NSAID use of longer duration have been suggested both as an alternative to intermittent as-needed use and as a replacement for chronic use[41]. However, the results of our study suggest that such a treatment regimen may lack long-term efficacy, while increasing the risks for adverse treatment effects. Even though repeated NSAID cycles are used by some clinicians, their efficacy trajectory is unknown because long-term clinical trial data on this treatment regimen is lacking. Future research should focus on incorporating study designs that mimic real-world clinical practice in order to better characterize the efficacy trajectory in these scenarios.

In summary, this study described the efficacy and safety trajectories of oral NSAIDs for knee OA over a 26-week period and concluded that oral NSAIDs provide statistically significant pain reduction and functional improvement from as early as 2 weeks and up to 26 weeks of use, with the magnitude of the effect decreasing over this time period and no longer attaining clinical significance after 8 weeks. At the same time, a statistically significant risk of minor GI AEs was evident from 4 weeks of exposure. This information should be taken into account together with patient-specific safety profiles and preferences, comorbid conditions, and concomitant medications to aid clinicians in their decisions on the prescription of oral NSAIDs

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES:

1. Lawrence RC, Felson DT, Helmick CG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum. 2008;58(1):26–35. [PubMed: 18163497]

2. Deshpande BR, Katz JN, Solomon DH, et al. Number of Persons With Symptomatic Knee Osteoarthritis in the US: Impact of Race and Ethnicity, Age, Sex, and Obesity. Arthritis Care Res (Hoboken). 2016;68(12):1743–1750. [PubMed: 27014966]

3. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2163–2196. [PubMed: 23245607]

4. Felson DT, Lawrence RC, Dieppe PA, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med. 2000;133(8):635–646. [PubMed: 11033593]

5. McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage. 2014;22(3):363–388. [PubMed: 24462672]

6. National Clinical Guideline C. National Institute for Health and Clinical Excellence: Guidance In: Osteoarthritis: Care and Management in Adults. London: National Institute for Health and Care Excellence (UK) Copyright (c) National Clinical Guideline Centre, 2014; 2014.

7. Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Use and costs of prescription medications and alternative treatments in patients with osteoarthritis and chronic low back pain in community-based settings. Pain Pract. 2012;12(7):550–560. [PubMed: 22304678]

8. Deveza LA, Hunter DJ, Van Spil WE. Too much opioid, too much harm. Osteoarthritis Cartilage. 2018;26(3):293–295. [PubMed: 29277676]

9. Nissen SE, Yeomans ND, Solomon DH, et al. Cardiovascular Safety of Celecoxib, Naproxen, or Ibuprofen for Arthritis. N Engl J Med. 2016;375(26):2519–2529. [PubMed: 27959716]

10. Schneider V, Levesque LE, Zhang B, Hutchinson T, Brophy JM. Association of selective and conventional nonsteroidal antiinflammatory drugs with acute renal failure: A population-based, nested case-control analysis. Am J Epidemiol. 2006;164(9):881–889. [PubMed: 17005625]

11. Sostres C, Gargallo CJ, Lanas A. Nonsteroidal anti-inflammatory drugs and upper and lower gastrointestinal mucosal damage. Arthritis Res Ther. 2013;15 Suppl 3:S3.

12. Fendrick AM, Greenberg BP. A review of the benefits and risks of nonsteroidal anti-inflammatory drugs in the management of mild-to-moderate osteoarthritis. Osteopath Med Prim Care. 2009;3:1. [PubMed: 19126235]

13. Richy F, Bruyere O, Ethgen O, et al. Time dependent risk of gastrointestinal complications induced by non-steroidal anti-inflammatory drug use: a consensus statement using a meta-analytic approach. Ann Rheum Dis. 2004;63(7):759–766. [PubMed: 15194568]

14. Yang M, Wang HT, Zhao M, et al. Network Meta-Analysis Comparing Relatively Selective COX-2 Inhibitors Versus Coxibs for the Prevention of NSAID-Induced Gastrointestinal Injury. Medicine (Baltimore). 2015;94(40):e1592. [PubMed: 26448006]

15. Hawkey CJ. COX-1 and COX-2 inhibitors. Best Pract Res Clin Gastroenterol. 2001;15(5):801–820. [PubMed: 11566042]

16. Warner TD, Giuliano F, Vojnovic I, Bukasa A, Mitchell JA, Vane JR. Nonsteroid drug selectivities for cyclo-oxygenase-1 rather than cyclo-oxygenase-2 are associated with human gastrointestinal toxicity: a full in vitro analysis. Proc Natl Acad Sci U S A. 1999;96(13):7563–7568. [PubMed: 10377455]

17. Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0 In: Higgins J, Green S, eds. www.handbook.cochrane.org: The Cochrane Collaboration; 2011.

18. Mitchell Markum. Engauge Digitizer. https://github.com/markummitchell/engauge-digitizer.

19. The Cochrane Risk of Bias Tool. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

20. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

21. The Cochrane Musculoskeletal Group. "List of Proposed Outcomes". http://musculoskeletal.cochrane.org/proposed-outcomes.

22. ICH E9 statistical principles for clinical trials. European Medicines Agency;1998.

23. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719–748. [PubMed: 13655060]

24. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–1558. [PubMed: 12111919]

25. Wandel S, Juni P, Tendal B, et al. Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis. Bmj. 2010;341:c4675. [PubMed: 20847017]

26. Ghosh S, Paul S, Das N, Bhattacharyya TK. A study on the effects of diclofenac sodium and etoricoxib in the treatment of osteoarthritis. Journal of the Indian Medical Association. 2007;105(5):260–262. [PubMed: 17915794]

27. Paul S, Das N, Ghosh S. The effects of aceclofenac and nabumetone in osteoarthritis. JNMA J Nepal Med Assoc. 2009;48(174):121–125. [PubMed: 20387351]

28. Bjordal JM, Klovning A, Ljunggren AE, Slordal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: A meta-analysis of randomised placebo-controlled trials. Eur J Pain. 2007;11(2):125–138. [PubMed: 16682240]

29. Bensen WG, Fiechtner JJ, McMillen JI, et al. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. Mayo Clinic proceedings Mayo Clinic. 1999;74(11):1095–1105.

30. Zhao SZ, McMillen JI, Markenson JA, et al. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis treated with celecoxib. Pharmacotherapy. 1999;19(11):1269–1278. [PubMed: 10555933]

31. Huerta C, Castellsague J, Varas-Lorenzo C, Garcia Rodriguez LA. Nonsteroidal anti-inflammatory drugs and risk of ARF in the general population. Am J Kidney Dis. 2005;45(3):531–539. [PubMed: 15754275]

32. Helin-Salmivaara A, Saarelainen S, Gronroos JM, Vesalainen R, Klaukka T, Huupponen R. Risk of upper gastrointestinal events with the use of various NSAIDs: a case-control study in a general population. Scand J Gastroenterol. 2007;42(8):923–932. [PubMed: 17613921]

33. Helin-Salmivaara A, Virtanen A, Vesalainen R, et al. NSAID use and the risk of hospitalization for first myocardial infarction in the general population: a nationwide case-control study from Finland. Eur Heart J. 2006;27(14):1657–1663. [PubMed: 16731535]

34. da Costa BR, Reichenbach S, Keller N, et al. Effectiveness of non-steroidal anti-inflammatory drugs for the treatment of pain in knee and hip osteoarthritis: a network meta-analysis. Lancet. 2017;390(10090):e21–e33. [PubMed: 28699595]

35. van Walsem A, Pandhi S, Nixon RM, Guyot P, Karabis A, Moore RA. Relative benefit-risk comparing diclofenac to other traditional non-steroidal anti-inflammatory drugs and cyclooxygenase-2 inhibitors in patients with osteoarthritis or rheumatoid arthritis: a network meta-analysis. Arthritis Res Ther. 2015;17:66. [PubMed: 25879879]

36. Bannuru RR, Schmid CH, Kent DM, Vaysbrot EE, Wong JB, McAlindon TE. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis: a systematic review and network meta-analysis. Ann Intern Med. 2015;162(1):46–54. [PubMed: 25560713]

37. Conaghan PG. A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity. Rheumatol Int. 2012;32(6):1491–1502. [PubMed: 22193214]

38. Trelle S, Reichenbach S, Wandel S, et al. Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis. BMJ. 2011;342:c7086. [PubMed: 21224324]

39. Horl WH. Nonsteroidal Anti-Inflammatory Drugs and the Kidney. Pharmaceuticals (Basel). 2010;3(7):2291–2321. [PubMed: 27713354]

40. Gore M, Sadosky A, Leslie D, Tai KS, Seleznick M. Patterns of therapy switching, augmentation, and discontinuation after initiation of treatment with select medications in patients with osteoarthritis. Clin Ther. 2011;33(12):1914–1931. [PubMed: 22088416]

41. Bruyere O, Cooper C, Pelletier JP, et al. An algorithm recommendation for the management of knee osteoarthritis in Europe and internationally: a report from a task force of the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum. 2014;44(3):253–263. [PubMed: 24953861]

Significance and Innovations:

- Current research on oral NSAIDs in osteoarthritis does not provide information about the durability of efficacy or the onset of early adverse events.

- We conducted meta-analyses of efficacy and safety at 2, 4, 8, 12, and 26 weeks to characterize the trajectory of efficacy and early adverse events for oral NSAIDs in knee osteoarthritis.

- Our results suggest that the beneficial effects of NSAIDs peak at 2 weeks and begin to decline by 8 weeks, whereas minor GI and CV adverse events begin to manifest as early as 4 weeks.

- Information on the efficacy and safety trajectory of oral NSAIDs can guide clinicians and patients in selecting an appropriate NSAID treatment regimen.
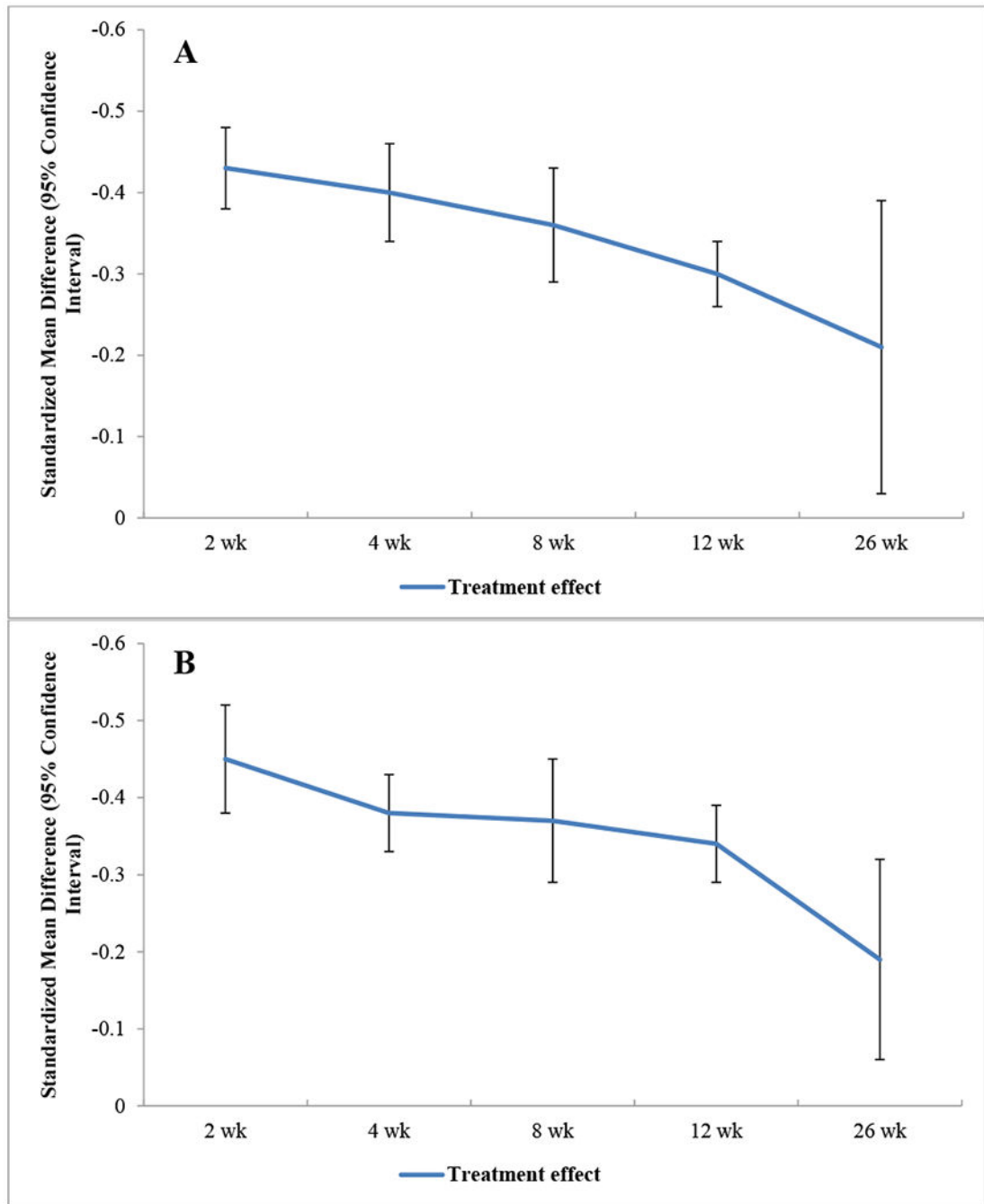
Figure 1:
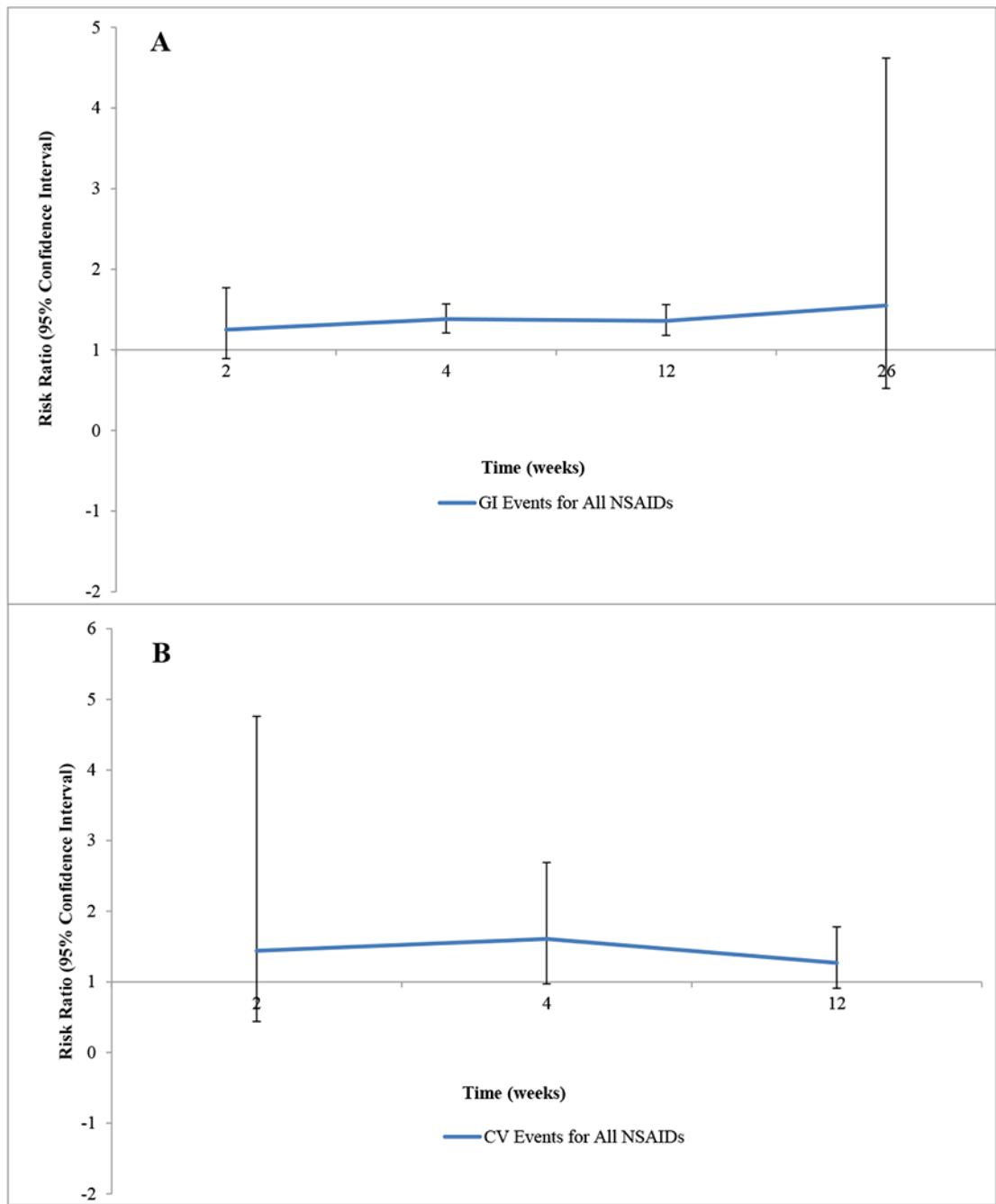Trajectory of Overall Effects of NSAIDs on Pain (A) and Function (B)

Figure 2:

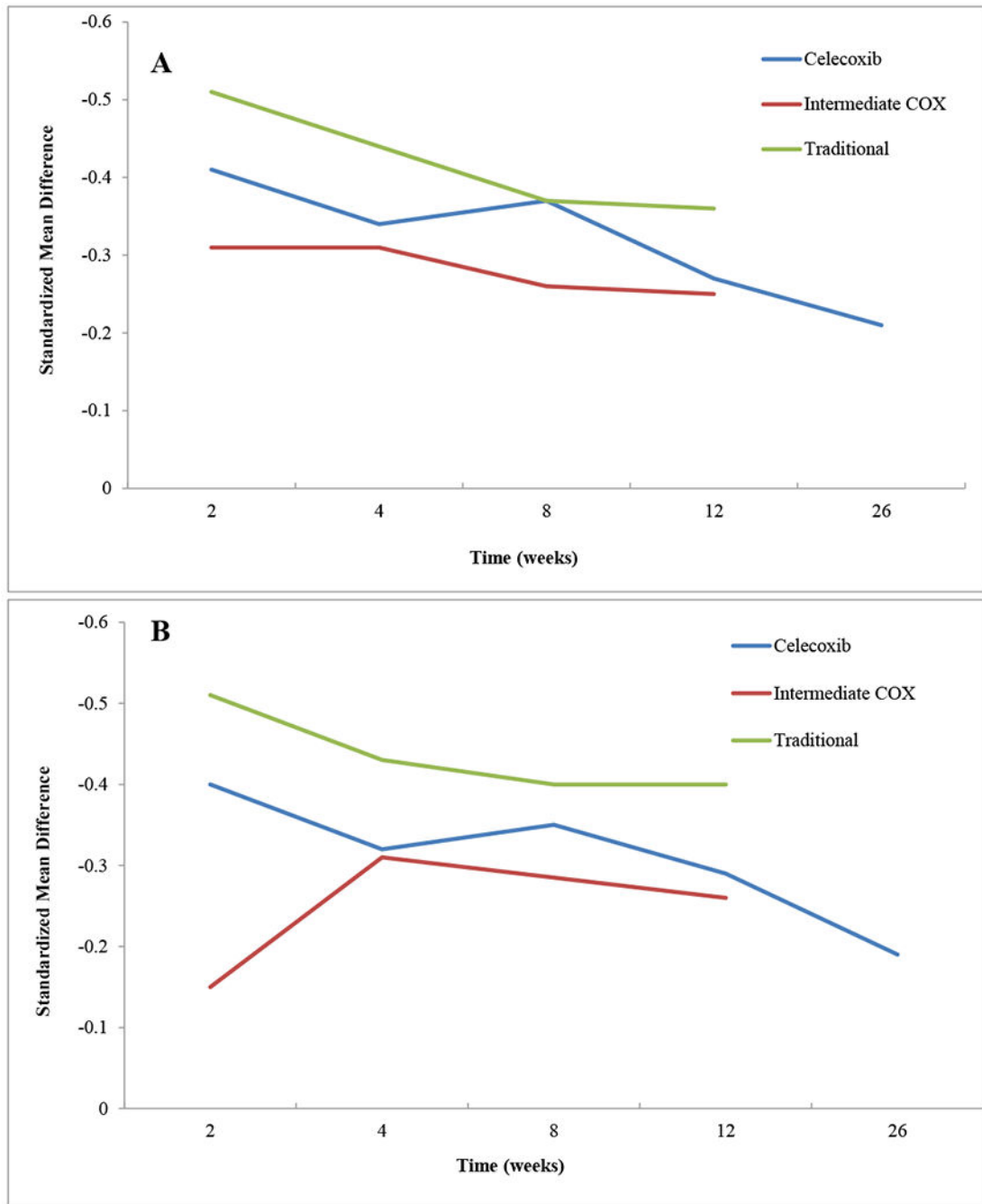Trajectories of the Gastrointestinal (A) and Cardiovascular (B) Safety of NSAIDs

Figure 3:
Trajectory of Effects of Different Classes of NSAIDs on Pain (A) and Function (B)

Table 1:

Effects of NSAIDs on Pain and Function

Outcome	All NSAIDs SMD (95% CI)	All NSAIDs- Very Low Quality* Removed	All NSAIDs- Knee OA only	Celecoxib	Intermediate COX Inhibitors	Traditional NSAIDs
Pain- all time points						
N RCTs; N Patients	48; 17,861	29; 11,741	32; 12,875	20; 7,996	9; 2,902	25; 7,303
Pain- 2 weeks	**-0.43 (-0.48, -0.38)**, $I^2=$ 55%	**-0.48 (-0.55, -0.41)**, $I^2=$ 64%	**-0.40 (-0.46, -0.35)**, $I^2=$ 44%	**-0.41 (-0.48, -0.34)**, $I^2=$ 54%	**-0.31 (-0.41, -0.21)**, $I^2=$ 40%	**-0.51 (-0.59, -0.44)**, $I^2=$ 55%
N RCTs; N Patients	59; 22,911	37; 15,287	41; 16,931	31; 11,699	9; 2,902[†]	29; 8,896
Pain- 4 weeks	**-0.40 (-0.46, -0.34)**, $I^2=$ 76%	**-0.40 (-0.44, -0.36)**, $I^2=$ 21%	**-0.38 (-0.45, -0.30)**, $I^2=$ 80%	**-0.34 (-0.39, -0.29)**, $I^2=$ 32%	**-0.31 (-0.38, -0.23)**, $I^2=$ 0[†]	**-0.44 (-0.54, -0.34)**, $I^2=$ 77%
N RCTs; N Patients	13; 6,341	9; 4,648	6; 3,849	9; 4,970	1; 308	4; 1,218
Pain- 8 weeks	**-0.36 (-0.43, -0.30)**, $I^2=$ 41%	**-0.42 (-0.49, -0.35)**, $I^2=$ 23%	**-0.37 (-0.49, -0.26)**, $I^2=$ 69%	**-0.37 (-0.46, -0.28)**, $I^2=$ 56%	**-0.26 (-0.49, -0.04)**, $I^2=$ NA	**-0.37 (-0.49, -0.25)**, $I^2=$ 0
N RCTs; N Patients	24; 11,096	17; 7,925	15; 7,762	13; 6,472	2; 571	13; 4,657
Pain- 12 weeks	**-0.30 (-0.34, -0.26)**, $I^2=$ 0	**-0.31 (-0.36, -0.27)**, $I^2=$ 0	**-0.27 (-0.32, -0.23)**, $I^2=$ 0	**-0.27 (-0.32, -0.22)**, $I^2=$ 0	**-0.25 (-0.41, -0.08)**, $I^2=$ 0	**-0.36 (-0.42, -0.30)**, $I^2=$ 0
N RCTs; N Patients	2; 976	2; 976	2; 976	2; 976	ND	ND
Pain- 26 weeks	**-0.21 (-0.39, -0.03)**, $I^2=$ 48%	**-0.21 (-0.39, -0.03)**, $I^2=$ 48%	**-0.21 (-0.39, -0.03)**, $I^2=$ 48%	**-0.21 (-0.39, -0.03)**, $I^2=$ 48%	NA	NA
Function- all time points						
N RCTs; N Patients	28; 9,595	21; 7,317	15; 5,619	13; 5,261	1; 263	16; 4,551
Function- 2 weeks	**-0.45 (-0.52, -0.38)**, $I^2=$ 59%	**-0.47 (-0.56, -0.37)**, $I^2=$ 69%	**-0.44 (-0.50, -0.37)**, $I^2=$ 23%	**-0.40 (-0.46, -0.35)**, $I^2=$ 0	-0.15 (-0.40, 0.09), $I^2=$ NA	**-0.51 (-0.63, -0.39)**, $I^2=$ 69%
N RCTs; N Patients	35; 11,979	26; 8,966	22; 8,002	20; 6,219	1; 263	21; 6,282
Function- 4 weeks	**-0.38 (-0.43, -0.33)**, $I^2=$ 45%	**-0.40 (-0.46, -0.33)**, $I^2=$ 49%	**-0.34 (-0.40, -0.28)**, $I^2=$ 31%	**-0.32 (-0.37, -0.26)**, $I^2=$ 19%	**-0.31 (-0.56, -0.07)**, $I^2=$ NA	**-0.43 (-0.52, -0.35)**, $I^2=$ 57%
N RCTs; N Patients	7; 2,492	5; 1,630	1; 460	4; 1,581	ND	3; 911
Function- 8 weeks	**-0.37 (-0.45, -0.29)**, $I^2=$ 0	**-0.41 (-0.52, -0.30)**, $I^2=$ 0	**-0.33 (-0.52, -0.15)**, $I^2=$ NA	**-0.35 (-0.45, -0.25)**, $I^2=$ 0	NA	**-0.40 (-0.61, -0.20)**, $I^2=$ 48%
N RCTs; N Patients	23; 10,118	17; 7,320	14; 6,784	13; 6,395	2; 571	12; 4,165
Function- 12 weeks	**-0.34 (-0.39, -0.29)**, $I^2=$ 28%	**-0.35 (-0.42, -0.28)**, $I^2=$ 44%	**-0.31 (-0.36, -0.25)**, $I^2=$ 17%	**-0.29 (-0.34, -0.24)**, $I^2=$ 0	**-0.26 (-0.43, -0.10)**, $I^2=$ 0	**-0.40 (-0.48, -0.31)**, $I^2=$ 39%

Outcome	All NSAIDs SMD (95% CI)	All NSAIDs- Very Low Quality* Removed	All NSAIDs- Knee OA only	Celecoxib	Intermediate COX Inhibitors	Traditional NSAIDs
N RCTs; N Patients	2; 976	2; 976	2; 976	2; 976	ND	ND
Function- 26 weeks	**−0.19 (−0.32, −0.07)**, I²= 0	**−0.19 (−0.32, −0.07)**, I²= 0	**−0.19 (−0.32, −0.07)**, I²= 0	**−0.19 (−0.32, −0.07)**, I²= 0	NA	NA

*
"Very Low Quality" was defined as studies that received ≥2 High Risk of Bias ratings OR one specific High Risk Rating in the "Other" category in addition to ≥2 Unclear Risk ratings OR ≥3 Unclear Risk of Bias ratings in dimensions other than the "Other" category.

†
Paul 2009 was removed due to I² value of 93% and extremely large effect.

Statistically significant effects are written in bold font. Negative standardized mean differences favor Treatment, and positive standardized mean differences favor Placebo. RCT= Randomized controlled trial; N= "number of…"; CI= Confidence Interval; SMD= Standardized Mean Difference; I²= measure of heterogeneity, with 100% being the maximum possible heterogeneity; ND= No data; NA= Not applicable

Table 2:

Results for Discontinuations & Safety

Outcome	N RCTs	N Patients	Effect estimate (95% CI)*	Range in follow-up (weeks)
All NSAIDs vs. Placebo				
Withdrawals due to Adverse Events	60	22,993	**RR 1.16 (1.02, 1.32)**, I^2= 22%	1-104 (median: 6)
Withdrawals due to Lack of Efficacy	50	18,309	**RR 0.38 (0.34, 0.43)**, I^2= 37%	1-104 (median: 6)
Treatment-Related Adverse Events	24	9,548	**RR 1.21 (1.04, 1.40)**, I^2= 54%	1-13 (median: 6)
Gastrointestinal Adverse Events	59	23,026	**RR 1.36 (1.25, 1.49)**, I^2= 38%	1-26 (median: 6)
Cardiovascular Adverse Events	36	14,654	**RR 1.37 (1.05, 1.77)**, I^2= 0	1-13 (median: 6)
Serious Adverse Events	40	17,278	RR 0.90 (0.68, 1.19), I^2= 0	1-13 (median: 12)
Celecoxib vs. Placebo				
Withdrawals due to Adverse Events	28	11,177	RR 1.01 (0.84, 1.22), I^2= 23%	1-26 (median: 12)
Withdrawals due to Lack of Efficacy	23	9,084	**RR 0.40 (0.34, 0.48)**, I^2= 31%	1-26 (median: 12)
Treatment-Related Adverse Events	13	4,722	RR 0.99 (0.86, 1.13), I^2= 7%	1-13 (median: 6)
Gastrointestinal Adverse Events	27	10,984	**RR 1.14 (1.03, 1.27)**, I^2= 0	1-26 (median: 12)
Cardiovascular Adverse Events	18	7,732	RR 1.24 (0.86, 1.80), I^2= 0	1-13 (median: 12)
Serious Adverse Events	23	9,723	RR 0.89 (0.60, 1.32), I^2= 0	1-13 (median: 12)
Intermediate COX Inhibitors vs. Placebo				
Withdrawals due to Adverse Events	11	3,419	RR 1.11 (0.78, 1.57), I^2= 29%	4-12 (median: 6)
Withdrawals due to Lack of Efficacy	9	2,906	**RR 0.49 (0.39, 0.62)**, I^2= 25%	4-12 (median: 6)
Treatment-Related Adverse Events	3	1,045	RR 1.05 (0.82, 1.33), I^2= 0	4-6 (median: 4)
Gastrointestinal Adverse Events	11	3,419	**RR 1.40 (1.06, 1.86)**, I^2= 59%	4-12 (median: 6)
Cardiovascular Adverse Events	5	2,029	RR 1.29 (0.63, 2.63), I^2= 0	4-12 (median: 6)
Serious Adverse Events	5	1,829	RR 1.98 (0.57, 6.93), I^2= 0	4-12 (median: 6)
Traditional NSAIDs vs. Placebo				
Withdrawals due to Adverse Events	33	10,302	**RR 1.36 (1.16, 1.59)**, I^2= 1%	2-104 (median: 6)
Withdrawals due to Lack of Efficacy	25	7,066	**RR 0.31 (0.25, 0.39)**, I^2= 43%	2-104 (median: 6)
Treatment-Related Adverse Events	13	4,263	**RR 1.42 (1.13, 1.78)**, I^2= 61%	4-13 (median: 6)

Outcome	N RCTs	N Patients	Effect estimate (95% CI)*	Range in follow-up (weeks)
Gastrointestinal Adverse Events	32	9,892	**RR 1.49 (1.31, 1.68)**, I²= 45%	1-13 (median: 6)
Cardiovascular Adverse Events	17	5,542	**RR 1.92 (1.17, 3.16)**, I²= 29%	1-13 (median: 6)
Serious Adverse Events	20	6,573	RR 0.93 (0.63, 1.38), I²= 0	1-13 (median: 6)

*
Statistically significant effects appear in bold font. Risk ratios less than one favor Treatment, and risk ratios greater than one favor Placebo.

N= "number of…"; RCT= Randomized controlled trial; CI= Confidence Interval; RR= Risk Ratio; I²= measure of heterogeneity, with 100% being the maximum possible heterogeneity