

Responder analysis for pain relief and numbers needed to treat in a meta-analysis of etoricoxib osteoarthritis trials: bridging a gap between clinical trials and clinical practice

R A Moore,¹ O A Moore,² S Derry,¹ P M Peloso,³ A R Gammaitoni,³ H Wang³

► Additional data are published online only at <http://ard.bmj.com/content/vol69/issue2>

¹Pain Research and Nuffield Department of Anaesthetics, University of Oxford, Oxford, UK; ²Department of Rheumatology, Musgrove Park Hospital, Belfast, UK; ³Merck Research Laboratories, Rahway, New Jersey, USA

Correspondence to:

Dr R A Moore, Pain Research and Nuffield Department of Anaesthetics, University of Oxford, John Radcliffe Hospital, Level 6 West Wing, Oxford OX3 9DU, UK; andrew.moore@pru.ox.ac.uk

Accepted 22 March 2009
Published Online First
12 April 2009

ABSTRACT

Background: Population mean changes from clinical trials are difficult to apply to individuals in clinical practice. Responder analysis may be better, but needs validating for level of response and treatment duration.

Methods: The numbers of patients with pain relief over baseline ($\geq 15\%$, $\geq 30\%$, $\geq 50\%$, $\geq 70\%$) at 2, 4, 8 and 12 weeks of treatment were obtained using the WOMAC 100 mm visual analogue pain subscale score for each treatment group in seven randomised placebo-controlled trials of etoricoxib in osteoarthritis lasting ≥ 6 weeks. Dropouts were assigned 0% improvement from baseline from then on. The numbers needed to treat (NNTs) were calculated at each level of response and time point.

Results: 3554 patients were treated with placebo, etoricoxib 30 mg and 60 mg, celecoxib 200 mg, naproxen 1000 mg or ibuprofen 2400 mg daily. Response rates fell with increasing pain relief: 60–80% experienced minimally important pain relief ($\geq 15\%$), 50–60% moderate pain relief ($\geq 30\%$), 40–50% substantial pain relief ($\geq 50\%$) and 20–30% extensive pain relief ($\geq 70\%$). NNTs for etoricoxib, celecoxib and naproxen were stable over 2–12 weeks. Ibuprofen showed lessening of effectiveness with time.

Conclusion: Responder rates and NNTs are reproducible for different levels of response over 12 weeks and have relevance for clinical practice at the individual patient level. An average 10 mm improvement in pain equates to almost one in two patients having substantial benefit.

Clinical trials are performed usually for regulatory purposes, with outcomes typically reported as statistical comparisons between treatment group population means. The results of clinical trials can be difficult to translate into clinical practice. A report that an intervention shows an average 10 mm reduction more than placebo on a 100 mm visual analogue scale has little immediate impact.

Moreover, few of us are average. Most drugs provide a good response in half or fewer of the patients treated,^{1,2} true in postoperative pain,³ neuropathic pain,^{4–6} migraine⁷ and tumour necrosis factor antagonists in rheumatoid arthritis.⁸ An 80/20 rule seems to apply in osteoarthritis, with 80% of patients experiencing 20% pain relief but only 20% experiencing 80% relief; about half have their pain halved.⁹

Genetic influences help determine the clinical response to analgesic drugs for non-specific anti-inflammatory drugs (NSAIDs),¹⁰ opioids¹¹ and

more generally,¹² as well as the clinical response to methotrexate.¹³ Pain is driven by complex pathways of neural mechanisms which are likely to be different between individuals.¹⁴ Imaging reveals loss of grey matter in chronic pain above that found with age alone.^{15,16}

Average data from skewed distributions can produce misleading results.¹⁷ Dichotomous responder analyses have been reported previously for acute¹⁸ and chronic pain.^{5,6,19} The validity of a dichotomous measure should be established before being widely used.²⁰

An added factor contributing to differences in treatment response observed in clinical practice compared with a clinical trial is the handling of dropouts. Commonly, a “last observation carried forward” technique is used in clinical trials, where data from patients with good pain control but intolerable adverse events will still be included in efficacy calculations using the population mean. In clinical practice, this same patient would be considered a treatment failure.

We used individual patient data from seven randomised placebo-controlled trials in osteoarthritis to investigate the effects of different levels of pain relief assessed at various time points on estimates of efficacy.

METHODS

Merck Research Laboratories provided pain response data from seven randomised placebo-controlled trials of etoricoxib in osteoarthritis lasting ≥ 6 weeks (protocols 007, 018, 019, 071, 073, 076 and 077).^{21–26} PDF copies of the company clinical trial reports were also available.

We calculated the number of patients in each treatment group in each trial achieving various Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) thresholds of pain relief over baseline of $\geq 15\%$ (minimal benefit), $\geq 30\%$ (moderate), $\geq 50\%$ (substantial)²⁷ and $\geq 70\%$ which we defined as extensive improvement. These were assessed at 2, 4, 8 and 12 weeks. All trials lasted 12 weeks except protocol 007 which lasted 6 weeks.

In each study patients were asked, “During the last 48 hours, how much pain do you have (1) walking on a flat surface; (2) going up or down stairs; (3) at night while in bed; (4) sitting or lying; (5) standing upright?”. On a 100 mm visual analogue scale, patients placed an “x” ranging from 0 (“no pain”) to 100 (“extreme pain”). The



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

Western Ontario and McMaster Universities (WOMAC) 100 mm visual analogue pain subscale score was calculated as the average of the responses to the five questions.

Criteria used in defining responders included:

- ▶ For patients who did not drop out, only actual measured values were used for calculations. Last observation carried forward was not used.
- ▶ For patients who withdrew for any reason, measurements made within 7 days of the last dose were used to calculate the response.
- ▶ Thereafter, patients were assigned 0% improvement.

We calculated the number and percentage of responders for each level of response for each drug and time point and the number needed to treat (NNT) compared with placebo (with 95% CI).²⁸ The relative risk with 95% CI was calculated using the fixed effects model²⁹ and considered statistically significant when the 95% CI did not include 1. Statistically significant differences between NNTs were established using the z test,³⁰ comparing different drug/dose combinations only in the trials in which they were used together.

Clinical trial reports were used to obtain, for each active treatment, the difference between active treatment and placebo for the WOMAC pain subscale score. This was defined in the company clinical trial reports as the mean time-weighted average change from baseline (flare/randomisation visit) over the 6- or 12-week treatment period ((WOMAC baseline – WOMAC treatment average) – (WOMAC baseline – WOMAC placebo average)). Average results for each treatment arm were pooled using RevMan 5.0.

RESULTS

Information was available on 3554 patients (two-thirds women) with an average age of 62 years (see online supplement). Six trials involved patients with osteoarthritis of the knee or hip and one of the knee only. Osteoarthritis was established clinically and radiographically. Initial pain had to be a minimum of 40/100 mm at inclusion, plus ≥ 15 mm increase and worsening in investigator global assessment since baseline. The actual numbers of responders for each level of response for each drug in each trial and at each time point are shown in the online supplement.

Table 1 gives the percentage of responders and NNTs for each level of response for each drug in each trial and at each time point. The percentage of responders and NNTs are also shown in figs 1 and 2, respectively.

The percentage of patients achieving levels of pain relief with placebo at each threshold rose between weeks 2 and 12 (table 1). At the end of 12 weeks the proportion of patients was about 45% for a response of $\geq 15\%$, 35% for $\geq 30\%$, 25% for $\geq 50\%$ and 15% for $\geq 70\%$. More patients achieved each level of response with active drug than with placebo (table 1, fig 1). Etoricoxib 30 mg and 60 mg daily and celecoxib 200 mg daily had similar response patterns, with constant percentages at lower response levels but a tendency over time for the proportion achieving $\geq 70\%$ pain relief by 12 weeks to increase. Naproxen 1000 mg daily and ibuprofen 2400 mg daily had different response patterns, with stable or increasing percentages at higher levels ($\geq 50\%$, $\geq 70\%$) but progressive decreases in the percentage achieving lower levels of response of $\geq 15\%$ and $\geq 30\%$.

Table 1 Percentage of responders with treatment/placebo and numbers needed to treat (NNTs) after 2, 4, 8 and 12 weeks

Outcome	Percentage of responders with treatment/placebo after different numbers of weeks				NNT (95% CI)			
	2	4	8	12	2	4	8	12
Etoricoxib 30 mg; 5 trials, 1486 patients								
$\geq 15\%$	70/41	70/45	70/43	68/45	3.5 (2.9 to 4.2)	4.1 (3.4 to 5.2)	3.6 (3.0 to 4.4)	4.4 (3.5 to 5.7)
$\geq 30\%$	54/27	56/30	59/31	59/36	3.7 (3.1 to 4.5)	3.9 (3.3 to 4.9)	3.7 (3.1 to 4.5)	4.3 (3.5 to 5.6)
$\geq 50\%$	35/13	40/18	44/23	44/24	4.6 (3.8 to 5.6)	4.4 (3.7 to 5.5)	5.0 (4.0 to 6.5)	5.1 (4.1 to 6.9)
$\geq 70\%$	17/5	22/9	26/13	28/17	8.9 (7.0 to 12)	7.7 (6.0 to 11)	7.8 (6.0 to 11)	8.9 (6.3 to 15)
Etoricoxib 60 mg; 3 trials, 711 patients								
$\geq 15\%$	77/46	76/51	73/44	71/47	3.3 (2.6 to 4.5)	4.1 (3.1 to 6.2)	3.5 (2.7 to 4.9)	4.2 (2.9 to 7.4)
$\geq 30\%$	61/31	63/32	60/30	59/35	3.2 (2.6 to 4.6)	3.2 (2.6 to 4.4)	3.3 (2.6 to 4.5)	4.0 (2.9 to 6.7)
$\geq 50\%$	43/16	46/19	47/20	44/23	3.7 (2.9 to 5.0)	3.7 (2.9 to 5.1)	3.7 (2.9 to 5.1)	4.7 (3.3 to 8.1)
$\geq 70\%$	25/4	27/9	30/11	30/15	4.7 (3.9 to 6.0)	5.7 (4.3 to 8.4)	5.3 (4.0 to 7.7)	6.7 (4.4 to 14)
Celecoxib 200 mg; 2 trials, 714 patients								
$\geq 15\%$	64/42	67/45	64/40	64/40	4.6 (3.4 to 7.0)	4.5 (3.3 to 6.7)	4.2 (3.2 to 6.1)	4.2 (3.2 to 6.1)
$\geq 30\%$	53/27	54/29	54/30	53/31	3.9 (3.0 to 5.3)	4.0 (3.1 to 5.6)	4.1 (3.2 to 5.9)	4.7 (3.5 to 7.1)
$\geq 50\%$	30/12	39/18	39/22	39/22	5.6 (4.2 to 8.3)	4.8 (3.6 to 6.9)	5.9 (4.2 to 9.9)	5.8 (4.2 to 9.5)
$\geq 70\%$	15/6	22/10	24/13	26/16	10 (7.1 to 18)	8.7 (5.9 to 16)	9.1 (6.0 to 19)	10 (6.3 to 27)
Naproxen 1000 mg; 2 trials, 531 patients								
$\geq 15\%$	78/54	75/57	70/48	64/47	4.2 (3.0 to 7.5)	5.6 (3.6 to 13)	4.6 (3.1 to 8.8)	5.9 (3.6 to 15)
$\geq 30\%$	63/35	65/36	60/35	55/35	3.5 (2.6 to 5.5)	3.5 (2.6 to 5.4)	3.9 (2.8 to 6.4)	4.8 (3.2 to 9.2)
$\geq 50\%$	44/20	47/22	45/22	44/23	4.1 (3.0 to 6.3)	3.9 (2.9 to 6.0)	4.3 (3.1 to 7.0)	4.8 (3.3 to 8.5)
$\geq 70\%$	22/5	27/10	28/12	27/15	5.9 (4.4 to 8.8)	5.8 (4.1 to 9.7)	6.1 (4.2 to 11)	8.0 (4.9 to 21)
Ibuprofen 2400 mg; 2 trials, 618 patients								
$\geq 15\%$	69/43	65/47	60/48	58/49	3.8 (2.9 to 5.5)	5.5 (3.8 to 10)	7.8 (4.8 to 22)	11 (5.8 to 121)
$\geq 30\%$	49/29	51/33	51/36	49/41	5.0 (3.6 to 8.1)	5.6 (3.9 to 10)	6.6 (4.3 to 14)	NS
$\geq 50\%$	29/14	35/20	40/24	39/27	6.5 (4.6 to 11)	6.9 (4.6 to 14)	6.4 (4.3 to 12)	8.4 (5.1 to 24)
$\geq 70\%$	16/6	20/8	22/13	26/18	10 (7.0 to 21)	7.9 (5.6 to 14)	11 (6.6 to 32)	13 (6.8 to 75)

Response with placebo is for placebo groups from trials in the particular comparison being made.

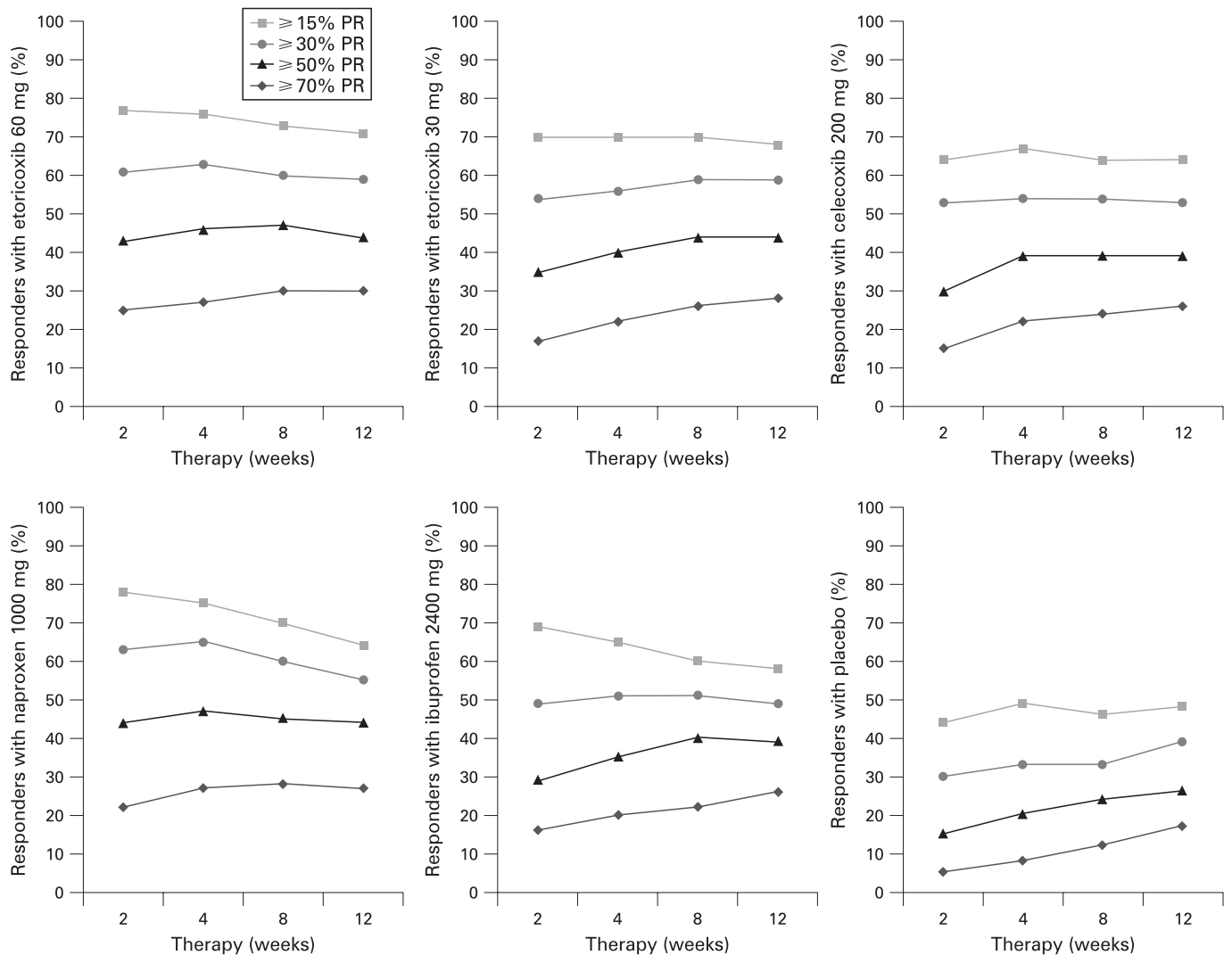


Figure 1 Percentage of responders over baseline at various levels of reductions in pain intensity (PR) for placebo and five active drugs over 12 weeks of treatment.

NNTs calculated for $\geq 15\%$, $\geq 30\%$ and $\geq 50\%$ pain relief were very similar with etoricoxib 60 mg and 30 mg, celecoxib 200 mg and naproxen 1000 mg (table 1, fig 2). NNT values were between 3 and 5 over the 12 weeks of measurement. For $\geq 70\%$ pain relief the NNT was distinctly higher with values between 6 and 10. The pattern for ibuprofen 2400 mg was different, with NNTs generally much higher (worse) at longer study duration and less consistency between the various levels of response.

There were three direct comparisons, each in two trials, of daily doses of etoricoxib 60 mg and naproxen 1000 mg, etoricoxib 30 mg and celecoxib 200 mg, and etoricoxib 30 mg and ibuprofen 2400 mg. No statistically significant differences were found between them at any level of response or any duration of treatment.

The additional reduction in the WOMAC pain subscale score for each active treatment over placebo between the flare/randomisation visit and end of treatment is shown in fig 3. The smallest mean difference above placebo was 8 mm for ibuprofen 2400 mg daily and the largest was 15 mm for etoricoxib 60 mg daily. This shows that about 60% of patients have moderate benefit (fig 1) while average reductions in pain over placebo appear modest.

DISCUSSION

Population mean changes have no easy resonance outside a clinical trial. An average of 10 mm out of 100 mm (10% improvement; fig 3) conveys no expectation of great benefit, with little to balance against known risks. A different approach is needed.

A pilot study of responder analysis in a single trial indicated that it might be a useful way of reporting pain results in osteoarthritis.⁹ It suggested that at least 50% pain relief after 6 weeks of treatment could be a useful discriminator between interventions of greater and lesser efficacy. This needs validation with regard to different levels of pain relief, and especially duration, given that arthritis treatments are used in the medium to long term.

IMMPACT provided recommendations for interpreting the clinical importance of treatment outcomes in clinical trials of chronic pain.²⁷ It suggested that a 10–20% decrease in pain intensity was minimally important, $\geq 30\%$ moderately important and $\geq 50\%$ substantial. This meta-analysis used these three discriminator points, together with the even higher discriminator of $\geq 70\%$ pain relief, to perform a responder analysis, calculate NNTs and examine the effects of trial duration.

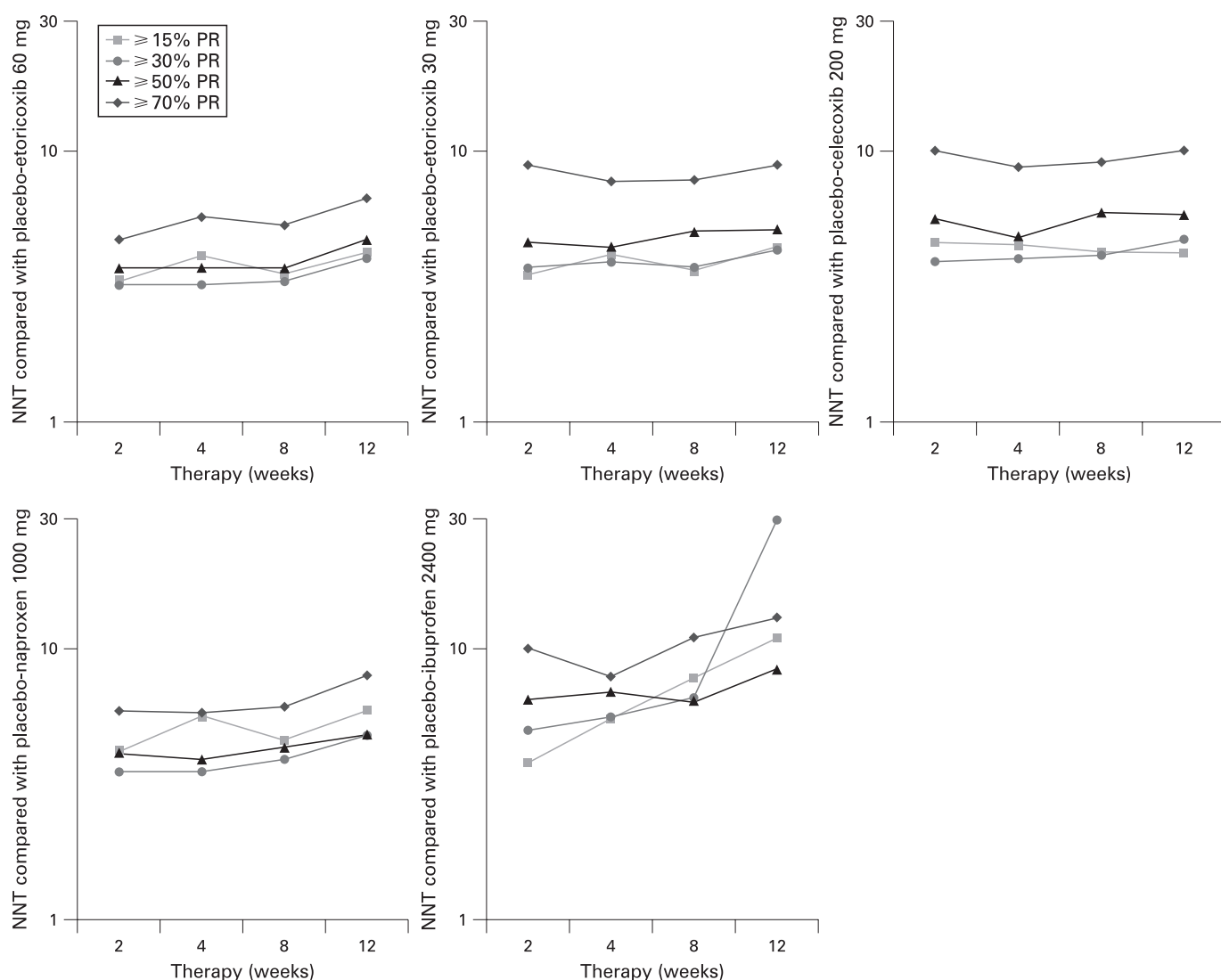


Figure 2 Numbers needed to treat (NNT) compared with placebo for five active drugs over 12 weeks of treatment using various levels of reductions in pain intensity (PR) over baseline.

Response rates declined as the discriminator level increased for all five active drugs and placebo (fig 1, table 1). Using the IMMPACT descriptions for commonly used NSAIDs, 60–80% of patients with osteoarthritis can expect minimally important pain relief, 50–60% moderate pain relief, 40–50% substantial pain relief and 20–30% extensive pain relief.

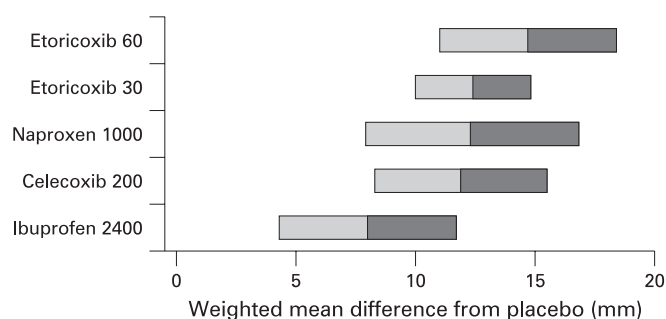


Figure 3 Weighted mean difference between treatment and flare/randomisation visit for WOMAC pain subscale: active treatment minus placebo. Shading shows upper and lower 95% confidence intervals.

With placebo and active drugs the proportion achieving higher levels of pain relief increased over time, perhaps due to a natural waning of pain inherent in the “flare” design. The tendency to less response at lower levels of pain relief over time with naproxen and ibuprofen may reflect higher withdrawal rates for traditional NSAIDs over cyclooxygenase-2 selective inhibitors.^{31 32} Patients may be balancing benefit and harm, with lower levels of relief perhaps not worthwhile in the face of adverse events. In the responder approach, dropouts contribute to the denominator only, whereas with “last observation carried forward” they appear to continue to benefit after withdrawal.

NNTs were comparable for pain relief of $\geq 15\%$, $\geq 30\%$ and $\geq 50\%$, with higher (worse) NNTs for $\geq 70\%$ for both doses of etoricoxib, celecoxib and naproxen. For these three drugs, NNTs were reasonably stable over 2–12 weeks. Ibuprofen 2400 mg daily was different, with NNT values generally increasing (worsening) with longer duration. The longitudinal responder analysis provides more insight than population average differences (fig 3). This apparently different behaviour with ibuprofen 2400 mg daily did not translate into a statistically significant difference in NNTs between ibuprofen and etoricoxib 30 mg, nor was a significant difference found between

etoricoxib 30 mg and celecoxib 200 mg, or etoricoxib 60 mg and naproxen 100 mg. Establishing a dose-response in analgesic trials is known to be difficult even where direct comparisons are available in relatively simple models such as postoperative pain.³³ Demonstrating an absolute difference in response of 10% requires a substantial number of trials and patients. Confirming statistically that ibuprofen 2400 mg is inferior to other NSAIDs at commonly used doses in osteoarthritis will require more data.

A range of ± 0.5 NNT was proposed to determine whether an NNT has “clinical relevance”—whether the NNT is within acceptable bounds of clinical “accuracy”.³⁴ A subsequent proposal was that ± 0.5 NNT could be used to determine that NNTs were different.²⁷ If a numerical difference between NNTs of 1 (eg, 3.5 vs 5.0) was taken to be important, then the application to NNTs in table 1 would begin to differentiate between drugs, with ibuprofen particularly being judged less effective.

The results of the responder analysis achieve the same global conclusions as the original trials—namely, that etoricoxib and its comparators have useful analgesic properties in osteoarthritis. Arguably the most important outcome from these analyses is that, for osteoarthritis, patients and professionals can be provided with trial data that translate into clinical practice by using realistic estimates of the chance of achieving a particular level of benefit. For pain, a mean difference of 10 mm over placebo translates into about 40% having substantial benefit and 30% not having even minimal benefit. Most people with osteoarthritis treated with an NSAID at an appropriate dose can expect to get at least a minimal benefit (though 1 in 5 will not), almost 1 in 2 can expect a substantial benefit and about 1 in 5 an extensive benefit. The prospect of a 1 in 2 chance of substantial benefit has considerably more impact than an average 10% improvement in pain. Moreover, the information is conveyed in terms of both the likelihood of benefit (1 in 2) and the extent of the benefit (substantial).

Patients and professionals are interested in the known associated risks of both common and rare harm. These, too, can be conveyed in terms of likelihood of harm and its consequences.³⁵ Balancing benefit and harm is easier when common language describes both. Providing information about the chance of response at various thresholds might produce a more realistic appreciation of benefits and risks of treatment. There is, of course, the caveat that these results come from flare designs in clinical trials where patient selection criteria may make the population different (less comorbidity, perhaps) than a clinical practice population.

Responders are defined not just by the level of response but by the outcome used to define response. We chose the WOMAC pain subscale, combining pain associated with walking, climbing stairs, sitting, lying down and at night while in bed. Whether pain is the most appropriate outcome for responder analysis or whether function, sleep, quality of life or compound outcomes like the OMERACT-OARSI set of responder criteria³⁶ are preferable remains to be examined. Finally, this type of analysis may differentiate between treatments in a manner that is not possible with population mean changes in pain intensity, and discriminatory power may reside in different outcomes.

Responder analysis looks promising and much more helpful than an average change of a few millimetres based on populations of responders combined with non-responders.

CONCLUSION

Population mean change in pain intensity reported in clinical trials may be difficult to translate into clinical decision-making and patient expectations of benefit. Responder analyses and

NNTs calculated from them are reproducible for different levels of response and over at least 12 weeks of treatment with effective drugs. This offers the possibility of providing patients and professionals with information on the chance of achieving particular degrees of pain relief, improving clinical decision-making and patient communication.

Acknowledgements: The authors are grateful to Merck Research Laboratories for making available data from the original trials.

Funding: Pain Research is supported in part by the Oxford Pain Research Trust. Neither the Trust nor Merck Research Laboratories had any role in the design, planning, execution of the study or in writing the manuscript. No financial support was received from Merck Research Laboratories for this work. RAM is funded by NIHR Biomedical Research Centre Programme.

Competing interests: RAM has received research grants, consulting or lecture fees from pharmaceutical companies including Pfizer, MSD, GSK, AstraZeneca, Grunenthal, Menarini, Futura and others. RAM and SD have also received research support from charities and government sources at various times. RAM is the guarantor. RAM, OAM and SD have no direct stock holding in any pharmaceutical company. PMP, ARG and HW are employees of Merck Inc.

Contributions: RAM, ARG and PMP were involved with the original concept, planning the study, searching, writing it, analysis and preparing the manuscript; HW provided the data; OAM and RAM performed calculations and analysis; OAM and SD were involved with planning and writing. All authors read and approved the final manuscript.

Provenance and peer review: Not commissioned; externally peer reviewed.

REFERENCES

1. Christakis NA. Does this work for you? *BMJ* 2008;**337**:a2281.
2. Moore RA, Straube S, Derry S, et al. Does this work for you? Individuals, averages, and evidence based medicine. *BMJ* 2008;**337**:a2585
3. Moore RA, Edwards JE, McQuay HJ. Acute pain: individual patient meta-analysis shows the impact of different ways of analysing and presenting results. *Pain* 2005;**116**:322–31.
4. Finnerup NB, Otto M, McQuay HJ, et al. Algorithm for neuropathic pain treatment: an evidence based proposal. *Pain* 2005;**118**:289–305.
5. Straube S, Derry S, McQuay HJ, et al. Enriched enrolment: definition and effects of enrichment and dose in trials of pregabalin and gabapentin in neuropathic pain. A systematic review. *Br J Clin Pharmacol* 2008;**66**:266–75.
6. Sultan A, Gaskell H, Derry S, et al. Duloxetine for painful diabetic neuropathy and fibromyalgia pain: systematic review of randomised trials. *BMC Neurol* 2008;**8**:29.
7. Dahlof CG, Pascual J, Dodick DW, et al. Efficacy, speed of action and tolerability of almotriptan in the acute treatment of migraine: pooled individual patient data from four randomized, double-blind, placebo-controlled clinical trials. *Cephalalgia* 2006;**26**:400–8.
8. Geborek P, Crnkic M, Petersson IF, et al. South Swedish Arthritis Treatment Group. Etanercept, infliximab, and leflunomide in established rheumatoid arthritis: clinical experience using a structured follow up programme in southern Sweden. *Ann Rheum Dis* 2002;**61**:793–8.
9. Moore RA, Moore OA, Derry S, et al. Numbers needed to treat calculated from responder rates give a better indication of efficacy in osteoarthritis trials than mean pain scores. *Arthritis Res Ther* 2008;**10**:R39.
10. Fries S, Grosser T, Price TS, et al. Marked interindividual variability in the response to selective inhibitors of cyclooxygenase-2. *Gastroenterology* 2006;**130**:55–64.
11. Klepstad P, Dale O, Skorpen F, et al. Genetic variability and clinical efficacy of morphine. *Acta Anaesthesiol Scand* 2005;**49**:902–8.
12. Lötsch J, Geisslinger G. Current evidence for a genetic modulation of the response to analgesics. *Pain* 2006;**121**:1–5.
13. Wessels JA, van der Kooij SM, le Cessie S, et al. Pharmacogenetics Collaborative Research Group. A clinical pharmacogenetic model to predict the efficacy of methotrexate monotherapy in recent-onset rheumatoid arthritis. *Arthritis Rheum* 2007;**56**:1765–75.
14. Harvey VL, Dickenson AH. Mechanisms of pain in nonmalignant disease. *Curr Opin Support Palliat Care* 2008;**2**:133–9.
15. Baliki MN, Geha PY, Apkarian AV, et al. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. *J Neurosci* 2008;**28**:1398–403.
16. Tracey I. Neuroimaging of pain mechanisms. *Curr Opin Support Palliat Care* 2007;**1**:109–16.
17. McQuay HJ, Carroll D, Moore RA. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain* 1996;**64**:331–5.
18. Moore RA, McQuay H. Single-patient data meta-analysis of 3,453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.
19. Farrar JT, Dworkin RH, Max MB. Use of the cumulative proportion of responders analysis graph to present pain data over a range of cut-off points: making clinical trial data more understandable. *J Pain Symptom Manage* 2006;**31**:369–77.
20. McQuay HJ, Moore RA. Using numerical results from systematic reviews in clinical practice. *Ann Intern Med* 1997;**126**:712–20.

21. **Gottesdiener K**, Schnitzer T, Fisher C, *et al*, Protocol 007 Study Group. Results of a randomized, dose-ranging trial of etoricoxib in patients with osteoarthritis. *Rheumatology (Oxford)* 2002;**41**:1052–61.
22. **Reginster JY**, Malmstrom K, Mehta A, *et al*. Evaluation of the efficacy and safety of etoricoxib compared to naproxen in two, 138-week randomized studies of osteoarthritis patients. *Ann Rheum Dis* 2007;**66**:945–51.
23. **Leung AT**, Malmstrom K, Gallacher AE, *et al*. Efficacy and tolerability profile of etoricoxib in patients with osteoarthritis: a randomized, double-blind, placebo and active-comparator controlled 12-week efficacy trial. *Curr Med Res Opin* 2002;**18**:49–58.
24. **Wiesenhutter CW**, Boice JA, Ko A, *et al*, Protocol 071 Study Group. Evaluation of the comparative efficacy of etoricoxib and ibuprofen for treatment of patients with osteoarthritis: a randomized, double-blind, placebo-controlled trial. *Mayo Clin Proc* 2005;**80**:470–9.
25. **Puopolo A**, Boice JA, Fidelholtz JL, *et al*. A randomized placebo-controlled trial comparing the efficacy of etoricoxib 30 mg and ibuprofen 2400 mg for the treatment of patients with osteoarthritis. *Osteoarthr Cartil* 2007;**15**:1348–56.
26. **Bingham CO**, Sebba AI, Rubin BR, *et al*. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies. *Rheumatology (Oxford)* 2007;**46**:496–507.
27. **Dworkin RH**, Turk DC, Wyrwich KW, *et al*. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;**9**:105–21.
28. **Cook RJ**, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
29. **Morris JA**, Gardner MJ. Calculating confidence intervals for relative risk, odds ratios and standardised ratios and rates. In: Gardner MJ, Altman DG, eds. *Statistics with confidence – confidence intervals and statistical guidelines*. London: BMJ, 1995:50–63.
30. **Tramer MR**, Reynolds DJ, Moore RA, *et al*. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997;**315**:635–40.
31. **Moore RA**, Derry S, Makinson GT, *et al*. Tolerability and adverse events in clinical trials of celecoxib in osteoarthritis and rheumatoid arthritis: systematic review and meta-analysis of information from company clinical trial reports. *Arthritis Res Ther* 2005;**7**:R644–65.
32. **Moore RA**, Derry S, McQuay HJ. Discontinuation rates in clinical trials in musculoskeletal pain: meta-analysis from etoricoxib clinical trial reports. *Arthritis Res Ther* 2008;**10**:R53.
33. **McQuay HJ**, Moore RA. Dose-response in direct comparisons of different doses of aspirin, ibuprofen and paracetamol (acetaminophen) in analgesic studies. *Br J Clin Pharmacol* 2007;**63**:271–8.
34. **Moore RA**, Gavaghan D, Tramèr MR, *et al*. Size is everything—large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. *Pain* 1998;**78**:209–16.
35. **Moore RA**, Derry S, McQuay HJ, *et al*. What do we know about communicating risk? A brief review and suggestion for contextualising serious, but rare, risk, and the example of COX-2 selective and non-selective NSAIDs. *Arthritis Res Ther* 2008;**10**:R20.
36. **Pham T**, Van Der Heijde D, Lassere M, *et al*, OMERACT-OARSI. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *J Rheumatol* 2003;**30**:1648–54.

Take advantage of BMJ Journals' remarkable catalogue of titles with Related Collections

No busy professional has time to browse through all pertinent journals to find relevant articles, but with Related Collections you no longer have to. Follow the "Related Collections" link from any article and use the "Show Collections from other Journals" to expand your search across all BMJ Journals. Or simply follow the "Browse by topic" link on the home page. By setting up your own collections and receiving email alerts every time an article is added to your chosen area, you can build up your own significant body of knowledge.