PAIN® 146 (2009) 238–244

PAIN®

www.elsevier.com/locate/pain

Review and recommendations

Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations

Robert H. Dworkin [a,*], Dennis C. Turk [b], Michael P. McDermott [c], Sarah Peirce-Sandner [d], Laurie B. Burke [e], Penney Cowan [f], John T. Farrar [g], Sharon Hertz [e], Srinivasa N. Raja [h], Bob A. Rappaport [e], Christine Rauschkolb [i], Cristina Sampaio [j]

[a] Departments of Anesthesiology and Neurology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
[b] Department of Anesthesiology, University of Washington, Seattle, WA, USA
[c] Departments of Biostatistics and Computational Biology and Neurology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
[d] Department of Anesthesiology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
[e] United States Food and Drug Administration, Bethesda, MD, USA
[f] American Chronic Pain Association, Rocklin, CA, USA
[g] Department of Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, USA
[h] Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, MD, USA
[i] Johnson & Johnson Pharmaceutical Research & Development LLC, Raritan, NJ, USA
[j] Faculdade de Medicina de Lisboa, Portugal

ARTICLE INFO

Article history:
Received 28 April 2009
Received in revised form 19 June 2009
Accepted 17 August 2009

Keywords:
Chronic pain
Randomized clinical trials
Group differences
Clinical importance
Clinical meaningfulness
Effect size

ABSTRACT

An essential component of the interpretation of results of randomized clinical trials of treatments for chronic pain involves the determination of their clinical importance or meaningfulness. This involves two distinct processes—interpreting the clinical importance of individual patient improvements and the clinical importance of group differences—which are frequently misunderstood. In this article, we first describe the essential differences between the interpretation of the clinical importance of patient improvements and of group differences. We then discuss the factors to consider when evaluating the clinical importance of group differences, which include the results of responder analyses of the primary outcome measure, the treatment effect size compared to available therapies, analyses of secondary efficacy endpoints, the safety and tolerability of treatment, the rapidity of onset and durability of the treatment benefit, convenience, cost, limitations of existing treatments, and other factors. The clinical importance of individual patient improvements can be determined by assessing what patients themselves consider meaningful improvement using well-described methods. In contrast, the clinical meaningfulness of group differences must be determined by a multi-factorial evaluation of the benefits and risks of the treatment and of other available treatments for the condition in light of the primary goals of therapy. Such determinations must be conducted on a case-by-case basis, and are ideally informed by patients and their significant others, clinicians, researchers, statisticians, and representatives of society at large.

1. Introduction

An essential component of the interpretation of results of randomized clinical trials involves the determination of their clinical importance or meaningfulness, which involves two distinct processes—interpreting the clinical importance of *individual patient* improvements and of *group differences*. Unfortunately, the distinction between the clinical importance of individual patient improvements and the clinical importance of group differences is frequently misunderstood. Although methods for the determination of the clinical importance of individual patient improvements have now been well described [11], the interpretation of group differences is less well defined.

The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recently reviewed and recommended specific methods that can be used for interpreting the clinical importance of treatment outcomes for individual patients in chronic pain clinical trials [11]. These recommendations included a set of provisional benchmarks for interpreting changes in measures of pain, physical and emotional functioning, and global improvement that represent outcome domains recommended previously by IMMPACT [10,35]. The methods that were discussed for determining clinical importance and the recommended bench-

* Corresponding author. Address: Departments of Anesthesiology and Neurology, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Box 604, Rochester, NY 14642, USA. Tel.: +1 585 275 3524; fax: +1 585 473 5007.
E-mail address: robert_dworkin@urmc.rochester.edu (R.H. Dworkin).

0304-3959/$36.00
doi:10.1016/j.pain.2009.08.019

marks all involved changes that occur within individuals from the beginning of a clinical trial to its conclusion. For example, decreases in patients' pain intensity of ⩾30% were considered "moderately important" improvements, whereas decreases of ⩾50% were considered "substantial" improvements [11].

The determination of criteria for clinically meaningful improvements for individual patients is necessary to categorize patients as "responders" or "non-responders." Identifying which patients can be considered responders is often a critical aspect of interpreting clinical trial results. Responder analyses make it possible to compare the percentages of patients who achieve meaningful outcomes between treatment and control groups or between different treatment conditions, a readily interpretable approach to presenting clinical trial outcomes [13]. However, recommendations for determining clinically meaningful improvements in patients do not address an equally important consideration in the interpretation of clinical trial results—specifically, what magnitude of difference *between* treatment groups should be considered clinically meaningful?

In the IMMPACT recommendations describing the determination of clinically important changes for individual patients, the authors emphasized that the importance of group differences "can only be established in the broader context of the disease being treated, the currently available treatments, and the overall risk–benefit ratio of the treatment" [11,p. 108]. In this review, our objectives are to amplify the IMMPACT recommendations by first briefly discussing the essential differences between interpreting the clinical importance of patient improvements and of group differences and by subsequently discussing the factors that must be considered when evaluating the clinical importance of group differences. We focus on interpreting the *results* of analyses of group differences and not methods for ascertaining the minimal group differences that would be considered clinically important for sample size determination, which require that a specific value or range of values be *pre-specified* for the differences. Important information relevant to the interpretation of group differences, such as adverse events, only becomes available at the completion of the trial and, hence, cannot be used for sample size determination. Detailed discussion of sample size determination is thus beyond the scope of this article; however, we briefly address this topic in Section 5.

2. The clinical importance of patient improvements

Change in pain intensity typically serves as the primary efficacy endpoint in clinical trials of interventions for chronic pain. However, patients with chronic pain have multiple symptoms as well as diminished physical, emotional, and social functioning [35]. Diverse measures have been developed to assess pain and its impact on health-related quality of life, but because pain and many of its consequences involve subjective experiences, these assessments frequently involve patient-reported outcomes [37].

Considerable effort has been devoted to quantifying the magnitude of change in pain intensity that is considered clinically meaningful to chronic pain patients [e.g., 7,12,14,15,19]. Efforts to identify changes in pain intensity and other chronic pain outcomes that are clinically meaningful to patients have typically used either distribution-based or anchor-based methods [11,24]. Examples of distribution-based methods for determining clinically important changes include using the standard error of measurement. Anchor-based methods examine the relationship between, for example, changes in pain and another measure that serves to "anchor" the clinical importance of these changes, such as patient ratings of treatment satisfaction [32]. Because primary and secondary endpoints in chronic pain trials typically consist of patient-reported outcomes, anchor-based methods are generally considered more informative for assessing what patients consider clinically meaningful, although distribution-based methods provide valuable supportive information [11]. Although not without shortcomings, the use of global measures of improvement or overall treatment satisfaction in chronic pain trials allows patients to provide their integrated evaluation of a treatment, including but not limited to relief of pain, and such measures therefore have unique value as anchors in establishing clinical importance.

A very influential example of an anchor-based approach to establishing criteria for clinically important changes in pain intensity is the analysis performed by Farrar et al. [14], who determined the changes in pain intensity that were associated with patient ratings of global improvement at the end of each of 10 placebo-controlled pregabalin clinical trials in patients with painful diabetic peripheral neuropathy (DPN), postherpetic neuralgia (PHN), fibromyalgia, osteoarthritis (OA), and low back pain. Consistent results were found across the five different chronic pain conditions and in patients administered pregabalin and in those administered placebo. Receiver operating characteristic curve analyses demonstrated that decreases from baseline of 1.7 or more points (approximately 30% or greater) on a 0–10 numerical rating scale of pain intensity discriminated well between patients who provided ratings of being "much improved" or "very much improved" and patients who thought they were minimally improved, unchanged, or worse. Multiple anchor-based analyses have confirmed that patients consider pain intensity reductions on numerical or visual analogue scales of at least 2 points or 30% to be moderately clinically meaningful, and that a reduction of 1 point or 10–20% represents a minimally important change [11]; the clinical importance of changes on categorical pain intensity scales, however, has received less attention.

3. The clinical importance of group differences

Equally important to the determination of the clinical importance of improvements for individual patients is the interpretation of the clinical importance of group differences between treatment and placebo or between different treatments, perhaps especially when such differences, although statistically significant, are relatively small. It is crucial to recognize that criteria for clinically important changes in individuals cannot be extrapolated to the evaluation of group differences [4,6,17,33,37]. For example, a 2-point decrease on a 0–10 pain intensity scale, as discussed above, can be considered a clinically important improvement for individual patients, but it should not be concluded that a 2-point difference in mean pain reduction between an analgesic treatment and placebo is therefore necessary for a treatment benefit to be considered clinically important [11].

Indeed, in an important discussion of patient-reported outcomes, the U.S. Food and Drug Administration noted that "When defining meaningful change on an individual patient basis (i.e., responder), that definition is generally larger than the minimum important difference for application to group mean comparisons" [37]. The generally larger magnitude of changes required for clinically meaningful improvements in individual patients compared with those that represent meaningful group differences is illustrated by a meta-analysis of OA knee pain trials [5]. The results demonstrated that differences in mean pain intensity (on a 100 mm visual analogue scale) between placebo and various treatments—acetaminophen (paracetamol), chondroitin, glucosamine, intraarticular steroids, opioid analgesics, and oral and topical non-steroidal anti-inflammatory drugs (NSAIDs)—following 12 weeks of treatment were 10 mm for chondroitin and opioid analgesics and ranged from 4 mm to approximately 6.5 mm for the other treatments. The authors concluded that the differences

in mean response between existing treatments for OA knee pain and placebo never exceeded the threshold of 10 mm for patient reports of a "minimal perceptible difference" and were always much less than the threshold of 20 mm for an "important improvement" derived from previous literature. Similarly, the differences in change from baseline to endpoint between patients administered duloxetine or pregabalin and placebo in eight trials of painful DPN ranged from 0.9 to 1.5 points on 0–10 pain intensity scales [3,16,23,27–29,34,38] and never approached the approximately 2-point change associated with patient reports of being "much improved" in the Farrar et al. study described above [14].

One possible interpretation of the results of these clinical trials of OA knee pain and painful DPN is that the mean differences in response that were found between the treatment and placebo groups are not clinically meaningful. However, of existing pharmacologic treatments, acetaminophen and NSAIDs are internationally considered either first- or second-line for OA pain [1,21,39] and duloxetine and pregabalin are internationally considered either first- or second-line for painful DPN [2,9,25]. Of course, widespread clinical use does not provide evidence of clinical meaningfulness. If differences in mean response of the magnitude found in the trials of these medications were not considered clinically relevant benefits, however, clinicians would have limited therapeutic options for the pharmacologic treatment of these and other chronic pain conditions.

It is not surprising that the improvements patients consider clinically meaningful are generally larger than the differences found between efficacious treatments and placebo in chronic pain clinical trials. Meaningful change in individual patients reflects any effects of the active treatment, placebo and other non-specific effects of the clinical setting, natural history and spontaneous resolution, and statistical regression to the mean. Differences between treatment and placebo groups, however, reflect the *incremental* benefits of active treatments that contribute to improvement after subtracting out placebo and other non-specific effects, natural history, and regression to the mean, for example, the pharmacologic effects of a medication. In addition, the differences between treatment and placebo groups in chronic pain clinical trials are limited by the magnitudes of the responses in the placebo groups, which can be substantial [22,26] and reflect multiple factors, especially placebo and other non-specific effects of clinical trial participation. Although these factors also affect response in patients receiving active treatment, a substantial response in the placebo group can attenuate the group difference with an active treatment if there is a "floor" below which treatment rarely reduces pain.

4. Factors to consider in determining the clinical meaningfulness of group differences

Given their critical differences, evaluations of the clinical meaningfulness of group differences in chronic pain trials should not be based on criteria for evaluating clinically meaningful changes in individual patients. Rather, the evaluation of group differences should be carried out on a case-by-case consideration of the various characteristics of a specific treatment, the population of patients to be treated, and the risk–benefit ratio [33].

4.1. Primary efficacy outcome

The statistical significance of the primary efficacy analysis must be established before evaluating whether the group difference is clinically important. Assuming that the primary efficacy analysis reveals a statistically significant treatment benefit, there are multiple factors that should be considered when evaluating the magnitude of the group difference (Table 1). Perhaps the most informative of these is a comparison of the magnitude of the treatment benefit with those of other available treatments that are generally considered to have clinically meaningful treatment benefits. A group difference (vs. placebo) in the primary efficacy outcome that is comparable to or better than what has been reported for such treatments provides support for the clinical meaningfulness of the treatment benefit, whereas a group difference that is appreciably smaller than those of other therapies is likely to be of less relevance.

In its discussion of patient-reported outcomes, the FDA noted that when clinical trials show small treatment effect sizes (i.e., standardized mean differences between treatment groups), "it may be more informative to examine the distribution of responses between treatment groups to more fully characterize the treatment effect" [37]. This can be done with various types of responder analyses [8], including the percentage of patients showing a clinically meaningful pain reduction (e.g., $\geqslant$30% decrease from baseline) [14]; the percentage of patients reporting different levels of response on ratings of overall improvement or treatment satisfaction; and a cumulative proportion of responder analysis [13], which provides a method for displaying the level of response at all possible cut-off points (e.g., $\geqslant$30% decrease from baseline, $\geqslant$40% decrease from baseline) rather than a single point. Such responder analyses provide important information beyond the group difference in mean response because they make it possible to determine whether a subgroup of patients may experience meaningful or even substantial benefits even though the overall mean difference is small [17,33]. Fig. 1 illustrates how a relatively small difference between group means can be consistent with the existence of a subgroup of patients experiencing a meaningful response to treatment (discussed in greater detail in Section 5).

On the basis of such considerations, responder outcomes have assumed a greater role in the analysis, interpretation, publication, and evaluation of the results of clinical trials of pain treatments. The percentages of patients who are responders are readily understandable and facilitate comparison of clinical trial results. The

Table 1
Factors to consider in determining the clinical meaningfulness of group differences.

- Statistical significance of the primary efficacy analysis (typically necessary but not sufficient to determine that the group difference is clinically meaningful)
- Magnitude of improvement in the primary efficacy outcome with treatment
- Results of responder analyses
- Treatment effect size compared to available treatments
- Rapidity of onset of treatment benefit
- Durability of treatment benefit
- Results for secondary efficacy endpoints (e.g., improvements in physical and/or emotional functioning)
- Safety and tolerability
- Convenience
- Patient adherence
- Cost
- Different mechanism of action vs. existing treatments
- Limitations of available treatments
- Other benefits (e.g., few or no drug interactions, availability of a test that predicts a good therapeutic response)

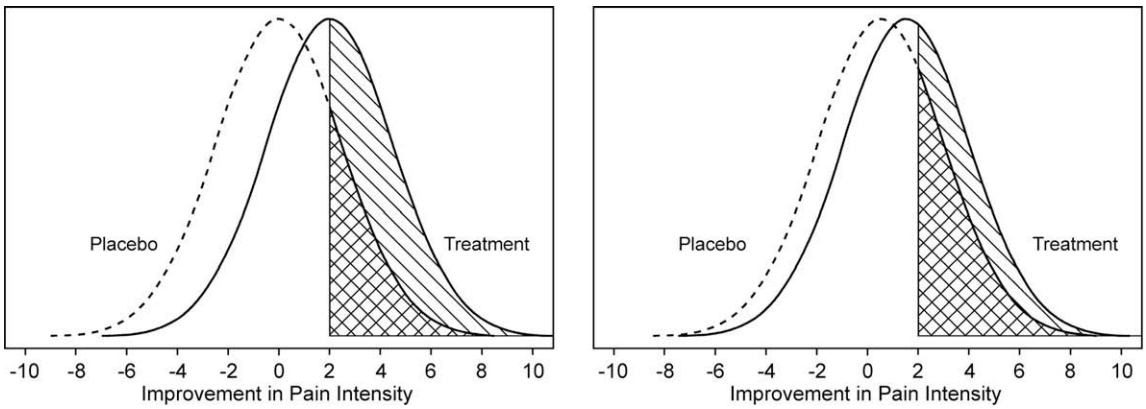

Fig. 1. Hypothetical distributions of changes from baseline in pain intensity (expressed as improvement on a 0–10 numerical rating scale). Both distributions are normal with a standard deviation of 2.5 points. Left-side panel: The placebo group has a mean improvement of zero and the treatment group has a mean improvement of 2.0. In this case, the percentage of placebo group subjects experiencing a clinically important improvement of 2.0 points is 21% (relative area under the placebo group curve indicated by hatched lines) compared to 50% of the subjects in the treatment group (corresponding relative area under the treatment group curve). The number needed to treat to achieve a clinically important improvement is 3.45. Right-side panel: The placebo group has a mean improvement of 0.5 and the treatment group has a mean improvement of 1.5. In this case, the percentage of placebo group subjects experiencing a clinically important improvement of 2.0 points is 27% (relative area under the placebo group curve indicated by hatched lines) compared to 42% of the subjects in the treatment group (corresponding relative area under the treatment group curve). The number needed to treat to achieve a clinically important improvement is 6.67.

magnitude of the group difference in the responder percentages, however, must still be interpreted. This process raises similar issues as the interpretation of group differences in mean response; for example, is a statistically significant difference between responder percentages of 50% in an active treatment group and 45% in a placebo group clinically meaningful? Also, the definition of a "responder" is somewhat arbitrary (e.g., should someone with a 29% change in pain intensity be considered a "non-responder" while someone with a 31% change in pain intensity is considered a "responder"?) and the dichotomization of a continuous outcome variable discards useful information and sacrifices statistical power [31]. Nevertheless, the conceptualization of group differences in terms of percentages of responders may be easier for clinicians and potentially useful in sample size determination (see Section 5).

Other aspects of the data collected for the primary efficacy outcome measure can contribute important information in evaluating the clinical importance of the treatment effect size. These include such potential secondary efficacy endpoints as the rapidity with which improvement occurs and the durability of the benefit; all other things being equal, treatments that have a rapid onset and last a long time are better than those lacking these features.

4.2. Secondary efficacy outcomes

The need to evaluate the pain relief associated with a chronic pain treatment in a context provided by assessments of physical functioning, emotional functioning, sleep, and other outcome domains is well recognized [35]. Improvements on such secondary endpoints, which almost always include assessments of what is broadly considered health-related quality of life [18], can provide important information regarding the overall therapeutic benefit beyond reduction in pain intensity. Lack of significant group differences on such secondary outcomes, however, may reflect inadequate statistical power when sample size estimates have been based on the primary outcome measure and not on the secondary endpoints or when corrections for multiplicity are necessary [36].

The use of secondary endpoints to interpret the meaningfulness of changes in pain intensity can be illustrated by considering two scenarios for a trial with a statistically significant but modest difference in mean pain intensity between the treatment and placebo groups. In the easiest to interpret scenario, one or more secondary outcome measures—for example, measures of physical functioning and sleep—show statistically significant improvement with treatment (that ideally is also clinically meaningful) and thereby corroborate the treatment difference in pain intensity and expand understanding of the therapeutic benefit of the intervention. In the second scenario, the secondary outcomes do not show clear numerical benefits of treatment and may even demonstrate statistically significant evidence of worsening. Assuming that this is not a result of selecting inappropriate measures, worsening with treatment for important secondary outcomes would suggest that a modest benefit in pain intensity might not be of sufficient clinical importance to warrant use of the treatment. Importantly, the determination of clinically meaningful changes for such secondary outcomes as measures of physical and emotional functioning faces the same challenges as for pain [17,18,30].

4.3. Safety and tolerability

The safety and tolerability of a treatment are essential components in interpreting the results of a clinical trial and determining the overall benefit that it can provide to patients [20,35]. Both the incidence and severity of adverse events must be taken into account in any interpretation of the magnitude of a treatment benefit. For example, a treatment with a small benefit relative to placebo that has excellent safety and tolerability would be viewed more favorably than one that has the same treatment benefit but is associated with relatively frequent moderate-to-severe adverse events. A rigorous examination of all adverse effects of treatment, including laboratory results and other assessments of any worsening in the patient's health or well-being, is essential to provide a balanced assessment of therapeutic benefit. Such evaluations would ideally take into account patient as well as clinician perspectives on these adverse outcomes.

4.4. Other factors

The overall benefit of a potential treatment must also be viewed in light of its anticipated ease of use by patients in the community and the likelihood that they will adhere to treatment. Clinical trials can provide some evidence of patient adherence to treatment, but their tightly controlled nature can make generalization to the community problematic and judgments about anticipated ease of use and adherence may require other considerations. For example, oral

medications are generally preferred to those that require injection, and a medication that can be taken in the morning and at bedtime is not only more convenient for patients than one that must be taken five times daily but will also promote better adherence with therapy. In addition, a treatment that either does not require titration or that can be rapidly titrated to optimum dosage will be preferred by patients (and clinicians) to one that requires a lengthy and closely monitored titration to achieve its efficacious dosage. The cost of a treatment is another important source of patient non-adherence with treatment, of course, and is also important in considerations of treatment cost-effectiveness.

In approximately half of patients with chronic pain, existing treatments are not effective or are poorly tolerated, and in patients who do respond, it is relatively rare for pain to be completely relieved or even reduced to mild severity. Because chronic pain treatments have incomplete efficacy, a treatment whose mechanism(s) of action is different from those of existing therapies may be effective in patients who are currently non-responders, presumably because the pathophysiologic mechanism of the patient's pain is targeted by the new treatment and not by existing therapies. For similar reasons, treatments with a different mechanism of action could also be used in combination with existing therapies to augment treatment response in those patients who are partial responders [9]. Moreover, because chronic pain treatments are often poorly tolerated, a treatment with different mechanisms of action may be better tolerated by patients who cannot tolerate the adverse effects of existing therapies. For these reasons, the benefit of a treatment with a different mechanism of action from those of existing therapies could be considered more clinically meaningful than if the treatment had the same mechanism of action as existing therapies.

The limitations of any available treatments should also be taken into account when interpreting group differences. Specifically, a modest benefit compared with placebo may be more clinically meaningful if existing therapies have important limitations, such as poor safety and tolerability or limited efficacy, compared with a similar benefit occurring in a context of existing therapies that have excellent safety, tolerability, and efficacy.

There are other factors specific to certain treatments that must be considered in evaluating the clinical meaningfulness of their benefits. These include an absence or limited number of drug interactions, especially because many chronic patients are older and may not only be receiving other analgesics but are likely to be taking a variety of non-analgesic medications. In addition, modest benefits with a treatment for which there is a simple test or procedure available that predicts whether a patient will have a positive therapeutic response (e.g., a trial of transcutaneous electrical nerve stimulation) may be more clinically meaningful than comparable benefits associated with a treatment for which response cannot be predicted. In many circumstances, patients, clinicians, and third-party payers would consider it beneficial to be able to predict therapeutic response because increasing the likelihood of a positive response reduces the time, risks, and expenses associated with therapeutic failure.

Finally, any research design features of the clinical trial that could have attenuated the magnitude of the group difference in response must also be considered in its evaluation. For example, insufficient pain at baseline could make it difficult or even impossible for an active treatment to show greater reduction in pain compared with placebo. In addition, patients in placebo groups are likely to use more rescue medication than patients treated with efficacious analgesics, and it is certainly possible that a substantially greater use of rescue analgesics in a placebo group could attenuate the difference in pain relief compared with the active treatment. Of course, clinical meaningfulness cannot be attributed to a group difference simply because the trial has design limitations. However, such factors can provide possible explanations for more modest benefits than would otherwise have been expected and thereby provide a basis for conducting additional studies (ideally, without such limitations) to determine whether the magnitude of the treatment benefit has been underestimated.

5. Connections between important individual changes and important group differences

The concepts of clinically important individual changes and clinically important group differences are frequently confused in the medical literature as well as in research, particularly when applied to the problem of sample size determination for a clinical trial. The following example illustrates the differences as well as some connections between the two concepts. Suppose that the primary outcome variable in a clinical trial is the change from baseline in pain intensity, as measured on a 0–10 numerical rating scale. Also, assume for simplicity that this outcome variable is normally distributed and that the standard deviation is 2.5 points in both the treatment group and the placebo group, so that treatment causes a shift in the distribution of outcomes (Fig. 1).

In the left-side panel of Fig. 1, it is assumed that there is no placebo effect (i.e., the placebo group has a mean improvement of 0), so 21% of the subjects administered placebo experience a moderately clinically important reduction in pain intensity of at least 2 points [11]. The group difference in mean response in this example is chosen to be exactly the same as the clinically important individual change, as has sometimes been done when calculating the sample size for a trial; this implies that 50% of the subjects in the treatment group experience a clinically important change. The group difference of 2 points in this example corresponds to a number needed to treat (NNT) of 3.45 on the basis of the proportion of subjects experiencing a clinically important change. This treatment effect size might be appropriate in some circumstances, but for the reasons discussed in Section 4, smaller values of the group difference in mean response may be more appropriate.

In the right-side panel of Fig. 1, with a small placebo effect of 0.5 point and a group difference in mean response of 1 point, 27% of the subjects in the placebo group but an additional 15% of the subjects in the treatment group (i.e., 42%) experience a clinically important change, resulting in an NNT of 6.67. These two examples illustrate that considering clinically meaningful percentages of responders for the treatment vs. control groups may be more helpful for sample size determination than directly considering the value of the difference in group means, although the latter value may ultimately be required. When the outcomes are not normally distributed, if the variances are unequal, or if the treatment causes something other than a simple shift in the distribution of outcomes, the same principles would apply but the connection between the clinically important individual change and the difference in group mean responses is not as straightforward.

6. Conclusions

The clinical importance of individual patient improvements and of group differences are distinct aspects of the results of a clinical trial, and both play an important role in the evaluation of its outcomes. The clinical importance of individual patient improvements in chronic pain trials can be determined by assessing what patients themselves consider meaningful improvement using well-defined methods. In contrast, the clinical meaningfulness of group differences must be determined by a multi-factorial evaluation of the benefits and risks of the treatment and of other available treatments for the condition in light of the primary goals of therapy. Differences in mean reductions in pain between active treatment

and placebo groups do not adequately describe the potential benefits of a treatment in the population of individuals with chronic pain. Although a statistically significant group difference in pain intensity is typically necessary for a demonstration of efficacy, a determination of the clinical meaningfulness of this difference should not be based solely on its magnitude. Group mean differences can obscure meaningful individual patient improvements and other benefits and risks. Information about percentages of responders and evaluation of secondary outcomes, safety and tolerability, and the other factors we have discussed (see Table 1) must all be considered to adequately understand the therapeutic benefit associated with a treatment for chronic pain.

Unfortunately, because this multi-factorial evaluation must consider the benefits and risks of the treatment and of other available treatments for the condition, it is impossible to provide specific guidelines for determining whether or not a specific group difference is clinically meaningful. Such determinations must be conducted on a case-by-case basis, and are ideally informed by patients and their significant others, clinicians, researchers, statisticians, and representatives of society at large. Although such a process may appear daunting, it bears striking resemblance to the advisory committees and similar groups that are routinely assembled by regulatory agencies around the world when evaluating the efficacy and safety of medical interventions. The involvement of multiple stakeholders ensures a comprehensive evaluation that considers the different and complementary perspectives necessary to determine whether the benefits of a treatment outweigh its risks and constitute a clinically meaningful addition to available therapies.

Acknowledgements

The views expressed in this article are those of the authors, none of whom have financial conflicts of interest related to the material presented in this manuscript. No official endorsement by the US Food and Drug Administration or the pharmaceutical companies that have provided unrestricted grants for the support of IMMPACT to the University of Rochester Office of Continuing Professional Education should be inferred.

The first author thanks Drs. Christopher Murray, Richard Polisson, and Lee Simon for stimulating discussions related to the topic of this article.

References

[1] American Pain Society. Guideline for the management of pain in osteoarthritis, rheumatoid arthritis, and juvenile chronic arthritis. Glenview, IL: American Pain Society; 2002.
[2] Attal N, Cruccu G, Haanpää M, Hansson P, Jensen TS, Nurmikko T, Sampaio C, Sindrup S, Wiffen P. EFNS guidelines on pharmacological treatment of neuropathic pain. Eur J Neurol 2006;13:1153–69.
[3] Arezzo JC, Rosenstock J, Lamoreaux L, Pauer L. Efficacy and safety of pregabalin 600 mg/d for treating painful diabetic peripheral neuropathy: a double-blind placebo-controlled trial. BMC Neurol 2008;8:33.
[4] Beaton DE, Bombardier C, Katz JN, Wright JG, Wells G, Boers M, Strand V, Shea B. Looking for important change/differences in studies of responsiveness. J Rheumatol 2001;28:400–5.
[5] Bjordal JM, Klovning A, Ljunggren AE, Slørdal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: a meta-analysis of randomised placebo-controlled trials. Eur J Pain 2007;11:125–38.
[6] Cella D, Bullinger M, Scott C, Barofsky I. Group vs individual approaches to understanding the clinical significance of differences or changes in quality of life. Mayo Clin Proc 2002;77:384–92.
[7] Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. Spine J 2008;8:968–74.
[8] Dionne RA, Bartoshuk L, Mogil J, Witter J. Individual responder analyses for pain: does one pain scale fit all. Trends Pharmacol Sci 2005;26:125–30.
[9] Dworkin RH, O'Connor AB, Backonja M, Farrar JT, Finnerup NB, Jensen TS, Kalso EA, Loeser JD, Miaskowski C, Nurmikko TJ, Portenoy RK, Rice AS, Stacey BR, Treede RD, Turk DC, Wallace MS. Pharmacologic management of neuropathic pain: evidence-based clinical recommendations. Pain 2007;132:237–51.
[10] Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9–19.
[11] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Haythornthwaite JA, Jensen MP, Kerns RD, Ader DN, Brandenburg N, Burke LB, Cella D, Chandler J, Cowan P, Dimitrova R, Dionne R, Hertz S, Jadad AR, Katz NP, Kehlet H, Kramer LD, Manning DC, McCormick C, McDermott MP, McQuay HJ, Patel S, Porter L, Quessy S, Rappaport BA, Rauschkolb C, Revicki DA, Rothman M, Schmader KE, Stacey BR, Stauffer JW, von Stein T, White RE, Witter J, Zavisic S. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105–21.
[12] Farrar JT. What is clinically meaningful: outcome measures in pain clinical trials. Clin J Pain 2000;16:S106–12.
[13] 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.
[14] Farrar JT, Young Jr JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149–58.
[15] Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. JAMA 1999;282:1157–62.
[16] Goldstein DJ, Lu Y, Detke MJ, Lee TC, Iyengar S. Duloxetine vs. placebo in patients with painful diabetic neuropathy. Pain 2005;116:109–18.
[17] Guyatt GH. Making sense of quality-of-life data. Med Care 2000;38:175–9.
[18] Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622–9.
[19] Hanley MA, Jensen MP, Ehde DM, Robinson LR, Cardenas DD, Turner JA, Smith DG. Clinically significant change in pain intensity ratings in persons with spinal cord injury or amputation. Clin J Pain 2006;22:25–31.
[20] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407–15.
[21] Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, Gunther K, Hauselmann H, Herrero-Beaumont G, Kaklamanis P, Lohmander S, Leeb B, Lequesne M, Mazieres B, Martin-Mola E, Pavelka K, Pendleton A, Punzi L, Serni U, Swoboda B, Verbruggen G, Zimmerman-Gorska I, Dougados M. EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2003;62:1145–55.
[22] Katz J, Finnerup NB, Dworkin RH. Clinical trial outcome in neuropathic pain: relationship to study characteristics. Neurology 2008;70:263–72.
[23] Lesser H, Sharma U, LaMoreaux L, Poole RM. Pregabalin relieves symptoms of painful diabetic neuropathy: a randomized controlled trial. Neurology 2004;63:2104–10.
[24] Lydick E, Epstein RS. Interpretation of quality of life changes. Qual Life Res 1993;2:221–6.
[25] Moulin DE, Clark AJ, Gilron I, Ware MA, Watson CP, Sessle BJ, Coderre T, Morley-Forster PK, Stinson J, Boulanger A, Peng P, Finley GA, Taenzer P, Squire P, Dion D, Cholkan A, Gilani A, Gordon A, Henry J, Jovey R, Lynch M, Mailis-Gagnon A, Panju A, Rollman GB, Velly A. Pharmacological management of chronic neuropathic pain: consensus statement and guidelines from the Canadian Pain Society. Pain Res Manage 2007;12:13–21.
[26] Quessy SN, Rowbotham MC. Placebo response in neuropathic pain trials. Pain 2008;138:479–83.
[27] Raskin J, Pritchett YL, Wang F, D'Souza DN, Waninger AL, Iyengar S, Wernicke JF. A double-blind, randomized multicenter trial comparing duloxetine with placebo in the management of diabetic peripheral neuropathic pain. Pain Med 2005;6:346–56.
[28] Richter RW, Portenoy R, Sharma U, Lamoreaux L, Bockbrader H, Knapp LE. Relief of painful diabetic peripheral neuropathy with pregabalin: a randomized, placebo-controlled trial. J Pain 2005;6:253–60.
[29] Rosenstock J, Tuchman M, LaMoreaux L, Sharma U. Pregabalin for the treatment of painful diabetic peripheral neuropathy: a double-blind, placebo-controlled trial. Pain 2004;110:628–38.
[30] Schunemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcome measures in clinical trials: the clinician's perspective. Health Qual Life Outcomes 2006;4:62.
[31] Senn S. Statistical issues in drug development. 2nd ed. Chichester: John Wiley and Sons; 2007.
[32] ten Klooster PM, Drossaers-Bakker KW, Taal E, van de Laar MA. Patient-perceived satisfactory improvement (PPSI): interpreting meaningful change in pain from the patient's perspective. Pain 2006;121:151–7.
[33] Testa MA. Interpretation of quality-of-life outcomes: issues that affect magnitude and meaning. Med Care 2000;38:166–74.

[34] Tolle T, Freynhagen R, Versavel M, Trostmann U, Young Jr JP. Pregabalin for relief of neuropathic pain associated with diabetic neuropathy: a randomized, double-blind study. Eur J Pain 2008;12:203–13.
[35] Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS, Hewitt DJ, Jadad AR, Katz NP, Kramer LD, Manning DC, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robinson JP, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Witter J. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 2003;106:337–45.
[36] Turk DC, Dworkin RH, McDermott MP, Bellamy N, Burke LB, Chandler JM, Cleeland CS, Cowan P, Dimitrova R, Farrar JT, Hertz S, Heyse JF, Iyengar S, Jadad AR, Jay GW, Jermano JA, Katz NP, Manning DC, Martin S, Max MB, McGrath P, McQuay HJ, Quessy S, Rappaport BA, Revicki DA, Rothman M, Stauffer JW, Svensson O, White RE, Witter J. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Pain 2008;139:485–93.
[37] U.S. Department of Health and Human Services. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. U.S. Department of Health and Human Services; 2006. Available from: http://www.fda.gov/cder/guidance/5460dft.pdf.
[38] Wernicke JF, Pritchett YL, D'Souza DN, Waninger A, Tran P, Iyengar S, Raskin J. A randomized controlled trial of duloxetine in diabetic peripheral neuropathic pain. Neurology 2006;67:1411–20.
[39] Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, Bierma-Zeinstra S, Brandt KD, Croft P, Doherty M, Dougados M, Hochberg M, Hunter DJ, Kwoh K, Lohmander LS, Tugwell P. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16:137–62.