Responder analyses in randomised controlled trials for chronic low back pain: an overview of currently used methods

Nicholas Henschke · Annefloor van Enst ·
Robert Froud · Raymond WG Ostelo

Received: 1 March 2013 / Revised: 23 December 2013 / Accepted: 24 December 2013 / Published online: 14 January 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract

Purpose To provide an overview and a critical appraisal of the use of responder analyses in published randomised controlled trials (RCTs) of interventions for chronic low back pain (LBP). The methodology used for the analyses, including the justification, as well as the implications of responder analyses on the conclusions was explored.

Methods A convenience sample of four systematic reviews evaluating 162 RCTs of interventions for chronic LBP was used to identify individual trials. Randomised trials were screened by two reviewers and included if they performed and reported a responder analysis (i.e. the proportion of participants achieving a pre-defined level of improvement). The cutoff value for responders, the period of follow-up, and the outcome measure used were extracted. Information on how RCT authors justified the

methodology of their responder analyses was also appraised.

Results Twenty-eight articles (17 %) using 20 different definitions of responders were included in this appraisal. Justification for the definition of responders was absent in 80 % of the articles. Pain was the most frequently used domain for the definition of response (50 %), followed by back-specific function (30 %) and a combination of pain and function (20 %). A reduction in pain intensity ≥ 50 % was the most common threshold used to define responders (IQR 33–60 %).

Conclusions Few RCTs of interventions for chronic LBP report responder analyses. Where responder analyses are used, the methods are inconsistent. When performing responder analyses authors are encouraged to follow the recommended guidelines, using empirically derived cut-offs, and present results alongside mean differences.

Keywords Randomised trial · Back pain · Analysis · Responder · Outcomes

N. Henschke (✉)
Institute of Public Health, University of Heidelberg,
Im Neuenheimer Feld 324, 69120 Heidelberg, Germany
e-mail: henschke@uni-heidelberg.de

N. Henschke · A. van Enst · R. WG Ostelo
Department of Epidemiology and Biostatistics and the EMGO+
Institute for Health and Care Research, VU University Medical
Center, Amsterdam, The Netherlands

R. Froud
University College of Health Studies, Campus Kristiania,
Oslo, Norway

R. Froud
Warwick Medical School, University of Warwick, Coventry, UK

R. WG Ostelo
Department of Health Sciences, Faculty of Earth and Life
Science, VU University, Amsterdam, The Netherlands

Introduction

Chronic low back pain (LBP) is a common and important cause of functional disability, work absenteeism and high medical expenses [1]. Many studies have been undertaken to search for effective therapies to reduce the burden of chronic LBP complaints and its consequences [2]. However, the effectiveness of many interventions for chronic LBP remains uncertain, as reflected in the findings of numerous randomised controlled trials (RCTs), systematic reviews and meta-analyses [3]. This uncertainty is often attributed to heterogeneous response to treatment among patients. A ‘responder analysis’ has been proposed to offer

an additional, clinically meaningful way to facilitate the interpretation of trial outcomes, especially when treatment response is heterogeneous [4, 5].

In the interpretation of clinical trial results, two different aspects of clinically relevant differences must be distinguished [6]. The first aspect deals with establishing the magnitude of difference in outcome between the treatment and control groups that is large enough to define the scientific or therapeutic importance of the results. This is usually achieved through comparison of summary measures (i.e. mean differences, when continuous outcome measures are used). The second aspect deals with determining the proportion of patients who, at an individual level, achieve a change in the outcome measure which represents a clinically important change. This can be explored by analysing the number of individuals who improve beyond a set threshold—a responder analysis. If, as is often the case, continuous outcomes are used, the chosen threshold cutoff must be informed by minimally important change (MIC) for an individual to ‘respond’ to an intervention as well as the minimal detectable change of the outcome measure in question [7, 8].

Reporting of the proportion of participants in a trial who achieve these changes at specific follow-up time points is suggested to facilitate comparison of the results of different clinical trials and may help to identify unique subgroups of patients who respond to certain interventions [6, 9]. Moreover, it allows the calculation of the number needed to treat (NNT) which can further improve interpretation [4]. It is currently unknown to what extent responder analyses are used in chronic LBP trials and what methodology is used to define a responder to an intervention. The aim of the current study was to explore the methodology used to perform responder analyses. The objectives were to describe the current use of responder analysis in terms of: justification for the methodology; the outcome measure; the cutoff values and how they were derived; and the follow-up times at which the proportion of responders were calculated.

Materials and methods

The current review utilised the included studies in four systematic reviews that were commissioned by the Dutch Health Care Insurance Council. These systematic reviews were intended to provide a comprehensive overview of all relevant RCTs and determine the effectiveness of the main non-surgical treatment options for chronic LBP. Four systematic reviews were performed which evaluated physical and rehabilitation interventions (such as exercise and behavioural therapy) [10], complementary and alternative medicine [11], pharmacological therapy [12], and injection and denervation procedures [13] in patients with chronic

LBP. The four reviews included all RCTs which evaluated therapeutic interventions in adult subjects (aged 18 years or older) with chronic (>12 weeks duration) non-specific LBP. All of the reviews followed similar methodology which was based on that recommended by the Cochrane Back Review Group [14]. Literature searching was performed on electronic databases such as MEDLINE and EMBASE, as well as by identifying eligible RCTs from previous systematic reviews within the Cochrane Library, published up to December 2008. The risk of bias in all eligible trials had been assessed independently by two authors using an 11-item criteria list [15], with disagreements resolved by discussion or consultation with a third author.

Data collection

Full-text versions of all included RCTs in the four systematic reviews were independently assessed for inclusion in the current study by two authors (NH, AvE). Both authors extracted study-level data from the selected articles. A third reviewer (RO) was consulted when disagreements could not be resolved. Studies were included in the current overview when a responder analysis was presented in “Results” or “Discussion” of the article. A responder analysis was defined as an analysis or presentation of the proportion of participants who achieved a pre-defined level of improvement on one of the main outcomes at a certain time point. Subjects could either have improved by the pre-defined level of improvement and ‘responded’ to therapy, or not achieved the pre-defined improvement, and thus failed to respond to therapy. For the purpose of this review, the response had to be based on either pain intensity (e.g. visual analogue scale (VAS), numerical rating scale (NRS), McGill pain questionnaire), back-specific functional status [e.g. Roland–Morris Disability Questionnaire (RMDQ), Oswestry Disability Index (ODI)], or a combination of both pain and functional improvement. Patient-reported measures of pain and function are the most commonly assessed primary outcomes in LBP trials [16].

A standardised list of items was used for data extraction, considering demographic and methodological data as well as qualitative and quantitative outcomes from the included RCTs. Data included the score for risk of bias, the type of intervention and comparison treatment, the outcome (e.g. RMDQ) and domain of outcome (pain, function, or combined), the criteria or cutoff value to define response, and the reference or source that was used for the definition of response.

Data analysis

A descriptive overview of the methods which are currently employed to perform a responder analysis in all included

Fig. 1 Flowchart of randomised controlled trials in the review

studies was conducted, focusing on differences between cutoffs chosen and the definition of responders, references or sources for definitions and cutoffs, and the follow-up time points used. Descriptions of the interpretation of the responder analysis, such as whether it was the primary or secondary analysis and its implications were also abstracted.

Results

From the four systematic reviews a total of 162 RCTs were screened (Fig. 1). The review of physical and rehabilitation interventions [10] included the most RCTs ($n = 83$), followed by the review on complementary and alternative medicine ($n = 35$), injection therapy and denervation ($n = 27$), and pharmacological therapy ($n = 17$). The screening led to the inclusion of 28 RCTs (17 %) which included a responder analysis of the data (Table 1).

Twenty-five (89 %) of the 28 RCTs satisfied 6 or more of the 11 risk of bias assessment criteria and were considered to have a low risk of bias [17]. A responder analysis was performed at more than one follow-up time point in only three RCTs (Table 1).

All included RCTs evaluated at least one intervention group and one comparison group. After classification of all 66 groups in the 28 RCTs, there were 19 placebo groups, 7 injection therapy, and 6 denervation procedure groups. There were nine pharmacological treatment groups and nine complementary and alternative medicine groups (five acupuncture and four herbal medicines). Two behavioural therapy and eight exercise therapy groups were also included in this review. The remaining six groups were control interventions such as usual care or no treatment.

Of the 28 included RCTs, responder analyses were reported using 20 different definitions of response (Table 2). A reference or source to justify the choice of cutoff or definition of response accompanied only 4 of the 20 definitions (20 %). These references are also provided in Table 2. Most responses were defined as a relative decrease of a specific outcome from baseline levels. Pain

Table 1 Characteristics of included studies

Study	Review	Sample size (n)	Groups	Follow-up (weeks)
Atkinson [38]	Pharmacological	83	4	12
Barendse [27]	Injection	28	2	8
Brinkhaus [52]	Complementary	298	3	8
Chrubasik [32]	Complementary	210	3	4
Chrubasik [33]	Complementary	88	2	6
Coats [39]	Pharmacological	293	2	4
Foster [25]	Injection	31	2	3, 8
Freeman [30]	Injection	57	2	26
Frerick [31]	Complementary	319	2	3
Geurts [29]	Injection	80	2	13
Haake [19]	Complementary	1,161	3	26
Katz [40]	Pharmacological	118	2	12
Keitel [34]	Complementary	150	2	3
Leeuw [18]	Non-medical	73	2	26
Manchikanti [35]	Injection	120	2	13, 26, 52
Manchikanti [36]	Injection	75	3	13, 26, 52
Mendelson [41]	Complementary	154	2	18
Pauza [42]	Injection	56	2	26
Peloso [44]	Pharmacological	336	2	13
Ruoff [43]	Pharmacological	318	2	13
Sherman [20]	Non-medical	96	3	12
Smeets [24]	Non-medical	212	3	52
Soriano [37]	Non-medical	71	2	2
van der Roer [21]	Non-medical	102	2	52
van Kleef [26]	Injection	31	2	8
van Wijk [28]	Injection	81	2	13
Witt [22]	Complementary	2,623	2	13
Yelland [23]	Non-medical	88	4	52

was the most frequently used domain for the definition of response (50 %), followed by back-specific function (30 %) and a combination of pain and function (20 %). A decrease of at least 50 % pain intensity (from baseline levels) using a VAS was the most common cutoff, which

Table 2 Definition of “responders” used in randomised controlled trials (RCTs) of interventions for chronic LBP

Outcome	Definition of ‘responder’	Reference or source	RCTs
Back-specific function (Hannover)	12 % improvement	–	[19]
Back-specific function (Hannover)	20 % improvement	–	[22]
Back-specific function (Oswestry)	Two-grade improvement in one or more functional subsets	–	[25]
Back-specific function (RMDQ)	2-point reduction	[53]	[20, 24]
Back-specific function (RMDQ)	30 % reduction	[54]	[21]
Back-specific function (RMDQ)	50 % reduction	–	[20, 23]
Combined	2-point reduction in VAS and 50 % improvement in general perceived effect	–	[26, 27]
Combined	50 % reduction in median VAS-back, without a drop in daily activities or rise in analgesics; or 25 % reduction in median VAS-back, with a simultaneous rise in daily activities of at least 25 % and a drop in analgesic intake of at least 25 %	–	[28]
Combined	50 % reduction in median VAS-leg without a drop in daily physical activities score, a rise in numerical analgesics rating scale, or both	–	[29]
Combined	No neurologic deficit, improvement in LBOS of >7, improvement in SF-36 bodily pain and physical function of >1 standard deviation	–	[30]
Pain intensity	21 mm reduction on a mean of three visual analogue scales of the MPQ	[55]	[18]
Pain intensity	25 % reduction (VAS)	–	[28, 42]
Pain intensity	30 % reduction (VAS)	[47]	[31, 34, 40, 43, 44]
Pain intensity	33 % improvement on three pain-related items on the Chronic Pain Grade scale	–	[19]
Pain intensity	33 % reduction (VAS)	–	[41]
Pain intensity	50 % reduction (VAS)	–	[23, 25, 28, 31, 34–36, 39, 40, 42–44, 52]
Pain intensity	60 % reduction (VAS)	–	[37]
Pain intensity	67 % reduction (VAS)	–	[41]
Pain intensity	75 % reduction (VAS)	–	[38, 42]
Pain intensity	100 % reduction (VAS)	–	[32, 33, 37]

RMDQ Roland–Morris Disability Questionnaire, *VAS* visual analogue scale, *LBOS* low back outcome scale, *MPQ* McGill Pain Questionnaire

was used in 13 (46 %) of the included RCTs (Fig. 2). All but one study [18] performing a responder analysis for pain intensity used a cutoff equal to or >25 % reduction in pain to define responders.

Studies which evaluated responders according to functional improvement used a variety of back-specific function measures and either a relative [19–23] or absolute [20, 24, 25] improvement as the definition for responders. Five studies [26–30] presented a responder analysis using a combination of pain, function, and/or clinical findings to define responders to the intervention. Some of these studies employed a combination of absolute and relative improvements [26, 27], or anchor and distribution-based approaches, [30] to define responders (Table 2).

Fourteen RCTs (50 %) calculated the proportion of responders as the primary analysis for the study [19, 25–37]. Of these, ten RCTs also provided data on continuous outcomes or mean differences between the intervention groups [19, 26–31, 34–36] and four RCTs did not [25, 32, 33, 37]. Twelve RCTs (43 %) calculated the proportion of responders as a secondary analysis (after evaluating group differences) [18, 20, 22–24, 27, 38–43]. In all 12 of these, the responder analysis supported the results of the primary analysis. Two RCTs (7 %) presented a responder analysis in the discussion section of the report [21, 44]. Finally, in three RCTs (11 %) a responder analysis was used to calculate and present the NNT for the intervention and control groups [31, 42, 44].

Fig. 2 Cutoff values for responder analyses using pain intensity (visual analogue scale unless stated otherwise) as the outcome measure

Discussion

Only 28 out of 162 (17 %) screened RCTs included a responder analysis. Where a responder analysis was included, the cutoff used to define a response varied widely across studies and methodology used to derive the cut point was inconsistent. This is problematic because different cut points can produce different results and eliminate the ability of responder analyses to facilitate comparisons across trials [5, 45]. As currently reported, responder analyses in RCTs of interventions for chronic LBP may offer little to facilitate interpretation as they offer little additional information. Efforts to standardise the methodology of responder analyses along with improvements in reporting may facilitate demonstration of a small mean difference for some interventions coexisting with an attractive number of individual responders to an intervention [4, 5]. This is possible in cases where treatment works well for a proportion of patients and could be especially useful in LBP trials. As responder analyses often require continuous outcomes to be categorised, with consequent loss of information and reduced power for statistical tests [46], they are only recommended as secondary analyses to improve the interpretability of the main analysis [9].

The most commonly reported responder analysis was the proportion of participants with a reduction of 50 % pain intensity from baseline. In a recent consensus statement to improve the interpretation of clinical pain trials, reductions in chronic pain intensity in individuals of at least 10–20 % were suggested to reflect MICs [6]. Reductions of >30 % appear to reflect at least moderate clinically important

changes, and consensus on this threshold has also been informed by a panel considering the effects of measurement error from distribution studies. It is recommended that the percentage of patients with this, or greater, changes in pain relief are reported in clinical trials of chronic pain treatments [6]. In addition, the authors could report reductions of >50 % which appear to reflect more substantial improvements [6]. Other research has suggested that for a range of commonly used outcome measures for back pain, a 30 % improvement from baseline may be considered clinically meaningful when comparing before and after measures for individual patients [8, 47]. It should be noted that these recommendations have only recently been published [6] and, while not expected to be reflected in the included studies in this review (as most were published prior to 2008), can be used as a reference point for future RCTs employing a responder analysis.

The presentation of outcomes from a clinical trial can have a considerable effect on the understanding of results and, similarly, there are potentially important variations in the methods used to describe the results of a responder analysis and the interpretation of the NNT [48, 49]. However, very few authors who performed a responder analysis used the findings to provide an estimate of the NNT, which may have limited the interpretation of these findings for patients and clinicians [4, 5]. Froud et al. [5] performed a responder analysis on a large trial (UK BEAM, $n = 1,334$) of physical treatments for back pain which originally showed small to moderate benefits of spinal manipulation over ‘best care’ in general practice. Using a reduction of five points on the Roland–Morris Disability Questionnaire at 3 months as the cutoff, 62 (24 %) individuals in the best care group and 125 (44 %) in the manipulation group were defined as responders. Calculating the absolute risk reduction and then inverting this gives the NNT [50]. The calculated NNT was 5.2 (95 % CI 3.7–8.8), with corresponding between-group mean difference of 1.4 (95 % CI 0.6–2.1). This means that every five to six patients referred to spinal manipulation instead of ‘best care’, on average, will yield one additional responder to treatment at 3 months. In contrast to the small mean difference originally reported, the NNT was small and could be attractive to clinicians, patients, and policy makers [5].

While these proposed values are not strict definitions of a clinically important change, they do offer a common starting point for future research. Ideally, a cutoff to define ‘responders’ should be backed by empirically derived evidence, decided upon a priori, and set out in the trial analysis plan to avoid the temptation of selecting the best cutoff post hoc. Alternatively, presenting the proportion of responders over a range of cutoff points may allow easier comparisons between treatment groups or applicability to

patient care [51]. One note of caution is that responder analyses open up the possibility of reporting outcomes using relative reporting methods (e.g. as percentage of participants who improved) which, in the case of small absolute differences, can make interventions seem more effective. We recommend that if an author chooses to report outcome in relative terms, these are also accompanied by absolute reporting methods.

One limitation of this study is that only four recent systematic reviews were used as the source of potentially eligible RCTs. However, this does cover the majority of conservative treatment, pharmacological therapy, and injection and denervation trials published before 2010 and therefore provide a good overview of all interventions for chronic LBP. Moreover, the four reviews were all performed according to recommended methods of the Cochrane Collaboration and had comparable search strategies and inclusion criteria.

Using summary measures to report outcomes of RCTs for chronic LBP may not give a full picture of the effect of a treatment, especially as individuals respond in different ways. As the methods used to analyse and present the results of RCTs have important implications on how they are interpreted, responder analyses should ideally be performed using empirically derived cutoffs and presented alongside mean differences. In this way, these analyses can aid in interpretation of outcomes as well as the implementation of findings from RCTs.

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Freburger JK et al (2009) The rising prevalence of chronic low back pain. *Arch Intern Med* 169:251–258
2. Airaksinen O et al (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15(Suppl 2):S192–S300
3. Koes BW, van Tulder MW, Thomas S (2006) Diagnosis and treatment of low back pain. *BMJ* 332:1430–1434
4. Guyatt GH et al (1998) Interpreting treatment effects in randomised trials. *BMJ* 316:690–693
5. Froud R et al (2009) Estimating the number needed to treat from continuous outcomes in randomised controlled trials: methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. *BMC Med Res Methodol* 9:35
6. Dworkin RH et al (2008) Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *Pain* 9:105–121
7. de Vet HC et al (2009) Three ways to quantify uncertainty in individually applied “minimally important change” values. *J Clin Epidemiol* 63:37–45
8. Ostelo RW et al (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 33:90–94
9. Froud R et al (2011) Reporting outcomes of back pain trials: a modified Delphi study. *Eur J Pain* 15:1068–1074
10. van Middelkoop M et al (2010) A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J* 20:19–39
11. Rubinstein SM et al (2010) A systematic review on the effectiveness of complementary and alternative medicine for chronic non-specific low-back pain. *Eur Spine J* 19:1213–1228
12. Kuijpers T et al (2010) A systematic review on the effectiveness of pharmacological interventions for chronic non-specific low-back pain. *Eur Spine J* 20:40–50
13. Henschke N et al (2010) Injection therapy and denervation procedures for chronic low-back pain: a systematic review. *Eur Spine J* 19:1425–1449
14. Furlan AD et al (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 34:1929–1941
15. van Tulder M et al (2003) Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 28:1290–1299
16. Müller U et al (2004) Condition-specific outcome measures for low back pain. Part I: validation. *Eur Spine J* 13:301–313
17. van Tulder MW et al (2009) Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine* 34:1685–1692
18. Leeuw M et al (2008) Exposure in vivo versus operant graded activity in chronic low back pain patients: results of a randomized controlled trial. *Pain* 138:192–207
19. Haake M et al (2007) German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. *Arch Intern Med* 167:1892–1898
20. Sherman KJ et al (2005) Comparing yoga, exercise, and a self-care book for chronic low back pain: a randomized, controlled trial. *Ann Intern Med* 143:849–856
21. van der Roer N et al (2008) Intensive group training protocol versus guideline physiotherapy for patients with chronic low back pain: a randomised controlled trial. *Eur Spine J* 17:1193–1200
22. Witt CM et al (2006) Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. *Am J Epidemiol* 164:487–496
23. Yelland MJ et al (2004) Prolotherapy injections, saline injections, and exercises for chronic low-back pain: a randomized trial. *Spine* 29:9–16
24. Smeets RJ et al (2008) Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain* 134:263–276
25. Foster L et al (2001) Botulinum toxin A and chronic low back pain: a randomized, double-blind study. *Neurology* 56:1290–1293
26. van Kleef M et al (1999) Randomized trial of radiofrequency lumbar facet denervation for chronic low back pain. *Spine* 24:1937–1942
27. Barendse GA et al (2001) Randomized controlled trial of percutaneous intradiscal radiofrequency thermocoagulation for chronic discogenic back pain: lack of effect from a 90-second 70°C lesion. *Spine* 26:287–292
28. van Wijk RM et al (2005) Radiofrequency denervation of lumbar facet joints in the treatment of chronic low back pain: a randomized, double-blind, sham lesion-controlled trial. *Clin J Pain* 21:335–344
29. Geurts JW et al (2003) Radiofrequency lesioning of dorsal root ganglia for chronic lumbosacral radicular pain: a randomised, double-blind, controlled trial. *Lancet* 361:21–26
30. Freeman BJ et al (2005) A randomized, double-blind, controlled trial: intradiscal electrothermal therapy versus placebo for the

- treatment of chronic discogenic low back pain. *Spine* 30: 2369–2377
31. Frerick H et al (2003) Topical treatment of chronic low back pain with a capsicum plaster. *Pain* 106:59–64
 32. Chrubasik S et al (2000) Treatment of low back pain exacerbations with willow bark extract: a randomized double-blind study. *Am J Med* 109:9–14
 33. Chrubasik S et al (2003) A randomized double-blind pilot study comparing Doloteffin and Vioxx in the treatment of low back pain. *Rheumatology* 42:141–148
 34. Keitel W et al (2001) Capsicum pain plaster in chronic non-specific low back pain. *Arzneimittelforschung* 51:896–903
 35. Manchikanti L et al (2008) Lumbar facet joint nerve blocks in managing chronic facet joint pain: one-year follow-up of a randomized, double-blind controlled trial: clinical Trial NCT00355914. *Pain Physician* 11:121–132
 36. Manchikanti L et al (2004) One day lumbar epidural adhesiolysis and hypertonic saline neurolysis in treatment of chronic low back pain: a randomized, double-blind trial. *Pain Physician* 7:177–186
 37. Soriano F, Rios R (1998) Gallium arsenide laser treatment of chronic low back pain: a prospective, randomized and double blind study. *Laser Therapy* 10:175–180
 38. Atkinson JH et al (2007) Efficacy of noradrenergic and serotonergic antidepressants in chronic back pain: a preliminary concentration-controlled trial. *J Clin Psychopharmacol* 27:135–142
 39. Coats TL et al (2004) Effects of valdecoxib in the treatment of chronic low back pain: results of a randomized, placebo-controlled trial. *Clin Ther* 26:1249–1260
 40. Katz N et al (2007) A 12-week, randomized, placebo-controlled trial assessing the safety and efficacy of oxymorphone extended release for opioid-naïve patients with chronic low back pain. *Curr Med Res Opin* 23:117–128
 41. Mendelson G et al (1983) Acupuncture treatment of chronic back pain. A double-blind placebo-controlled trial. *Am J Med* 74:49–55
 42. Pauza KJ et al (2004) A randomized, placebo-controlled trial of intradiscal electrothermal therapy for the treatment of discogenic low back pain. *Spine J* 4:27–35
 43. Ruoff GE et al (2003) Tramadol/acetaminophen combination tablets for the treatment of chronic lower back pain: a multicenter, randomized, double-blind, placebo-controlled outpatient study. *Clin Ther* 25:1123–1141
 44. Peloso PM et al (2004) Analgesic efficacy and safety of tramadol/acetaminophen combination tablets (Ultracet) in treatment of chronic low back pain: a multicenter, outpatient, randomized, double blind, placebo controlled trial. *J Rheumatol* 31:2454–2463
 45. Norman GR et al (2001) Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 39:1039–1047
 46. Streiner DL (2002) Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Canadian journal of psychiatry Revue canadienne de psychiatrie* 47:262–266
 47. Farrar JT et al (2001) Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 94:149–158
 48. Covey J (2007) A meta-analysis of the effects of presenting treatment benefits in different formats. *Med Decis Making* 27:638–654
 49. Croft P, Froud R, Lewis AM (2010) Dropouts and sub-groups—statistics can help but not cure. *Pain* 151:563–564
 50. Cook RJ, Sackett DL (1995) The number needed to treat: a clinically useful measure of treatment effect. *Br Med J* 310:452–454
 51. Farrar JT, Dworkin RH, Max MB (2006) 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* 31:369–377
 52. Brinkhaus B et al (2006) Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Arch Intern Med* 166:450–457
 53. Vollenbroek-Hutten MM et al (2004) Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multi-axial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clin Rehabil* 18:566–579
 54. Jordan K et al (2006) A minimal clinically important difference was derived for the Roland–Morris Disability Questionnaire for low back pain. *J Clin Epidemiol* 59:45–52
 55. Beurskens AJ, de Vet HC, Koke AJ (1996) Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 65:71–76