ISSN 0008-3194 (p)/ISSN 1715-6181 (e)/2023/38–49/$2.00/©JCCA 2023

Impact of hype on clinicians' evaluation of trials – a pilot study

Neil Millar, PhD[1]
Brian Budgell, DC, PhD[2]

Objective: *The purpose of this study was to determine the practicality of using a teleconferencing platform to assess the effect of hype on clinicians' evaluations of reports of clinical trials in spinal care.*

Methods: *Twelve chiropractic clinicians were interviewed via a videoconferencing application. Interviews were recorded and timed. Participant behaviour was monitored for compliance with the protocol. Differences between participants numerical ratings of hyped and non-hyped abstracts based on four measures of quality were analysed using pairwise comparisons (Wilcoxon signed rank test for independent samples). In addition, a linear mixed effects model was fitted with condition (i.e. hype vs. no hype) as a fixed effect and participant and abstract as random effects.*

Impact du battage médiatique sur l'évaluation des essais par les cliniciens - une étude pilote

Objectif : *L'objectif de cette étude était de déterminer s'il était possible d'utiliser une plateforme de téléconférence pour mesurer l'effet du battage médiatique sur les évaluations par les cliniciens des rapports d'essais cliniques dans le domaine des soins de la colonne vertébrale.*

Méthodes : *Douze chiropracticiens ont été interrogés par le biais d'une application de vidéoconférence. Les entretiens ont été enregistrés et chronométrés. Le comportement des participants a été contrôlé pour s'assurer qu'ils respectaient le protocole. Les différences entre les évaluations numériques des participants pour les résumés avec et sans publicité, basées sur quatre mesures de qualité, ont été analysées en utilisant des comparaisons par paire (test de rang signé de Wilcoxon pour les échantillons indépendants). En outre, un modèle linéaire à effets mixtes a été ajusté avec la condition (c'est-à-dire avec ou sans battage publicitaire) comme effet fixe et le participant et le résumé comme effets aléatoires.*

[1] Faculty of Engineering, Information and Systems, University of Tsukuba
[2] Canadian Memorial Chiropractic College

Corresponding author: Brian Budgell, Canadian Memorial Chiropractic College, 6100 Leslie St., Toronto, Ontario, Canada M2H 3J1
E-mail: bbudgell@cmcc.ca
Tel: (416) 482-2340 ext 151

This work was supported by the Japanese Society for the Promotion of Science JSPS under Grant #18K00735 (NM). The authors have no disclaimers, competing interests, or additional sources of support or funding to report in the preparation of this manuscript.

Results: *The interviews and data analysis were conducted without significant technical difficulty. Participant compliance was high, and no harms were reported. There were no statistically significant differences in the quality rankings of hyped versus non-hyped abstracts.*

Conclusion: *The use of a videoconferencing platform to measure the effects of hype on clinicians' evaluations of abstracts of clinical trials is practical and an adequately powered study is justified. Lack of statistically significant results may well be due to low participant numbers.*

(JCCA. 2023;67(1):38-49)

KEY WORDS: pilot study, videoconferencing, hype, medical writing, spinal care, chiropractic

Résultats : *Les entretiens et l'analyse des données se sont déroulés sans difficulté technique majeure. Les participants se sont montrés très coopératifs et aucun problème n'a été signalé. Il n'y a pas eu de différences statistiquement significatives dans le classement de la qualité des résumés avec ou sans battage médiatique.*

Conclusion : *L'utilisation d'une plate-forme de vidéoconférence pour mesurer les effets du battage médiatique sur les évaluations des résumés d'essais cliniques par les cliniciens est pratique et une étude suffisamment puissante est justifiée. L'absence de résultats statistiquement significatifs pourrait bien être due au faible nombre de participants.*

(JCCA. 2023;67(1):38-49)

MOTS CLÉS : étude pilote, vidéoconférence, battage médiatique, rédaction médicale, soins de la colonne vertébrale, chiropratique

Introduction

Abstracts of reports of randomised controlled trials (RCTs) provide primary care physicians with quick and easy access to information regarding the efficacy or effectiveness of treatment. The abstract may be the only source of information on which some readers base their assessment of a trial.[1] Therefore, to allow readers to make a critical and accurate appraisal, information in abstracts should be transparent, objective and sufficiently detailed.

However, there is a growing tendency for researchers to use subjective language to make the research field, methods or results seem more appealing to the target readers – a phenomenon referred to as 'hype'. This is evidenced by an 880% increase in the use of selected hype words (e.g. *crucial, novel, innovative, unprecedented*) in PubMed abstracts from 1974 to 2014[2], with similar trends in fundamental and clinical research journals[3], and in other disciplines[4].

A previous study of hype in a sample of 24 RCTs in orthopaedic medicine and spine care identified a total of 161 instances of hype occurring in all but two reports, with most hype targeted at methods (e.g. *robust, exhaustive, expert*), the outcome of the research (e.g. *vital importance*), and the novelty (e.g. *novel, innovative*).[5] In a follow-up interview study[6], authors identified their motivation for using hype as mainly promotional, but also related it to external editorial intervention, linguistic ability, and replication of conventionalised discourse, underlined by pressure to publish.

Of concern is that hype may undermine objective and disinterested interpretation, and so bias readers' evaluation of new knowledge. There is some evidence clinicians who read abstracts containing *spin*, a phenomenon akin to hype, may be biased towards a positive appraisal of treatment.[7,8] Spin is a related but broader concept involving distortion or misrepresentation of the findings so as to portray the study in a more favourable light[9] – for example, by presenting *post hoc* hypotheses as *a priori*, selectively reporting positive results or recommending a treatment without a clinically important effect.

However, notwithstanding the concerns cited above, there is currently no convincing evidence that hyping actually does bias readers' evaluation. In this pilot study, we explore the feasibility of using a videoconferencing platform to assess whether hype in reports of RCTs influences clinicians' appraisal of research.

Methods

This study was approved by the institutional research ethics board of the Canadian Memorial Chiropractic College (REB approval #2006X01)

Trial design

We conducted a double blind randomised controlled pilot trial to compare how clinicians evaluated a trial when the abstract did or did not contain hype. We randomly assigned clinicians to read and evaluate sets of four abstracts in which two abstracts were in the original form (no hype) and two were manipulated versions (hype). The clinicians were asked to rank the abstracts according to the scale shown below under *outcome measures* and to recall information from each abstract.

Materials

From recent reports of RCTs published in the two leading journals in spinal care (*Spine, European Spine Journal*), we selected four structured abstracts that reported clinical research with human subjects, did not contain hype, and were of comparable length (mean word count 265, s.d. 44). We then added six hyping items (single words and short phrases) to each abstract, resulting in an original and a hyped version of each abstract – example shown in Table 1; all abstracts included as supplemental material in Appendix 1. Based on previous research[5], the additions were typical of hype in RCTs and targeted four aspects of the research: (1) how well the study was implemented (e.g. *carefully designed*); (2) the novelty (*this is the first study to …*); (3) the outcome (*convincing evidence*); and (4) the competence of the researchers (*an experienced radiologist*).

Twelve packages of four abstracts each were assembled by a research assistant not involved in the interviews or data analysis. While each package contained two original and two hyped abstracts, no one package contained the original and hyped version of the same abstract. This allowed for six permutations of each combination of abstracts, so that each permutation was tested on two participants. Assignment of the participants to abstracts was conducted by a researcher not involved in interviewing the participants and was conducted using a random number generator to create the sequence of presentation of packages to participants.

Participants and procedure

Participants were recruited by circulating an email to all clinical faculty at the chiropractic institution of one of the researchers. Inclusion criteria were that the participants were currently clinical instructors in the chiropractic institution of the corresponding author. Thirteen participants volunteered and completed the informed consent document after which an interview time was scheduled on Zoom (https://zoom.us). One participant's data were excluded as detailed below, resulting in analysis of 12 complete sets of data. Participants were not informed that the study was investigating hype in abstracts of RCTs. Rather, they only knew that they were being asked to read and evaluate abstracts. The actual interviews took between 20 and 25 minutes to execute, although no time limits were imposed on the participants. Each interview involved the single participant, the interviewer, who was blinded to the abstracts that the participant was viewing, and a research assistant on a separate computer who displayed abstracts only to the participant after the interviewer had explained the procedure of the interview. When the participant indicated that they had finished reading the abstract, the interviewer posed the four questions listed below under '*Outcome measures*.' Thereafter, the interviewer posed three open-ended questions regarding each abstract. Interviews were recorded to permit off-line transcription of participant comments.

Outcome measures

The primary outcome measures were numerical ratings on a scale of 0 to 10 given in response to the following questions:

1. Based on your reading of this abstract, rate how likely it is that you would implement the findings of this study in the same sorts of patients. (0 = I **certainly would not** implement this treatment; 10 = I **certainly would** implement this treatment)
2. Based on your reading of this abstract, rate how rigorous the study was. (0 = not at all rigorous; 10 = very rigorous)
3. Based on your reading of this abstract, rate how novel the study was. (0 = not at all novel; 10 = very novel)
4. Based on your reading of this abstract, how competent were the researchers to conduct a study of this sort? (0 = incompetent; 10 = extremely competent)

Table 1.
Example abstract without and with hype (in bold font, underlined).
From: Eur Spine J. 2014 Jun;23(6):1204-14 (reprinted with permission).

Original version	Hyped version
PURPOSE: To evaluate the effect of a programme of active self-correction and task-oriented exercises on spinal deformities and health-related quality of life (HRQL) in patients with mild adolescent idiopathic scoliosis (AIS) (Cobb angle <25°).	**PURPOSE:** **<u>This is the first study</u>** to evaluate the effect of a programme of active self-correction and task-oriented exercises on spinal deformities and health-related quality of life (HRQL) in patients with mild adolescent idiopathic scoliosis (AIS) (Cobb angle <25°).
METHODS: This was a parallel-group, randomised, superiority-controlled study in which 110 patients were randomly assigned to a rehabilitation programme consisting of active self-correction, task-oriented spinal exercises and education (experimental group, 55 subjects) or traditional spinal exercises (control group, 55 subjects). Before treatment, at the end of treatment (analysis at skeletal maturity), and 12 months later (follow-up), all of the patients underwent radiological deformity (Cobb angle), surface deformity (angle of trunk rotation) and HRQL evaluations (SRS-22 questionnaire). A linear mixed model for repeated measures was used for each outcome measure.	**METHODS:** This was a parallel-group, randomised, superiority-controlled study in which 110 patients were randomly assigned to a **<u>carefully designed</u>** rehabilitation programme consisting of active self-correction, task-oriented spinal exercises and education (experimental group, 55 subjects) or traditional spinal exercises (control group, 55 subjects). Before treatment, at the end of treatment (analysis at skeletal maturity), and 12 months later (follow-up), for all patients, **<u>an experienced radiologist</u>** undertook **<u>detailed</u>** evaluations of spinal deformity (Cobb angle), surface deformity (angle of trunk rotation) and HRQL evaluations (SRS-22 questionnaire). A linear mixed model for repeated measures was used for each outcome measure.
RESULTS: There were main effects of time ($p < 0.001$), group ($p < 0.001$) and time by group interaction ($p < 0.001$) on radiological deformity: training in the experimental group led to a significant improvement (decrease in Cobb angle of >5°), whereas the control group remained stable. Analysis of all of the secondary outcome measures revealed significant effects of time, group and time by group interaction in favour of the experimental group.	**RESULTS:** There were main effects of time ($p < 0.001$), group ($p < 0.001$) and time by group interaction ($p < 0.001$) on radiological deformity: training in the experimental group led to a significant improvement (decrease in Cobb angle of >5°), whereas the control group remained stable. Analysis of all of the secondary outcome measures revealed significant effects of time, group and time by group interaction in favour of the experimental group.
CONCLUSIONS: The programme of active self-correction and task-oriented exercises was superior to traditional exercises in reducing spinal deformities and enhancing the HRQL in patients with mild AIS. The effects lasted for at least 1 year after the intervention ended.	**CONCLUSIONS:** **<u>The findings provide convincing evidence that</u>** a programme of active self-correction and task-oriented exercises is superior to traditional exercises in reducing spinal deformities and enhancing the HRQL in patients with mild AIS. **<u>It is noteworthy that</u>** the effects lasted for at least 1 year after the intervention ended.

Secondary outcome measures were open-ended verbal responses to the following questions:

1. What 4 or 5 words or phrases do you recall from the abstract?
2. In just a few words, what would you describe as the strengths, if any, of the study?
3. In just a few words, what would you describe as the weaknesses, if any, of the study?

Analytical methods

The researcher performing data analysis was blinded as to which abstracts were original and which were hyped until

all data analysis had been completed. Differences between the numerical ratings for hyped and non-hyped abstracts were analysed using pairwise comparisons (Wilcoxon signed rank test for independent samples). In addition, a linear mixed effects model was fitted with *condition* (i.e. *hype* vs. *no hype*) as a fixed effect and *participant* and *abstract* as random effects. Quantitative analyses were carried out in R.[10]

Verbal responses were transcribed and analysed independently by two researchers for reference to the hyping items. Adjectives and adjectival phrases used in the open-ended responses were coded according to whether or not they could be classified by hype, based on previously published identification of the lexicon of hype.[2-4] Thus, each adjective and adjectival phrase was assigned to a semantic category corresponding to: (1) implementation (the clinical implications of the findings), (2) rigour of design and execution, (3) novelty of the study, (4) the competence of the researchers. Differences were resolved through discussion.

In an analysis conceived *post hoc*, the researchers also coded the open-ended responses according to which CONSORT item(s) for abstracts the response corresponded to.[11] Again, differences were resolved through discussion.

Results

The final sample comprised data from twelve participants: eight males; mean age 42.5 years (±11.8), minimum = 27, maximum = 64; mean years since licensure 14.8 years (±11.4), minimum = 0, maximum= 38; five with post-graduate qualifications. Although data from a thirteenth participant were collected, the data set was discarded as they did not remain on topic during the experiment. The participants read each abstract in an average time of 94 seconds (s.d.= 26s) and reading times did not differ significantly abstract by abstract (p=0.174) nor by the order of presentation (p=0.899).

All interviews were conducted without significant technical issues, but that in one instance there was a transient loss of connection traced to a faulty modem, not to the application used for the interviews. There were no adverse events related to the interview process.

Numerical ratings

Table 2 compares the mean ratings given to original and hyped abstracts. The spread of the ratings is shown in the boxplots in Figure 1. All differences were statistically insignificant, although abstracts containing hype were rated somewhat more favourably in terms of novelty (mean difference = 1.14). Across the other criteria, the mean ratings were slightly lower when the abstracts contained hype.

The results of the Wilcoxon signed rank test for independent samples (effect size, as measured by *r*, and p-value) are given in the final two columns of Table 1. All comparisons were non-significant at $p<0.05$. Post hoc power analyses based on the effect observed on ratings for *novelty* showed that with an alpha = 0.05 and power = 0.80, the projected sample size needed is approximately 90 items, which is equivalent to N = 45 participants. In addition,

Table 2.
Mean ratings by criterion and condition

Rating criteria	Condition	Mean rating	(sd)	Mean difference	r	p
(1) implementation	*original*	6.67	(2.18)	-0.38	0.086	0.560
	hyped	6.29	(2.26)			
(2) rigour	*original*	6.12	(1.85)	-0.04	0.035	0.818
	hyped	6.08	(2.26)			
(3) novelty	*original*	5.50	(2.15)	1.14	0.215	0.139
	hyped	6.46	(2.36)			
(4) competence	*original*	6.46	(1.47)	-0.52	0.088	0.548
	hyped	6.12	(1.92)			

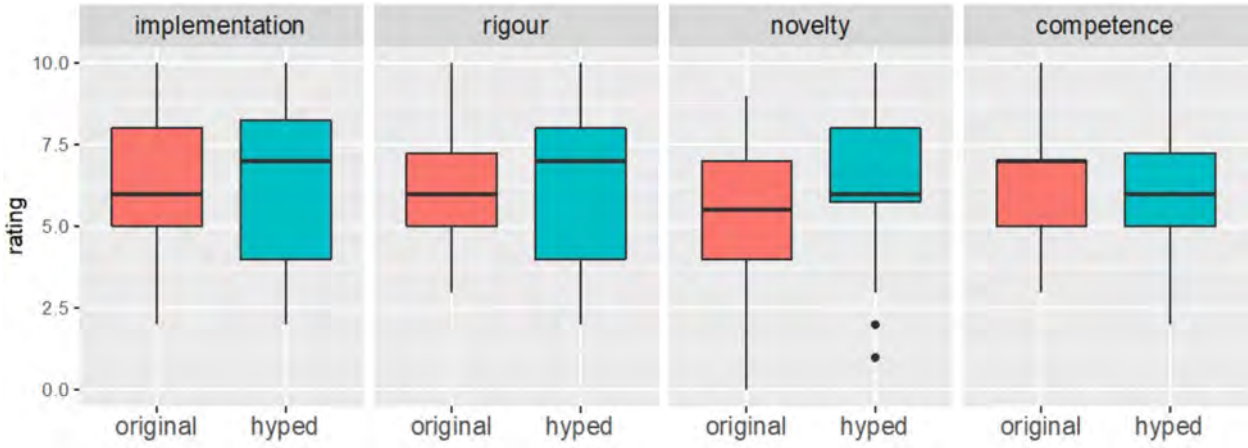

Figure 1.
Comparison of participant ratings of original and hyped abstracts

for each of the outcome measures, a linear mixed effects model was fitted with *condition* as a fixed effect and *participant* and *abstract* as random effects. In all four models, condition (i.e. the presence or absence of hype) had no statistically significant effect on the ratings.

Verbal responses

The design of the trial created 144 exposures to hype: 12 participants x two hyped abstracts x six hype items per abstract. In open-ended responses to hyped abstracts, five participants mentioned a total of seven hyping items either as 'phrases recalled' (*experienced, first study, unique*) or as strengths (*experienced* x2, *novel, qualified*). One of these participants also described the methodology of a hyped abstract as 'rigorous'.

Table 3 shows the classification of participants comments according to their reference to hype and items in the CONSORT 2010 checklist of information to include when reporting a pilot or feasibility randomized trial in a journal or conference abstract.[11]

Discussion

While unstructured conversations via Zoom may be challenged to a degree by small delays in transmission, this did not present a problem for the structured interviews conducted in this study.[12] Furthermore, and perhaps because of the convenience with regard to timing and place for participation, all participants attended their scheduled appointments and all interviews were conducted within the scheduled time. Data collection proceeded as planned, and others have demonstrated that data collection in structured interviews is as effective via Zoom as it is in person.[13] This study was conducted in the midst of the COVID pandemic, and so the use of remote interviews mitigated any risk of disease transmission between participant and researcher. Further, there were no dropouts and no adverse responses to the interview process.

With regard to the experimental results, in this small study there were no statistically significant impacts of hype on clinicians' evaluations of abstracts of randomized clinical trials in spinal care. A post hoc power analysis suggested that in order to demonstrate with confidence the effect of one use of hype (novelty*)* on clinicians' evaluations of abstracts would require a cohort of 45 participants, not an impractical number. In response to open-ended questions, there were relatively few references to hype items and references to CONSORT items in discussions of study designs and weaknesses were sparse and unevenly distributed across items.

Limitations

This was a pilot study with a small cohort, and so the experimental results should not be interpreted as having any implications for the results of an adequately powered

Table 3.
Open-ended responses referring to strengths and weaknesses of studies.

	Strengths (hyped)	Strengths (original)	Weaknesses (hyped)	Weaknesses (original)
Title	4	-	-	-
Trial design	16	15	2	2
Methods per CONSORT				
Participants	4	2	5	1
Interventions	13	10	10	15
Objective	0	0	0	1
Outcome	10	10	6	2
Randomization	10	8	1	1
Blinding (masking)	4	4	3	0
Sum	**57**	**49**	**27**	**22**
Results/Conclusions per CONSORT				
Numbers randomized	7	6	4	3
Recruitment	0	0	0	0
Numbers analysed	1	4	6	5
Outcome	1	2	2	0
Harms	1	0	0	0
Conclusions	0	1	2	2
Sum	**10**	**13**	**14**	**10**

study. Similarly, the facility of conducting interviews in this study should not imply facility with other videoconferencing platforms in other locations where internet connectivity may differ. Additionally, as the study participants were all teaching faculty, the results should not be extrapolated to other populations, for example non-teaching community-based practitioners.

Conclusion

The use of a videoconferencing platform, Zoom, to measure the effects of hype on clinicians' evaluations of abstracts of clinical trials was practical and essentially problem-free in this exercise. Hype in abstracts did not appear to affect clinicians' evaluations of articles in this small pilot study, and so a larger, adequately powered study is justified.

Acknowledgements

This study was supported by JSPS KAKENHI under Grant #18K00735 (XX). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank Jennifer Santoso for able assistance in the randomization of abstracts presented to the participants, and the authors thank Hanna Yang for able assistance in the blinding of the investigators to the abstracts presented to the participants.

References

1. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. Int J Surg. 2011;9(8): 672-677. doi:10.1016/j.ijsu.2011.09.004

2. Vinkers CH, Tijdink JK, Otte WM. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. BMJ. 2015; 351: 1-6. doi:10.1136/bmj.h6467
3. Fraser VJ, Martin JG. Marketing data: has the rise of impact factor led to the fall of objective language in the scientific article? Respir Res. 2009;10: 1-5. doi:10.1186/1465-9921-10-35
4. Weidmann NB, Otto S, Kawerau L. The use of positive words in political science language. PS Polit Sci Polit. 2018;51(03): 625-628. doi:10.1017/S1049096518000124
5. Millar N, Salager-Meyer F, Budgell B. "It is important to reinforce the importance of …": 'Hype' in reports of randomized controlled trials. English Specif Purp. 2019; 54. doi:10.1016/j.esp.2019.02.004
6. Millar N, Budgell B, Salager-Meyer F. Hype in reports of clinical research: the authors' perspectives. English Specif Purp. 2020;60: 53-64. doi:10.1016/j.esp.2020.07.001
7. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. J Clin Oncol. 2014;32(36): 4120-4126. doi:10.1200/JCO.2014.56.7503
8. Tsujimoto Y, Aoki T, Shinohara K, et al. Physician characteristics associated with proper assessment of overstated conclusions in research abstracts: a secondary analysis of a randomized controlled trial. PLoS One. 2019;14(1): e0211206. doi:10.1371/journal.pone.0211206
9. Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. Proc Natl Acad Sci. 2018;115(11): 2613-2619. doi:10.1073/pnas.1710755115
10. R Core Team. R: A Language and Environment for Statistical Computing. 2021. https://www.r-project.org/.
11. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016; 355. doi:10.1136/bmj.i5239
12. Boland JE, Fonseca P, Mermelstein I, Williamson M. Zoom disrupts the rhythm of conversation. J Exp Psychol Gen. 2021.
13. Chuey A, Asaba M, Bridgers S, et al. Moderated online data collection for developmental research: methods and replications. Front Psychol. 2021; 12. doi:10.3389/fpsyg.2021.734398

Appendix 1.
Abstracts

Eur Spine J. 2014 Jun;23(6):1204-14 (reprinted with permission).

Original version	Hyped version
PURPOSE: To evaluate the effect of a programme of active self-correction and task-oriented exercises on spinal deformities and health-related quality of life (HRQL) in patients with mild adolescent idiopathic scoliosis (AIS) (Cobb angle <25°).	**PURPOSE:** **This is the first study** to evaluate the effect of a programme of active self-correction and task-oriented exercises on spinal deformities and health-related quality of life (HRQL) in patients with mild adolescent idiopathic scoliosis (AIS) (Cobb angle <25°).
METHODS: This was a parallel-group, randomised, superiority-controlled study in which 110 patients were randomly assigned to a rehabilitation programme consisting of active self-correction, task-oriented spinal exercises and education (experimental group, 55 subjects) or traditional spinal exercises (control group, 55 subjects). Before treatment, at the end of treatment (analysis at skeletal maturity), and 12 months later (follow-up), all of the patients underwent radiological deformity (Cobb angle), surface deformity (angle of trunk rotation) and HRQL evaluations (SRS-22 questionnaire). A linear mixed model for repeated measures was used for each outcome measure.	**METHODS:** This was a parallel-group, randomised, superiority-controlled study in which 110 patients were randomly assigned to a **carefully designed** rehabilitation programme consisting of active self-correction, task-oriented spinal exercises and education (experimental group, 55 subjects) or traditional spinal exercises (control group, 55 subjects). Before treatment, at the end of treatment (analysis at skeletal maturity), and 12 months later (follow-up), for all patients, **an experienced radiologist** undertook **detailed** evaluations of spinal deformity (Cobb angle), surface deformity (angle of trunk rotation) and HRQL evaluations (SRS-22 questionnaire). A linear mixed model for repeated measures was used for each outcome measure.
RESULTS: There were main effects of time ($p < 0.001$), group ($p < 0.001$) and time by group interaction ($p < 0.001$) on radiological deformity: training in the experimental group led to a significant improvement (decrease in Cobb angle of >5°), whereas the control group remained stable. Analysis of all of the secondary outcome measures revealed significant effects of time, group and time by group interaction in favour of the experimental group.	**RESULTS:** There were main effects of time ($p < 0.001$), group ($p < 0.001$) and time by group interaction ($p < 0.001$) on radiological deformity: training in the experimental group led to a significant improvement (decrease in Cobb angle of >5°), whereas the control group remained stable. Analysis of all of the secondary outcome measures revealed significant effects of time, group and time by group interaction in favour of the experimental group.
CONCLUSIONS: The programme of active self-correction and task-oriented exercises was superior to traditional exercises in reducing spinal deformities and enhancing the HRQL in patients with mild AIS. The effects lasted for at least 1 year after the intervention ended.	**CONCLUSIONS:** **The findings provide convincing evidence that** a programme of active self-correction and task-oriented exercises is superior to traditional exercises in reducing spinal deformities and enhancing the HRQL in patients with mild AIS. **It is noteworthy that** the effects lasted for at least 1 year after the intervention ended.

Spine (Phila Pa 1976). 2000 Jun 15;25(12):1523-32 (reprinted with permission).

Original version	Hyped version
STUDY DESIGN: A prospective randomized controlled trial of exercise therapy in patients who underwent microdiscectomy for prolapsed lumbar intervertebral disc. Results of a pilot study are presented.	**STUDY DESIGN**: A prospective randomized controlled trial of exercise therapy in patients who underwent microdiscectomy for prolapsed lumbar intervertebral disc. Results of a pilot study are presented.
OBJECTIVE: To determine the effects of a postoperative exercise program on pain, disability, psychological status, and spinal function.	**OBJECTIVE**: To determine the effects of a postoperative exercise program on pain, disability, psychological status, and spinal function.
SUMMARY OF BACKGROUND DATA: Microdiscectomy is often used successfully to treat prolapsed lumbar intervertebral disc. However, some patients do not have a good outcome and many continue to have low back pain. The reasons for this are unclear but impairment of back muscle function due to months of inactivity before surgery may be a contributing factor. A postoperative exercise program may improve outcome in such patients.	**SUMMARY OF BACKGROUND DATA:** Microdiscectomy is often used successfully to treat prolapsed lumbar intervertebral disc. However, some patients do not have a good outcome and many continue to have low back pain. The reasons for this are unclear but impairment of back muscle function due to months of inactivity before surgery may be a contributing factor. A postoperative exercise program may improve outcome in such patients.
METHODS: Twenty patients who underwent lumbar microdiscectomy were randomized into EXERCISE and CONTROL groups. After surgery, all patients received normal postoperative care that included advice from a physiotherapist about exercise and a return to normal activities. Six weeks after surgery, patients in the EXERCISE group undertook a 4-week exercise program that concentrated on improving strength and endurance of the back and abdominal muscles and mobility of the spine and hips. Assessments of spinal function were performed in all patients during the week before surgery and at 6, 10, 26, and 52 weeks after. The assessment included measures of posture, hip and lumbar mobility, back muscle endurance capacity and electromyographic measures of back muscle fatigue. On each occasion, patients completed questionnaires inquiring about pain, disability and psychological status.	**METHODS:** Twenty patients who underwent lumbar microdiscectomy were randomized into EXERCISE and CONTROL groups. After surgery, all patients received normal postoperative care that included **detailed advice** from an **experienced physiotherapist** about exercise and a return to normal activities. Six weeks after surgery, patients in the EXERCISE group undertook an **innovative** 4-week exercise program that concentrated on improving strength and endurance of the back and abdominal muscles and mobility of the spine and hips. **Comprehensive** assessments of spinal function were performed in all patients during the week before surgery and at 6, 10, 26, and 52 weeks after. The assessment included measures of posture, hip and lumbar mobility, back muscle endurance capacity and electromyographic measures of back muscle fatigue. On each occasion, patients completed questionnaires inquiring about pain, disability and psychological status.
RESULTS: Surgery improved pain, disability, back muscle endurance capacity and hip and lumbar mobility in both groups of patients. After the exercise program, the EXERCISE group showed further improvements in these measures and also in electromyographic measures of back muscle fatigability. All these improvements were maintained 12 months after surgery. The only further improvement showed by the CONTROL group between 6 and 52 weeks was an increase in back muscle endurance capacity.	**RESULTS:** Surgery improved pain, disability, back muscle endurance capacity and hip and lumbar mobility in both groups of patients. After the exercise program, the EXERCISE group showed further improvements in these measures and also in electromyographic measures of back muscle fatigability. **Notably**, all these improvements were maintained 12 months after surgery. The only further improvement showed by the CONTROL group between 6 and 52 weeks was an increase in back muscle endurance capacity.
CONCLUSION: A 4-week postoperative exercise program can improve pain, disability, and spinal function in patients who undergo microdiscectomy.	**CONCLUSION**: **This is the first randomized controlled trial to demonstrate that** a 4-week postoperative exercise program can improve pain, disability, and spinal function in patients who undergo microdiscectomy.

Spine (Phila Pa 1976). 2005 Apr 1;30(7):711-21 (reprinted with permission).

Original version	Hyped version
STUDY DESIGN: A randomized clinical trial with blinded assessment.	**STUDY DESIGN:** A randomized clinical trial with blinded assessment.
OBJECTIVES: To investigate the clinical efficacy of 2 active interventions for patients with chronic low back pain.	**OBJECTIVES:** To investigate the clinical efficacy of 2 active interventions for patients with chronic low back pain.
SUMMARY OF BACKGROUND DATA: Manual therapy and exercise prescription are treatments frequently prescribed for patients with chronic low back pain. The evidence for the relative benefit of these treatments is limited, and questions concerning the most appropriate type of intervention remain unanswered.	**SUMMARY OF BACKGROUND DATA:** Manual therapy and exercise prescription are treatments frequently prescribed for patients with chronic low back pain. The evidence for the relative benefit of these treatments is limited, and questions concerning the most appropriate type of intervention remain unanswered.
METHODS: Eighty patients with chronic low back pain (>3 months) were randomized to one of the following treatments, involving 8 treatments over 8 weeks; 1) one-to-one treatment involving 30 minutes of manual therapy (mobilizations to the spine) and spinal stabilization exercises, and 2) a 10 station exercise class involving aerobic exercises, spinal stabilization exercises, and manual therapy. Three physiotherapists led the hour long group with a maximum of 10 patients. Questionnaires were completed, and physical measurements were taken by a blinded observer before randomization, at the completion of treatment, and at 6 months and 12 months after the completion of treatment. The intention-to-treat principle was used in data analysis.	**METHODS:** Eighty patients with chronic low back pain (>3 months) were randomized to one of the following treatments, involving 8 treatments over 8 weeks; 1) one-to-one treatment involving 30 minutes of manual therapy (mobilizations to the spine) and spinal stabilization exercises, and 2) a 10 station exercise class involving aerobic exercises, spinal stabilization exercises, and manual therapy. Three **experienced, qualified physiotherapists** led the hour long group with a maximum of 10 patients. **Carefully designed** questionnaires were completed, and **detailed** physical measurements were taken by a **trained** blinded observer before randomization, at the completion of treatment, and at 6 months and 12 months after the completion of treatment. The intention-to-treat principle was used in data analysis.
RESULTS: Eleven patients dropped out of the individual treatment sessions and 7 dropped out of the exercise group. There was a significant reduction (reduced disability) in the questionnaire score in both groups, and there were significant increases in range for all the physical movements tested in both groups. The exercise group was 40% more cost effective than the individual treatments.	**RESULTS:** Eleven patients dropped out of the individual treatment sessions and 7 dropped out of the exercise group. There was a significant reduction (reduced disability) in the questionnaire score in both groups, and there were significant increases in range for all the physical movements tested in both groups. **A novel finding was that** the exercise group was 40% more cost effective than the individual treatments.
CONCLUSION: Both forms of intervention were associated with significant improvement. On-going clinical research is necessary to provide guidance as to the clinical efficacy of various forms of intervention.	**CONCLUSION**: **The findings provide convincing evidence that** both forms of intervention are associated with significant improvement. On-going clinical research is necessary to provide guidance as to the clinical efficacy of various forms of intervention.

Spine (Phila Pa 1976). 2009 Jun 15;34(14):1436-40 (reprinted with permission).

Original version	Hyped version
STUDY DESIGN: This study was a prospective, randomized, controlled study.	**STUDY DESIGN:** This study was a prospective, randomized, controlled study.
OBJECTIVE: To compare the effectiveness of aquatic exercise interventions with land-based exercises in the treatment of chronic low back pain (CLBP).	**OBJECTIVE:** To compare the effectiveness of aquatic exercise interventions with land-based exercises in the treatment of chronic low back pain (CLBP).
SUMMARY OF BACKGROUND DATA: Land-based exercise and physiotherapy are the main treatment tools used for CLBP. Clinical experience indicates that aquatic exercise may have advantages for patients with musculoskeletal disorders.	**SUMMARY OF BACKGROUND DATA:** Land-based exercise and physiotherapy are the main treatment tools used for CLBP. Clinical experience indicates that aquatic exercise may have advantages for patients with musculoskeletal disorders.
METHODS: A total of 65 patients with CLBP were included in this study. Patients were randomly assigned to receive aquatic exercise or land-based exercise treatment protocol. Aquatic exercise program consisted of 20 sessions, 5 x per week for 4 weeks in a swimming pool at 33 degrees C. Land-based exercise (home-based exercise) program were demonstrated by a physiotherapist on one occasion and then they were given written advice. The patients were assessed for spinal mobility, pain, disability, and quality of life. Evaluations were performed before treatment (week 0) and after treatment (week 4 and week 12).	**METHODS:** A total of 65 patients with CLBP were included in this study. Patients were **carefully** randomized to receive aquatic exercise or land-based exercise treatment protocol. **The innovative** aquatic exercise program consisted of 20 **specially designed** sessions, 5 x per week for 4 weeks in a swimming pool at 33 degrees C. Land-based exercise (home-based exercise) program were demonstrated by a **qualified, experienced physiotherapist** on one occasion and then they were given written advice. The patients **underwent detailed assessment of** spinal mobility, pain, disability, and quality of life. Evaluations were performed before treatment (week 0) and after treatment (week 4 and week 12).
RESULTS: In both groups, statistically significant improvements were detected in all outcome measures (except modified Schober test) compared with baseline. However, improvement in modified Oswestry Low Back Pain Disability questionnaire and physical function and role limitations due to physical functioning subpart of Short-Form 36 Health Survey were better in aquatic exercise group ($P < 0.05$).	**RESULTS:** In both groups, statistically significant improvements were detected in all outcome measures (except modified Schober test) compared with baseline. However, improvement in modified Oswestry Low Back Pain Disability questionnaire and physical function and role limitations due to physical functioning subpart of Short-Form 36 Health Survey were better in aquatic exercise group ($P < 0.05$).
CONCLUSION: It is concluded that water-based exercises produced better improvement in disability and quality of life of the patients with CLBP than land-based exercise.	**CONCLUSION: This is the first randomized controlled trial to demonstrate** that water-based exercises produce better improvement in disability and quality of life of the patients with CLBP than land-based exercise.