

**Critical Appraisal and Comments on the
'Physiotherapy briefings for physicians. Low
Back Pain study sparks discussions.'**

By

**WCB Evidence Based Practice Group
*Dr. Craig W. Martin, Senior Medical Advisor***

January 2005

TABLE OF CONTENTS

Page

Background	1
Materials and methods	1
Results	1
1. Frost H, Lamb SE, Doll HA e al. Randomized controlled trial of physiotherapy compared with advice for low back pain. BMJ. Sept 17 2004;329:708-712.....	1
2. Physiotherapy briefing for physicians. Low Back Pain Study Sparks Discussions. Leaflet enclosure with the BCMJ. Fall 2004.....	3
3. Moseley L. Combined physiotherapy and education is efficacious for chronic low back pain. Australian Journal of Physiotherapy. 2002;48:297-302.....	4
4. Saal JA and Saal JS. Nonoperative treatment of herniated lumbar interval disc with radiculopathy. An outcome study. Spine. 1989;14(3):433-437.....	5
5. Hurwitz EL, Morgenstern H, Harber P et al. A randomized trial of medical care with and without physical therapy and chiropractic care with and without physical modalities for patients with low back pain: 6 month follow-up outcomes from the UCLA Low Back Pain Study. Spine. 2002;27:2193-2204.....	5
6. Moffett JK, Torgerson D, Bell-Syer S et al. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs and preferences. BMJ. 31 July 1999;319:279-283.....	6
7. Rainville J, Hartgan C, Martinez E et al. Exercise as a treatment for chronic low back pain. Spine Journal. 2004;4:106-115.....	7
8. Grunnesjo MI, Bogefeldt JP, Svardsudd KF et al. A randomized controlled clinical trial of stay active care versus manual therapy in addition to stay active care: functional variables and pain. Journal of Manipulative and Physiological Therapeutics. Sept 2004;27(7):431-441.....	8
9. Niemisto L, Lahtinen-Suopanki T, Rissanen P et al. A randomized trial of combined manipulation, stabilizing exercises and physician consultation compared to physician consultation alone for chronic low back pain. Spine. 2003;19:2185-2191.....	8
10. Over the last five years... ..	9
Summary	10
Bibliography	11
Appendix 1	12

Background.

A leaflet entitled 'Physiotherapy briefings for physicians. Low Back Pain Study Sparks Discussions' was provided in the November 2004 issue 46 number 9 of the BC Medical Journal. In this leaflet the Physiotherapy Association of British Columbia and Saskatchewan Physiotherapy Association disputed the way in which the results of a randomized controlled trial of physiotherapy compared with advice for low back pain was reported by the media (Frost H et al).

In late November 2004, the EBPG received a request from the Health Care Services group to investigate if the statements made in this leaflet were consistent with the supporting articles provided by the Associations and listed in the leaflet. As such, the purpose of this review is to conduct a critical appraisal of these articles and provide comments based on these appraisals.

Materials and methods.

Nine articles, including the leaflet, were retrieved and appraised according to standard methodology on critical appraisal of primary research (based on study design) and systematic review. Comments on the effectiveness of physiotherapy in treating low back pain is also provided by summarizing the results on the current primary research and systematic reviews on the topic.

Results.

1. Frost H, Lamb SE, Doll HA et al. Randomized controlled trial of physiotherapy compared with advice for low back pain. *BMJ*. Sept 17 2004;329:708-712.

This is a multi-center double blind randomized controlled trial investigating the effectiveness of advice to remain active + routine physiotherapy (up to 5x) compared with advice to remain active only from physiotherapist in a mixture of acute, subacute and chronic low back pain (LBP) patients referred by GPs and specialists. *Assuming that advice to remain active given by all participating physiotherapists was equal, this trial can be seen as looking into the incremental effect of adding routine physiotherapy treatments regardless of the methods being employed by different participating physiotherapists.* Randomization and blinding were clearly defined in this study.

Power calculation and statistical analysis were clearly stated and conducted. Intention to treat analysis was employed; the last known value carried forward is a recognized method. Sample size was adequate. The authors stated that the primary outcome considered in this study was the Oswestry disability index at 12 months. Further, the authors also stated that clinically meaningful differences were defined as a 4% difference in Oswestry score between groups. *Given this statement on primary outcome, other*

statistical analysis (outside of 12 months) should be seen as 'interim analysis'. As such, p-values should be adjusted to take into account the number of statistical analysis conducted in order to conclude that such hypothesis testing was statistically significant. Further, the a priori statement regarding criteria for 'clinically significant' as $\geq 4\%$ in Oswestry score difference should be accepted as that difference below 4% between advice and therapy group as clinically not significant.

The results of this study were clearly described, statistical analysis and hypothesis testing was done whenever appropriate. It should be noted that:

- The patient selection flow was clearly described. *This description is important in order to assess the possibility/impact of selection bias.*
- There was an imbalance in the baseline characteristics between groups. Namely that the advice only group contained a higher proportion of smokers and men. *The difference in the distribution of LBP between smokers and non smokers as well as between males and females is well recognized in the literature. It should be noted that the authors took these differences into account in their repeated measures analysis of covariance as reported in Table 2 and 3.*
- The authors also reported the difference between 12 month responders and non responders. *This is an exercise to assess the possibility of selection bias. Even though there were some differences noted between responders and non-responders, its effect toward the primary outcome assessment at 12 months was most likely very small, if any. Further, the authors conducted an intention to treat analysis (to preserve randomization) and last observation carried forward (to uniformly mitigate the problem of missing data). In terms of data analysis, these practices are appropriate and enhance the validity of the study results.*
- Outcomes were reported appropriately and according to the stated objectives. *It should be noted that none of the p-values reported in the repeated measures analysis (Table 2 and 3) were < 0.05 (standard accepted criteria to assess statistical significance) and the mean difference in 95% confidence interval, with the exception on some SF-36 domains as reported in Table 3, encompassed 0 (meaning the difference observed could be coincidence instead of true difference since it contained 0 or no different).*
- Some individual SF-36 domain analyses of covariance, as reported in Table 3, showed p values of < 0.05 . These domains included bodily pain at 12 months ($p=0.035$), social functioning at 6 months ($p=0.03$) and mental health at 2 months ($p=0.002$). *It is tempting to conclude that these differences were statistically significant. However, it is not true due to the fact that the size of p-value necessary to judge statistical significance needs to be adjusted to the number of statistical tests being conducted (in Table 3 alone, there were 24 test reported). Further, SF-36 was not the primary outcome.*

- Patient perceived benefits were also reported in Table 4. All differences reported in this table had p-values < 0.05. *It should be noted that this outcome was not the primary outcome. Further, the rule that $p < 0.05$ as statistically significant did not apply due to the problem of multiple comparison as previously noted.*

EBPG comments: Overall this is a well designed, well conducted and well reported study. The authors appropriately concluded that '...routine physiotherapy seemed to be no more effective than one session of assessment and advice from a physiotherapist...'. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence.

2. Physiotherapy briefing for physicians. Low Back Pain Study Sparks Discussions. Leaflet enclosure with the BCMJ. Fall 2004.

- The authors of this article stated that '...The results showed that after two and six months the patients who received treatment from physiotherapist reported that they were feeling better and stated a return to greater function. Those patients who received treatment from a physiotherapist at the earliest point saw a quicker return to function....'

This statement is somewhat misleading. The 2 and 6 months outcomes, as stated previously, were not the primary outcome. Further, the results did not show any statistical difference once multiple comparisons was taken into account, as discussed previously. In the original study, the authors did not look into the effect of timing of the treatment delivery nor there were data providing such conclusions.

- This article also stated that '.....Physiotherapy works. Both groups benefited from the intervention of physiotherapy...'. *This statement is somewhat misleading. Even though both interventions were delivered by physiotherapists, the original study by Frost et al⁽¹⁾ looked into the incremental effect of adding routine physiotherapy treatments over giving advice to stay active only. In order to be able to say that 'physiotherapy works', one would simply compare physiotherapy with no physiotherapy (can be with no treatment or placebo or other treatment). As well, the original BMJ article was not designed to assess the impact of the 'messengers' (i.e. physiotherapist in this case).*
- Two other points related to 'early treatment works' and 'early treatment helps a return to active life' were also mentioned in this BCMJ article. *However, this statement was not supported by the original study⁽¹⁾ reported in the BMJ.*

EBPG comments: There are some misleading statements in this 'briefing'.

3. Moseley L. Combined physiotherapy and education is efficacious for chronic low back pain. Australian Journal of Physiotherapy. 2002;48:297-302.

This is a single blinded randomized controlled trial investigating the effectiveness of 2x/week for 4 weeks physiotherapy + trunk muscle training + 1 hour education/week for 4 weeks v.s. medical management by GPs in treating chronic low back pain patients. Randomization was done through coin tossing. Outcome measures include the Roland Morris Disability Questionnaire (RMD) and 0-10 numerical rating scale for pain (NRS). The authors stated, a priori, that clinically significant change in the NRS and RMD were set at 2 and 4 points, respectively. Outcome measurements were done at week 4 and 12 months post intervention. At 12 months additional outcomes regarding number of health care professional consultations was also included. Inclusion and exclusion criteria were given. Sixty two subjects, identified through advertisement, responded and 5 were excluded from the study. Twenty nine and 28 subjects were assigned to physiotherapy and control (GP treatment) groups, respectively. Five and 3 subjects from the physiotherapy and control groups, respectively, dropped out from week 4 assessment. At 12 months follow-up, there were 19 subjects on each group. Statistical analysis was discussed adequately. Analysis was done by intention to treat.

- *It should be commended that the authors provided a flow chart (Figure 1) of subject recruitment. However the authors did not provide the total numbers of people presenting to the clinics and analyzed the differences between those volunteered and not volunteered. This type of analysis is necessary to assess the possibility of selection bias. Further, the authors did not provide information on the recruitment time frame.*
- *At 4 weeks follow-up, the authors presented RMD and NRS as outcome measurement tools. At 12 months follow-up, the authors added frequency of health care professional consultation as an outcome measurement. The authors did not state which of these tools was considered as the primary outcome measure. As such, there was no sample size calculation stated.*
- *Even though the authors discussed the statistical methods employed in this paper, the discussion is somewhat misleading with regard to the adjustment of p-value. In this paper, the authors adjusted the p-value based on the number of analyses done on each of the outcome variables, i.e. twice at 4 weeks and 12 months. However, there were, at least, 2 outcome variables measured and reported in this paper. As such, the p-value should have been adjusted further by a factor of 2. Thus, (at least) a p-value of 0.0125 should have been employed as a threshold to judge statistical significance.*
- *The authors stated the adoption of intention to treat analysis which should be commended. However, it was not clear nor was it stated the method being employed to handle missing data, given the fact that some subjects dropped out at the 4 weeks and 12 month follow-up periods.*

- *At 4 weeks follow-up, the difference in mean improvement (95% CI) for NRS (10 point scale) and RMD (18 point scale) were 1.5 (0.7 – 2.3) and 3.9 (2.0 – 5.8), respectively. At 12 months the mean different were 1.9 (1 – 2.8) and 3.9 (2.3 – 5.8) for NRS and RMD, respectively. The mean improvement on these measurements was hardly clinically significant (as stated a priori by the authors). It should be noted that the 95% confidence interval on the NRS encompassed 0.*

EBPG comments: There were flaws in the design, conduct and reporting of this randomized controlled trial. As such, the conclusions of this article should be treated with caution. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence.

4. Saal JA and Saal JS. Nonoperative treatment of herniated lumbar interval disc with radiculopathy. An outcome study. Spine. 1989;14(3):433-437.

The authors reported this study as a 'retrospective cohort study' (level 2 evidence). In fact this study is a case series (level 4 evidence). Regardless of the outcome, case series types of study (on which there is no control group), is only valid as a 'hypothesis generator' type of study. These types of studies are not appropriate for investigating the effectiveness of a particular treatment.

5. Hurwitz EL, Morgenstern H, Harber P et al. A randomized trial of medical care with and without physical therapy and chiropractic care with and without physical modalities for patients with low back pain: 6 month follow-up outcomes from the UCLA Low Back Pain Study. Spine. 2002;27:2193-2204.

This is a non-blinded, randomized controlled trial in patients with various spectrums of low back pain (acute – chronic). In this study, the outcomes of four treatment modalities, including medical care only, chiropractic care only, medical care with physical therapy and chiropractic care with physical modalities, were compared at 2, 6 and 24 weeks follow-up. The outcomes measured in this study included the 24 item Roland Morris Disability Questionnaire and changes in the most severe and average of low back pain intensity within the past week prior to follow-up. Differences between treatment groups of ≥ 2 points on the pain scale and ≥ 3 on the Roland Morris score was considered to be clinically significant. It should be noted that in this study third party liability or workers' compensation patients were excluded. This study was well designed, analyzed and reported. However:

- *There was a possibility of selection bias considering that > 50% of the initially eligible patients refused to participate and there was no information on the characteristics of these potential participants.*

- *About one third of patients assigned to the medical + physical therapy group did not have any physical therapy visits. These patients were still included in the medical + physical therapy group due to the intention to treat analysis. As such, the outcome of the medical + physical therapy group may be underestimated.*
- *The authors mentioned at least 2 primary outcomes that they would like to explore across 2 different time periods. As such, there was the possibility of multiple comparisons that were not addressed. Traditionally, this problem has been addressed through the adjustment of p-value by the number of statistical tests being conducted. However, statistical/epidemiological research suggests there are at least two different ways of thinking regarding the essence of statistical significance. In particular, one group of thought prefers the use of confidence intervals instead of p-values due to the essence of information provided by confidence intervals. It seems that in this paper, the authors adopted 95% confidence intervals as a measure of statistical significance. Regardless of this argument, with the exception of 0 – 6 months difference between medical care + physical therapy and medical care only, the 95% confidence intervals of various measurements presented in Table 4 and 5 contained 0 or 1. With regard to the difference between medical care + physical therapy and medical care only at 0 – 6 months follow-up, this difference is very small and not likely to be meaningful.*

EBPG comments: In general, this study was well designed, analyzed and presented. However, in this paper, the authors may under estimate the effect of physical therapy in patients with low back pain. As such, this study did not provide conclusive evidence to accept or refute the effectiveness of physical therapy. It should be noted that there was also the possibility of volunteer bias within the study. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence.

6. Moffett JK, Torgerson D, Bell-Syer S et al. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs and preferences. BMJ. 31 July 1999;319:279-283.

This is a single blind, randomized controlled trial examining the impact of a standard exercise program delivered by a physiotherapist against standard care delivered by medical practitioners in treating sub acute and chronic non specific low back pain patients. The primary outcome measures in this study were the Roland Morris Disability 24 items score and the Aberdeen Back Pain scale. Outcome measures were conducted at 6 weeks and 6 and 12 months. The study was well designed, analyzed and reported. The pitfalls of smaller sample size was discussed appropriately. The smaller number within the

sample size may actually affect the ability of this study to detect smaller differences between treatment groups as it was observed in the outcomes. However;

- *The study did not report the primary outcome and the primary measurement time period on which the sample size was calculated, thus the need to adjust for the p-value in assessing statistical significance.*
- *> 50% of eligible participants were not randomized and the differences between those who volunteered and those who were not was not presented. As such there was a possibility of selection/volunteer bias in this study.*
- *The authors concluded that exercise class was more clinically effective than traditional GP management. This conclusion is somewhat misleading. With the (possible) exception of the Aberdeen Back Pain scale at 1 year, the differences in Roland Morris and Aberdeen Back Pain (as primary outcome measurements) scores between intervention and control groups were not statistically significant (or the 95% confidence interval include 0) at 6 weeks and 6 and 12 months. However, the fact that statistical significance, as defined by its p-value, was not detected may be due to insufficient power of the study.*

EBPG comments: In general, this study was well designed, analyzed and presented. However, the conclusion provided in this paper may be misleading. At best, it can be concluded that exercise as delivered by physiotherapist, provides small beneficial effect, as measured by the Roland Morris and the Aberdeen Back Pain scales, at 6 weeks, 6 and 12 months. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence

7. Rainville J, Hartgan C, Martinez E et al. Exercise as a treatment for chronic low back pain. Spine Journal. 2004;4:106-115.

This is an 'expert' review article examining key aspects about the safety and efficacy of exercise in treating chronic low back pain. In this review, the authors searched Medline medical database by employing several keywords and provided narrative synthesizes of the primary research articles that the authors thought to be relevant (no specific criteria being given). The authors did not provide any critical appraisal data on these primary research papers. The authors concluded that exercise is safe for individuals with back pain because it does not increase the risk of future injuries or work absence. The authors also concluded that exercise improves impairments in back flexibility and strength.

EBPG comments: Even though the authors provided some information on the methodology of this paper, the paper is still considered as an expert based review article instead

of a systematic review. In the grades of quality of evidence (Appendix 1), this study provides level 5 evidence

8. **Grunnesjo MI, Bogefeldt JP, Svardsudd KF et al. A randomized controlled clinical trial of stay active care versus manual therapy in addition to stay active care: functional variables and pain. Journal of Manipulative and Physiological Therapeutics. Sept 2004;27(7):431-441.**

This is a single blinded, randomized controlled trial investigating the impact of manual therapy and stay active concept in a factorial study design among acute and subacute low back pain patients.

- *Given the treatment 'factors' in this study, i.e. stay active and manual therapy (it is also called a 2 x 2 design), there should be 4 study groups presented in this paper. The authors only reported data on 2 groups.*
- *The authors did not provide a clear explanation on the methodology of the research.*
- *The study lacked in crucial basic information on the conduct of clinical trials such as number of patients in each cell of the group studied, hypothesis, power calculation and statistical methods employed in the analysis (e.g. it was not appropriate to do statistical tests on the baseline characteristics of the participants, the authors did not take into account the nature of multiple statistical tests being conducted).*

EBPG comments:

This is a confusing and not well reported (probably not well executed as well) trial. There were many flaws detected in this paper that most likely will jeopardize its conclusions. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence.

9. **Niemisto L, Lahtinen-Suopanki T, Rissanen P et al. A randomized trial of combined manipulation, stabilizing exercises and physician consultation compared to physician consultation alone for chronic low back pain. Spine. 2003;19:2185-2191.**

This is a single blind, randomized controlled trial investigating the impact of manipulative therapy + exercise + education (treatment group) versus physician consultation + education (control group) among chronic low back pain patients. The study was well designed, analyzed and presented. The hypothesis was partially stated and the objectives of the study were clearly stated. Patient recruitment, including the inclusion and exclusion criteria, was well described and accounted. The interventions and the intended outcome measurements, including the costing, were well described. Randomization procedure, power calculation and statistical analysis were well explained. The

results of the study were presented according to the standard practice of well designed randomized controlled trials.

- *The EBPG rate this primary research paper as one of the best in this series. All the necessary attributes of good conduct in clinical trials were reported by the authors.*

EBPG comments: In general, this study was well designed, analyzed and presented. It should be noted that the difference between treatment and control groups with regard to their visual analog (VAS) pain score and Oswestry Disability score at 12 months were very small (Table 2) in favor of the treatment group. However, it should also be noted that the treatment group had less cost at 12 months compared to the control group. In the grades of quality of evidence (Appendix 1), this study provides level 1 evidence.

10. Over the last five years (see some bibliography cited below), numerous systematic reviews (level 1 evidence) have been published with regard to treatment for chronic or acute low back pain patients. There were only 2 systematic reviews (published in 1999) specifically assessing the impact of physiotherapy in treating low back pain and neck pain. However, most of the published systematic reviews also examined the effectiveness of physiotherapy compared to other treatments. The latest systematic review, reported by Bogduk N, showed that physiotherapy was not more effective than manipulative therapy, care by general practitioners or back school.

Summary.

- The article by Frost H, Lamb SE, Doll HA et al. (number 1) is a well designed, well conducted and well reported study. The authors appropriately concluded that '..routine physiotherapy seemed to be no more effective than one session of assessment and advice from a physiotherapist..'
- With the exception of articles no 1 and 9, above, the majority of articles presented by the two provincial physiotherapy associations in the 'leaflet' of the November 2004 BC Medical Journal are of much lower quality and their conclusions are not always convincing.

Bibliography.

- Harms-Ringdahl K, Holmstrom E, Jonsson T et al.(1999). Evidence-based physiotherapy for patients with low back pain. Summary and conclusions. The Swedish Council on Technology Assessment in Health Care.
- Nachemson AL, Jonsson E (eds.) (2000). Neck and Back Pain. The scientific evidence of causes, diagnosis and treatment. Lippincott Williams and Wilkins. Philadelphia.
- .(2000). Back Pain Neck Pain. An evidence based review. Summary and conclusions. The Swedish Council on Technology Assessment in Health Care.
- .(2003). Therapeutic Exercise. Health Council of the Netherlands. Executive summary. Publication No. 2003/22
- MW Tulder, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. *The Cochrane Database of Systematic Review 2000. Issue 2.*
- Bogduk N. Management of chronic low back pain. *Medical Journal of Australia.* 2004;180:79-83.
- Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works? *Pain.* 2004;107(1-2):176-190.
- Maher CG. Effective physical treatment for chronic low back pain. *Orthopaedics Clinics of North America.* Jan 2004;35(1):57-64.
- Hagen KB, Hilde G, Jamtvedt G et al. The Cochrane Review of bed rest for acute low back pain and sciatica. *Spine* 2000;25:2932-2939.
- Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic non specific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine.* 1997;22:2128-2156.
- Pengel HM, Maher CG, Refshauge KM. Systematic review of conservative interventions for subacute low back pain. *Clinical Rehabilitation.* 2002;16:811-820.
- Torill H, Veito MH, Eiksen HR. Low back pain interventions at the workplace: a systematic literature review. *Occupational Medicine.* 2004;54:3-13.
- Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low back pain. *The Cochrane Database Systematic Review 2004. Issue 4.*

Appendix 1.

Workers' Compensation Board of BC - Evidence-based Practice group. Grades of quality of evidence (adapted from 1,2,3,4).

1	Evidence from at least 1 properly randomized controlled trial (RCT) or systematic reviews of RCTs.
2	Evidence from well-designed controlled trials without randomization or systematic reviews of observational studies.
3	Evidence from well-designed cohort or case-control analytic studies, preferably from more than 1 centre or research group.
4	Evidence from comparisons between times or places with or without the intervention. Dramatic results in uncontrolled experiments could also be included here.
5	Opinions of respected authorities, based on clinical experience, descriptive studies or reports of expert committees.

Reference.

1. Canadian Task Force on the Periodic Health Examination: The periodic health examination. CMAJ. 1979;121:1193-1254.
2. Houston TP, Elster AB, Davis RM et al. The US Preventive Services Task Force Guide to Clinical Preventive Services, Second Edition. AMA Council on Scientific Affairs. American Journal of Preventive Medicine. May 1998;14(4):374-376.
3. Scottish Intercollegiate Guidelines Network (2001). SIGN 50: a guideline developers' handbook. SIGN. Edinburgh.
4. Canadian Task Force on Preventive Health Care. New grades for recommendations from the Canadian Task Force on Preventive Health Care. CMAJ. Aug 5, 2003;169(3):207-208