

臨床試験

**Dose-response and efficacy of spinal manipulation for care of chronic low back pain: a randomized controlled trial**

慢性腰痛治療のための脊椎徒手整復の用量反応と有効性：無作為化比較対照試験

*Mitchell Haas, DC<sup>a</sup>, Darcy Lavre, ND<sup>a</sup>, David Peterson, DC<sup>b</sup>, Naya Polissar, PhD, Monica Neradile, MS<sup>c</sup>*

<sup>a</sup>Center for Outcomes Studies, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA <sup>b</sup>Division of Chiropractic Sciences, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA

<sup>c</sup>The Mountain-Whisper-Light Statistics, 1827 23rd Ave. E., Seattle, WA 98122, USA □

受付：2012年11月2日、修正：2013年7月17日、受理：2013年7月22日



## 情報収集事業

療術に関する情報収集事業は、諸外国がカイロプラクティックなど手技療法の施術が有効だと考えている疾患・患者像と、有効でないと考えている疾患・患者像、及びそれらを前提とした患者の安全確保対策などを比較調査するために、療術に関する資料、情報の収集を行うもので、平成 27 年度は下記の実施計画に基づき、

文献（THE SPINE JOURNAL）から

「慢性下背部痛治療のための脊椎徒手整復の用量反応と有効性：無作為化比較対照試験」を翻訳し発刊した。

《翻訳文中の用語解説》

- 1、慢性下背部痛（慢性腰痛 cLBP）
- 2、腰痛（LBP）
- 3、脊椎徒手整復療法（SMT 高速矯正 = スラスト）
- 4、用手療法／用手操作（SMT）
- 5、治療医師（カイロプラクター）

平成 28 年 3 月 一般財団法人 全国療術研究財団

## カイロプラクティックの海外文献調査研究実施計画（平成 27 年度）

◎研究者及び事務担当者

研究者

宇都宮 光 明 一般財団法人全国療術研究財団 常務理事

東京慈恵会医科大学医学部医学科 昭和 36 年 3 月卒業 医学博士

松 本 徳太郎 一般財団法人全国療術研究財団 理事

アメリカ合衆国 ロサンゼルス・カイロプラクティック大学

カイロプラクティック科 昭和 48 年 9 月卒業 D.C.

事務担当者

佐 賀 哲 夫 一般財団法人全国療術研究財団 事務局長

◎分担した研究事業の概要

研究者名及び分担した研究項目

宇都宮光明・・・医学的見地からの文献研究の総括

松本徳太郎・・・カイロプラクティックの海外文献の調査と検討

佐賀 哲夫・・・海外文献の原文入手と著作権等の調査と検討

（衆議院の外国語センターと協議し調査依頼）



## カイロプラクティックの海外文献の翻訳について (THE SPINE JOURNAL)

D.C. 松本徳太郎

この研究の目的として書かれているように、これまで腰痛に対するカイロプラクティックのアジャストメント（矯正）の有効性及び安全性についての報告は多くあっても、最適な施療回数と症状の軽減の関係を明らかにしたものはなかったように思う。それも、100人の被検者数での最大52週間、施療回数18回という長期にわたる研究や、問診による疼痛の日数・機能障害のあった日数や不快感について等々、臨床に則した方法での結果判定は高い評価が出来るものとする。

また、カイロプラクティックと、軽いマッサージによる有効性との比較検討も興味深かった。

施療前に、静的（スタティック）／動的（モーション）触診（パルペーション）により、脊柱の彎曲・左右の対称性・脊椎の可動域・軟部組織の状態・圧痛（CARTT）を調べ、適応性を確認してからマニピュレーションと呼ぶ脊椎矯正を行っている。この施療法は、腰痛施療で標準的に用いる側臥位での腰椎および胸腰椎移行部の高速矯正（ランバーロール）であり、この方法はわが国でも最もオーソドックスなテクニックであろう。

被検者の中には施療への適応を認められなかった者もあり、その場合は高速矯正は行わず、また、急性痛の例では緩やかな操作（モビリゼーション）や軽いスラストを用いたとある。これも、我々が臨床における事故防止の観点から最重要視している安全安心の注意点と合致している。

施療の効果については、6週間で12回の施療が慢性腰痛に対して最も好ましい改善をもたらしたとある。6週間で12回となると、週2回の施療頻度になるが、この回数は研究に協力したことによる施療費の負担がないことで、この回数まで施療を続けたのかもしれない。

しかし現実の臨床においては、この頻度で通院してくれる患者は少ないのではないかと思う。この点は日米を問わず、共通しているのではないだろうか。この点は今少し深く知りたいと思った点である。

この研究報告書を読んだ時に、アメリカのカイロプラクティック大学での授業で、全患者に「どの患者も必ず10回受療させる方法」という講義を受けたことを思い出した。患者に対し責任を持って施療するうえで、受療回数は大事なことなのだと改めて思った。この回数を基に、腰痛を軽減させるベストな回数を知ることが出来たのは、今後のカイロ手技教育に大きな参考になった。



臨床試験

## **Dose-response and efficacy of spinal manipulation for care of chronic low back pain: a randomized controlled trial**

### **慢性下背部痛治療のための脊椎徒手整復の用量反応と有効性： 無作為化比較対照試験**

Mitchell Haas, DC<sup>a,\*</sup>, Darcy Vavrek, ND<sup>a</sup>, David Peterson, DC<sup>b</sup>, Nayak Polissar, PhD<sup>c</sup>, Moni B. Neradilek, MS<sup>c</sup>

<sup>a</sup>Center for Outcomes Studies, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA <sup>b</sup>Division of Chiropractic Sciences, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA

<sup>c</sup>The Mountain-Whisper-Light Statistics, 1827 23rd Ave. E., Seattle, WA 98122, USA

受付：2012年11月26日、修正：2013年7月17日、受理：2013年7月22日

#### **要約**

**背景状況：**何らかの状態に対する脊椎徒手整復によるケアの最適な通院回数に関する大規模な臨床試験は、これまでに行われていない。

**目的：**脊椎徒手整復のためのカイロプラクティックへの通院と慢性下背部痛（cLBP）の転帰との間の用量反応関係を明らかにすること。また、脊椎徒手整復の有効性を、軽マッサージによるコントロールと比較して検討すること。

**研究デザイン／設定：**治療に基づく無作為化比較対照試験

**患者サンプル：**慢性下背部痛（cLBP）がある参加者 400 人

**結果判定法：**cLBP の主要転帰は、12 週目と 24 週目の主要評価項目で評価した、100 ポイントでの修正 Von Korff 疼痛強度および機能的障害尺度とした。副次転帰は、疼痛があった日数、機能的障害があった日数、疼痛による不快感、全体として感じられた改善、薬剤の使用、および全体的健康状態を含めた。

**方法：**cLBP がある 100 人の参加者を、カイロプラクターによる脊椎徒手整復を 0、6、12、または 18 セッションを受ける、4 段階の用量レベルのそれぞれに無作為化した。参加者は、週に 3 回、6 週間の治療を受けた。脊椎徒手整復が割り当てられなかったセッションでは、集中的な軽マッサージ・コントロールを受けた。共変量で補正した直線的用量効果と、非脊椎徒手整復コントロール群との比較を 6、12、18、24、39、および 52 週目に評価した。

**結果：**主要転帰について、脊椎徒手整復群における疼痛と障害の平均の改善は、12 週目までで 20 ポイントで、52 週間まで維持可能であった。直線的な用量反応効果は小さく、両方の変量について、12 週目と 52 週目で、6 回の脊椎徒手整復セッションあたり約 2 ポイントに達するのみであった（ $p < .025$ ）。12 週目に、12 セッションで非脊椎徒手整復コントロールとの最大の差（疼

痛ポイント 8.6、および障害ポイント 7.6、 $p<.025$ ) が認められたが、24 週目では、差は無視できる程度であり、52 週目では通院 18 回で最大の集団差が認められた (疼痛ポイント 5.9、および障害ポイント 8.8、 $p<.025$ )。

**結論：** 脊椎徒手整復のための通院回数は、カイロプラクターへの実地通院が 18 回を越えた場合には、cLBP の転帰に対する効果はわずかであった。全体として、12 回の通院で最も好ましい結果が得られたが、他の用量レベルと比べてあまり明確な違いは認められなかった。

キーワード：慢性下背部痛 (Chronic low back pain)、用量反応 (Dose-response)、脊椎徒手整復 (Spinal manipulation)、カイロプラクティック (Chiropractic)、無作為化比較対照試験 (Randomized controlled trial)

-----  
FDA への機器/医薬品申請状況：該当なし

著者に関する開示：MH：助成金：NIH: NCCAM (I、施設に直接支払い済み)。DV：助成金：NIH: NCCAM (I、施設に直接支払い済み) DP:助成金:NIH: NCCAM (I、施設に直接支払い済み) NP：データモニタリング委員会や統計解析、エンドポイント委員会などの審査活動への参加費用：ウェスタン・ステーツ大学 (B) MBN：データモニタリング委員会や統計解析、エンドポイント委員会などの審査活動への参加費用：ウェスタン・ステーツ大学 (B)。

開示に関する凡例は、[www. TheSpineJournalOnline.com](http://www.TheSpineJournalOnline.com) の「目次」で閲覧可能である。

本試験は、[ClinicalTrials.gov](http://ClinicalTrials.gov) に NCT00376350 にて登録された。

\* 連絡先の著者：Center for Outcomes Studies, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA. 電話：(503) 251-5728、ファックス：(503) 251-2832。

E メールアドレス：haasmitch@comcast.net (M. Haas)

## エビデンスと方法

### 状況

著者らは、コントロールの役割を果たす軽マッサージと比較した、慢性下背部痛に対する脊椎徒手整復の影響を評価した。

### 貢献

適切に実施された本 RCT では、軽マッサージと比較して、18 回の通院による脊椎徒手整復が転帰にわずかな改善をもたらした過去とが明らかになったが、この改善は臨床的に有意でない可能性がある。著者らは、脊椎徒手整復の最良の治療効果は、対象と比較して 12 回のセッションで得られることも明らかにしたが、18 回のセッションではそれ以上の利益は得られなかった。



## 意義

本研究は、これまでにほとんど示されていない、いくつかのガイドランスを提供している。12回のカイロプラクティックのセッションは妥当であり、この評価項目では、脊椎徒手整復は軽マッサージよりもわずかに良好であるかもしれないが、24週目ではそうとも言えない。重要なこととしては、本研究が診療に基づく無作為化試験の研究デザインの妥当なモデルとしての役割を果たしている点が重要である。

## — 編集

---

## 序論

下背部痛（LBP）は高頻度に見られる、医療費のかかる疾患であることが知られており<sup>[1,2]</sup>、カイロプラクターが行う脊椎徒手整復の大部分の施術がLBPを対象に行われており<sup>[3]</sup>、カイロプラクターが米国におけるLBPの治療の大部分を行っている<sup>[4]</sup>。そのため、特に慢性下背部痛（cLBP）に対する脊椎徒手整復療法（SMT）の最適量を定めることが重要である<sup>[5]</sup>。

cLBPに対するSMT／カイロプラクティックによるケアの期間と頻度に関する推奨は、幅広く様々であり、長い間、臨床的な経験と評価に基づいて行われてきた<sup>[6]</sup>。1990年代始め、学際的RANDパネルは、その評価がエビデンスに基づく正式コンセンサスに至るにはあまりに多様である（2~24週間のケア、週1回~5回の通院）ことを見いだしたが、パネルメンバーは概して、典型的な患者は週3回の通院で4~6週間以内に改善すると期待した<sup>[7]</sup>。対照的に、全員がカイロプラクティックで構成されるRANDの専門家パネルは、14週間にわたる30回の通院を推奨した<sup>[8]</sup>。Shekelleら<sup>[3]</sup>が発表したカイロプラクティックによるケアの研究では、1回から19回の範囲の通院を示した。その後、Nyiendoら<sup>[9]</sup>は、526人の非特異的cLBP患者の診療ベースのコホートにおいて、平均6.7回の通院（標準偏差[SD] 57.5回、範囲51~56回）を明らかにした。

今日までに、cLBPのケアにおけるSMTの有効性とその役割に関するコンセンサスは得られていない。いくつかの系統的レビューは、SMTを支持する質に関するエビデンスを報告しており<sup>[10,11]</sup>、一方で、最新のコクランレビューを含む他のレビューは、SMTが他の介入と同程度であることを明らかにした<sup>[12]</sup>。系統的レビューの結果は、メタ解析か最良のエビデンスの統合であるかにかかわらず、レビューに含められた治験で用いられたケアの量に左右される可能性がある。研究者は、用量反応研究から實際上、SMTのセッション提供回数に関して何のエビデンスを得ていない。

ケアの期間と頻度に関するエビデンスが不足しているため、我々はSMT用量反応性を評価する初の試験的無作為化試験を実施した（n=72）<sup>[5]</sup>。我々は、通院回数が多いほど結果が好ましいという、カイロプラクターへの通院数（週に1~4回の通院で3週間）と短期間の疼痛と障害緩和との間の臨床的に重要な関連性を見いだした。続いて、我々は、SMTによる最適なcLBP

ケアの特定と比較有効性研究のデザインに関する情報の提供を目的とした、初の大規模用量反応試験である、現在進行中の5カ年の研究を実施した。我々はSMTと実地コントロールの間に差は無いという仮説の検証を行い、SMTの用量レベルの有効性も評価を行った。注意（通院数）と患者への接触、履歴、背景をコントロールするための軽マッサージの効果を上回る、転帰に対するSMT独自の寄与を評価した<sup>[13]</sup>。

## 方法

### デザイン

前向き非盲検無作為化比較対照研究において、非特異的cLBPがある400人の参加者を、カイロプラクターによる0、6、12、または18回のSMTセッションを受ける、4段階の用量レベルのそれぞれに無作為化した。全患者を週に3回、6週間の、18回の治療通院に割り当てた。脊椎徒手整復療法を割り当てた通院回数で実施し、施術者の注意と患者への接触をコントロールするために、非SMT通院では簡単な軽マッサージ・コントロールを実施した<sup>[14]</sup>。例えば、被験者は12回の通院分のSMTと6回の通院分の軽マッサージをカイロプラクターから受ける（図1）。追跡評価は、郵送による質問票、または盲検化した電話によるインタビューで無作為化後6、12、18、24、39、および52週目に実施した。主要転帰は、12週目と24週目のエンドポイントにおける自己報告による疼痛強度と機能的障害と予め規定した。

無作為化は、コンピューターで生成された、試験デザインに適応した割り当てを用いて実施し<sup>[15,16]</sup>、各群にわたって6種類のベースライン変量、すなわち疼痛と障害のスコア、年齢、性別、SMTとマッサージに対する相対的信頼、およびSMTまたはマッサージによるケアの履歴をバランスよく割り当てた。従って、試験群への割り当ては、収集したデータを無作為化直前に全データをコンピュータープログラムに入力を求めることで、試験関係者と参加者全員に伏せられた（疼痛、障害、および治療成功に対する信頼）。患者コーディネーターは、割り当てコンピュータープログラムにデータを入力した研究スタッフに電話で割り当て変量を求めた。その後、コーディネーターは患者のクリニックのファイルにケア方法を示した無印の密封した封筒を入れることで、参加者を各群に割り当てた。参加者と治療を行う医師は、無作為化後は介入について無作為化されなかった。但し、電話インタビューで一部の転帰結果を収集した患者コーディネーターには、試験を通じて治療群への割り当てを盲検化したままとした。本試験はウェスタン・ステーツ大学の施設内倫理委員会の承認（FWA 851）を受けた。治験はClinicalTrials.govでNCT00376350にて登録された。

### プロトコルの概要

参加者は、クレイグスリスト、郵便広告、および地方紙を通じて募集した。参加者には、本試験はcLBPに異なる2種類の用手療法の異なる組み合わせで構成される18回の通院の調査を

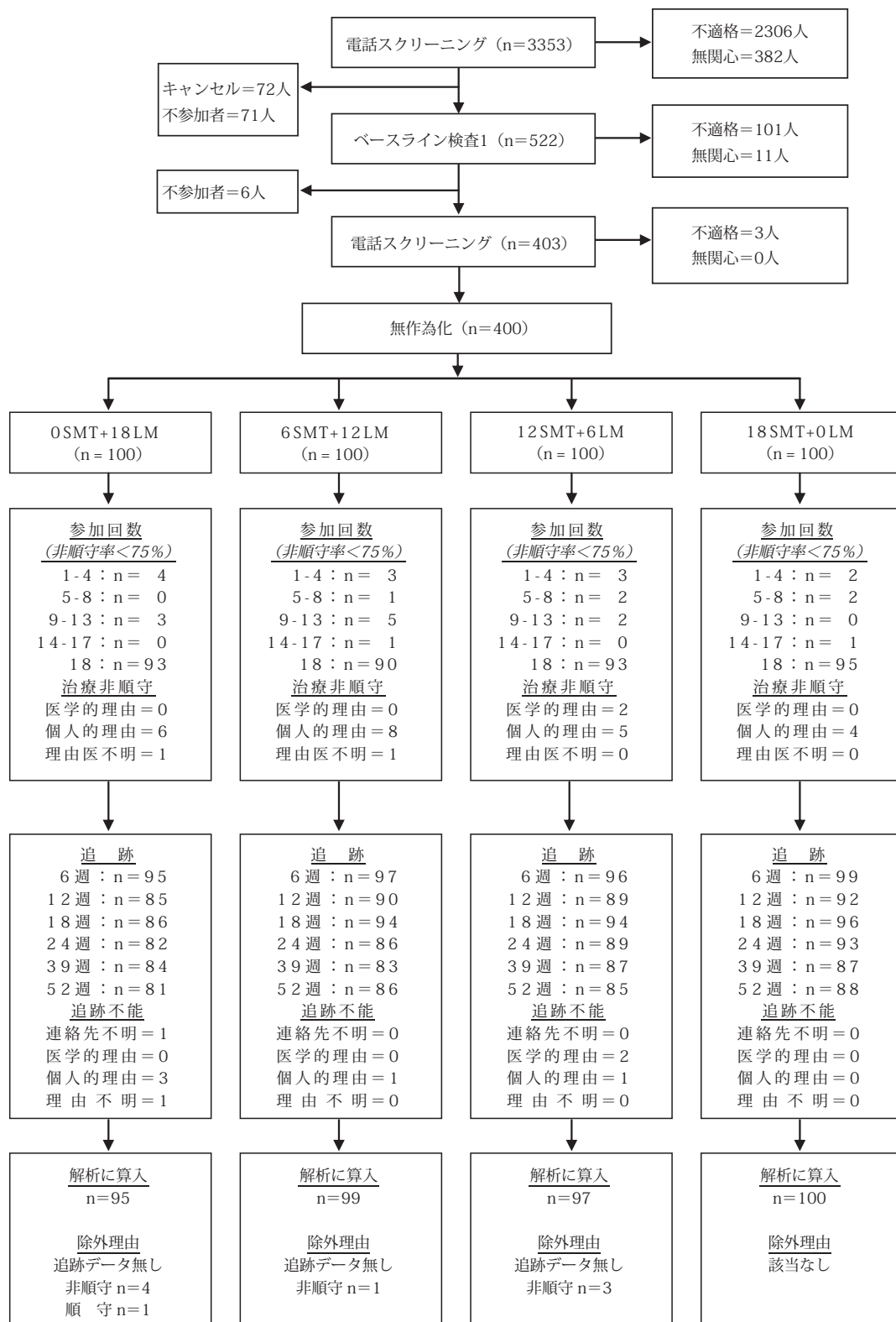


図1 試験フローチャート。全参加者を、18回の治療通院に割り当てた。参加者は、任意の1回の通院時に、脊椎徒手整復療法（SMT）か軽マッサージ・コントロール（LM）のいずれかを受けた。

行っていると知らせた。事前スクリーニングは、研究スタッフによる電話インタビューにより行った。最初のベースライン通院で、参加者はインフォームドコンセント用紙に署名し、ベースライン調査に記入した。その後、2人の有資格カイロプラクター(20年以上の経験がある教員)のうちの1人による大学の中央クリニックにて、履歴、身体検査、および腰部 X 線写真による適格性スクリーニング検査を実施した<sup>[17]</sup>。

適格とされた参加者に、試験ケアの実施に便利なクリニックを選ばせた。ケアは、ポートランド地域の9カ所のクリニックの4年から24年の経験がある12人の有資格カイロプラクターのうちの1人が提供した。治療を行うカイロプラクターは、著者らが知るカイロプラクターの技量に基づいて選ばれ、一部のカイロプラクターは我々の試験で過去に経験があるもの達であった。クリニックでの2回目のベースライン調査後、参加者に治療を行う医師による簡単な確認スクリーニング検査を実施した。その後、参加者を無作為化し、最初の治療を実施した。参加者には各治療通院、郵便による質問調査、および電話インタビュー毎に、報酬(10~20ドル)が与えられた。参加者は、インフォームドコンセント用紙に署名した。

#### 参加者

参加者は、18歳以上であり、現在、3カ月以上の期間にわたる機械的原因<sup>[19]</sup>によるcLBPの症状<sup>[18]</sup>の発現があることを必要条件とした。さらに、床効果を回避するために、過去6週間のうち30日にかなりのLBPがあり、また、最低LBPインデックスが100ポイント中25ポイントであったことを必要条件とした。過去90日以内に用手療法を受けた参加者または試験介入が禁忌とされる参加者<sup>[17,20]</sup>、ならびに活発な癌、脊椎病変、炎症、炎症性関節症、自己免疫疾患、抗凝固状態、神経変性疾患、膝関節より下への放散痛、器質性関連痛、妊娠、および身体傷害補償などの合併症がある参加者は除外した。

#### 介入

各通院治療は、カイロプラクティック治療と一致した、カイロプラクターによる15分間の治療とした<sup>[21]</sup>。参加者は、脊椎筋をリラックスさせるために、5分間の温パックを、その後、5分間のSMTまたはコントロール介入を施した。通院治療は、5分間の非常に低用量のパルス超音波(0.5ワット/cm<sup>2</sup>で20%の動作サイクル)を行って終了した。これは、治療に対する信頼性とケアの順守を強化するための擬似治療として用いた。<sup>[13]</sup>

脊椎徒手整復療法は、主にサイド・ポスチャー位置での、胸腰椎移行部の徒手スラスト(高速、低振幅)脊椎徒手整復で構成された<sup>[22]</sup>。実施する具体的な用手操作は、参加者の経過、自己報告と誘発による疼痛、脊椎可動域、および脊椎と脊椎周囲の軟部組織の触診など、参加者の継続的評価により、カイロプラクターが各通院時に決定した<sup>[17,22]</sup>。治療を行うカイロプラクターが適応を認められなかった場合には、用手操作は実施しなかった。腰椎の痛みの急性増悪の場

合には、機械的補助によるバネ仕掛けのテーブルの使用および部分的低速モビリゼーションなどの軽いスラスト用手操作が認められた<sup>[22]</sup>。

軽マッサージ・コントロールは、症状がある部位に集中した、腰部（腰部と下部胸椎）の傍脊柱筋群の5分間の穏やかな軽擦法と揉捏法で構成された<sup>[22,23]</sup>。使用したマッサージは、推奨される治療的マッサージ治療よりも穏やか、かつ短時間であった<sup>[21,24]</sup>。そのため、型通りの疑似介入ではなく、必要最低限度の介入で患者との接触のコントロールを行った。治療を実施するカイロプラクターは、施術者が与える治療の成功に対する期待のバランスを取るのに役立つよう、SMTとコントロール介入を等しい意欲で提供するようにも求めた。プロトコルの標準化と治療群間にわたる治療提供者の均衡は、年4回の研修を通じて維持し、診察室の観察と患者に対する電話インタビューによりモニターを行った<sup>[14,25]</sup>。

#### 転帰およびベースライン変量

本報告は、事前に規定した主要転帰である、Underwoodらが検証した自己報告に基づく修正Von Korff疼痛および障害尺度<sup>[26]</sup>を重視している。疼痛スコアは、3つの11段階の数値化尺度（今日の背部痛、過去4週間で最悪だった背部痛、および過去4週間の平均背部痛）の平均で、100ポイントの尺度に変換される。障害スコアも3つの尺度（日常的活動の障害、社会的活動と余暇活動、および作業能力（家の外または家事））の平均である。

副次転帰としては、疼痛による不快感<sup>[27]</sup>、12項目の身体的健康要素および精神的健康要素のサマリースコア<sup>[28]</sup>、EuroQolの健康状態視覚的アナログ尺度<sup>[29]</sup>、認識した疼痛と障害の改善、および疼痛や障害があった日数、およびの各要因ならびに投薬使用それぞれについての過去4週間の回数を含めた。追加のベースライン変量として、人口統計学的要因、恐怖回避思考質問票<sup>[30]</sup>、治療成功への信頼<sup>[14]</sup>、および以下の併発状態のリストのいずれか（関節炎、ぜんそくまたはアレルギー、胃腸障害、婦人科的問題、高血圧、またはその他の慢性疾患）<sup>[31]</sup>を含めた。

#### 統計解析

最初の割り当て群に含まれる各参加者により、治療意図に基づく解析を実施した。欠測データは線形補間により補完し、最後のデータを繰り越した。追跡データが無い場合、9人の参加者を解析から除外した。2つの主要転帰について、全欠測データを除外した感度解析を実施した。

事前に定めた主要解析は、SMT用量（傾き＝転帰増分／SMTセッション6回）の一次効果を明らかにするための、また各SMT群を非SMTコントロール群と比較（補正平均差）するための、回帰モデルで構成された。Zellner<sup>[32]</sup>による一見したところ無関係（同時）の回帰を用いて、個別のタイムポイントの転帰をモデル化した<sup>[33]</sup>。さらに、主要転帰のみについては、個人内のタイムポイント間の相互関係を用いた標準化推定式により、全ての追跡に渡る長期的影響をモデル化した<sup>[33]</sup>。傾きと群比較は、参加者の無作為化に用いた6つのベースライン調整変

量<sup>[15,16]</sup>、すなわち疼痛と障害の尺度、年齢、性別、SMTとマッサージに対する相対的信頼、およびSMTまたはマッサージによるケアの履歴で補正した。6つの調整変量に既に含まれていない場合には、転帰の評価基準のベースライン値を共変量として加えた。

事前に規定した副次的「応答者」解析では、2つの主要転帰を二分し、50%回復した参加者の割合を示した。次に、2項回帰を用いて、前述の解析を繰り返し、応答者の割合における傾きと群間差を明らかにした<sup>[33]</sup>。

標本サイズは、2つの主要転帰の100ポイント中10ポイントの群間の影響の検出に対して、両側検定で最低80%の検出力を持つよう、事前に決定した。10%の中断率を考慮した。2つの主要転帰を得るための補正に、有意水準として.025を用いた。我々の過去の研究と一致するよう、10段階による群間差の検出を選択した。<sup>[31,34]</sup>。すべての解析をStata 11.2 (StataCorp, College Station, TX, USA)を用いて実施した。

## 結果

参加者は、2007年3月から2010年5月までの間に登録され、2011年5月を最後の追跡の終了として1年間追跡を行った。割当は、クリニックと施術提供者にわたって等しく展開し、クリニックあたりのグループあたり平均で25% (SD=6%) および治療医師あたりのグループあたり25% (SD=12%) でグループを割り当てた。

図1の試験フローチャートは、全18回の試験通院に対して参加者が90%から95%参加し、ケアがしっかりと順守されたことを示している。18SMTセッションに割り当てられ、全18回のSMTに参加した4人の参加者は、合計5回の治療通院でSMTが不適切と見なされ、プロトコル通りの治療は控えた。3人は適応外であったため、1回の通院でSMTを控え、2人は急性増悪のため、1回の通院でモビリゼーションの施術を受けた。プロトコル逸脱が1件あり、患者が12回ではなく誤って13回SMT通院した。データ収集の順守は、全追跡タイムポイントの80%を上回った。9人の参加者が完全に追跡不能となった。

cLBPに対する投薬使用と試験外の施術提供者によるケアは、各タイムポイントで群間の釣り合いが取れていた。治療段階中、各治療群の参加者の93%から97%が、試験外の専門家の治療を控え、また94%から95%が処方薬を控えた。従って、各フォローアップ時に、およそ4分の3が外部の専門家の治療を受けなかったと報告し、最大群間差は参加者の4%から11%の範囲だった。また、各フォローアップ時に90%が処方薬を控え、最大群間差は1%から8%の範囲だった。非処方鎮痛薬は、グループ間で均衡がとれていた。2種類の介入の成功に対する信頼はほぼ同等であり、グループ間で均衡がとれていた。

ベースライン特性(表1)は、喫煙を除いてグループ間で均衡がとれており、解析に喫煙を含めた場合、効果量に実質的な変化は生じなかった。平均年齢は41.3歳で、大多数の参加者が非ヒスパニック系の白人だった。参加者の半数が、以下の特性、すなわち女性、大学学位、併

表1 基本特性

社会人口学的情報	SMT 0 (n=95)	SMT 6 (n=99)	SMT 12 (n=97)	SMT 18 (n=100)	全体 (n=391)
年齢 (歳)	40.9 (14.1)	41.4 (14.8)	41.8 (14.0)	41.2 (13.8)	41.3 (14.1)
女性 (%)	49	49	49	52	50
非白人またはヒスパニック系 (%)	14	18	11	16	15
既婚者 (%)	37	28	41	36	36
大学学位 (%)	58	63	51	53	56
収入 2 万ドル未満 / 年 (%)	31	27	19	28	26
喫煙者 (%)	17	13	6	8	11
期待					
治療成功への信頼					
脊椎徒手整復 (1 ~ 6 尺度) *	3.6 (1.2)	3.8 (1.1)	3.7 (1.2)	3.8 (1.1)	3.7 (1.2)
軽マッサージ (1 ~ 6 尺度) *	3.4 (1.2)	3.5 (1.2)	3.4 (1.2)	3.5 (1.2)	3.5 (1.2)
過去の治療					
脊椎徒手整復 (%)	71	70	74	72	72
軽マッサージ (%)	52	56	43	54	51
LBP 病状					
疼痛強度 (0 ~ 100 尺度) †	52.2 (16.3)	51.0 (18.2)	51.6 (17.5)	51.5 (16.8)	51.6 (17.2)
機能的障害 (0 ~ 100 尺度) †	45.2 (21.8)	44.8 (24.0)	46.1 (23.4)	45.2 (21.8)	45.3 (22.7)
疼痛による不快感 (0 ~ 100 尺度) †	41.7 (19.5)	41.1 (21.1)	40.3 (22.8)	42.4 (22.2)	41.4 (21.4)
疼痛があった日数 (過去 4 週間)	24.8 (4.8)	24.1 (5.5)	23.3 (5.7)	24.1 (4.6)	24.1 (5.2)
障害があった日数 (過去 4 週間)	7.4 (8.1)	6.7 (7.5)	6.8 (7.5)	6.5 (7.2)	6.8 (7.6)
継続期間 (年)	11.6 (9.5)	11.2 (9.8)	11.7 (10.4)	12.5 (9.5)	11.8 (9.8)
健康状態					
SF-12 身体的健康要素 ‡	43.0 (9.5)	43.8 (8.9)	44.3 (8.4)	42.3 (8.8)	43.3 (8.9)
SF-12 精神的健康要素 ‡	50.2 (10.5)	48.6 (10.5)	47.6 (11.2)	49.4 (9.6)	48.9 (10.5)
健康状態 (0 ~ 100 視覚的アナログ尺度) ‡	70.1 (17.2)	72.1 (13.8)	73.5 (14.4)	68.2 (17.4)	70.9 (15.8)
その他の併存疾患 (%)	58	57	52	54	55
恐怖回避思考質問票					
仕事に関連する思考 (0 ~ 100 尺度) ‡	36.6 (23.2)	32.0 (23.4)	31.0 (18.9)	32.2 (21.4)	32.9 (21.8)
活動に関連する思考 (0 ~ 100 尺度) ‡	55.0 (20.1)	53.8 (23.1)	56.4 (17.7)	58.8 (19.8)	56.0 (20.3)
経口薬の使用 (過去 4 週間の回数)					
処方薬	0.9 (3.6)	0.3 (1.1)	0.6 (2.6)	0.3 (1.5)	0.5 (2.4)
非処方薬	7.6 (10.0)	8.9 (10.8)	9.5 (10.0)	7.6 (9.4)	8.4 (10.1)

SMT : 脊椎徒手整復、SF-12 : 12 項目健康関連 QOL 質問票、LBP : 下背部痛。  
数値は平均値 (標準偏差) またはパーセンテージ。  
\*6 段階リッカート尺度、1 は信頼が最低であることを示し、6 は信頼が最高であることを示す。  
†スコアが低い方が好ましい。下背部痛の強度は、修正 Von Korff 尺度で評価した。  
‡スコアが高い方が好ましい。12 項目健康関連 QOL 質問票スコアは、米国一般集団 (平均 550、標準偏差 510) に対して標準化した。

存疾患、および試験介入の経験を報告した。LBP の平均期間は 11.8 年だった。参加者が経験した LBP の週当たりの平均日数は 6 日であり、その LBP に対して、週に 2 回服薬していた。

## 疼痛

疼痛は治療終了までに改善し、すべてのグループで無作為化後 52 週目まで持続した（表 2 および図 2）。グループ内の平均的な疼痛の減少は、SMT 治療群については 20 ポイントを上回った。

主要解析の、補正後の傾きと平均差（AMD）と信頼区間を表 2 に示す。疼痛強度において、各治療レベルにわたるわずかな統計的有意な直線的用量反応効果が、12 週目の主要評価項目で

表 2 主要転帰

時間	観察された未補正平均値 (SD)				傾き (95% CI)	補正平均差 (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	SMT セッション 6 回あたり	SMT 0 対 6	SMT 0 対 12	SMT 0 対 18
疼痛強度 (0～100 尺度)								
0 週間	52.2(16.3)	51.0(18.2)	51.6(17.5)	51.5(16.8)	-	-	-	-
6 週間	34.5(18.4)	32.3(15.8)	27.1(14.7)	30.2(19.0)	1.8(0.3,3.2) *	1.7(-2.6,5.9)	7.2(2.8,11.6) *	4.1(-0.5,8.6)
12 週間	37.9(20.4)	32.7(19.4)	29.0(20.8)	31.4(19.8)	2.2(0.6,3.8) *	4.5(-0.6,9.6)	8.6(3.2,14.0) *	6.1(1.0,11.2) *
18 週間	35.7(19.5)	31.4(18.4)	30.3(19.3)	29.3(19.7)	2.0(0.4,3.5) *	3.6(-1.3,8.5)	5.1(0.2,10.0)	6.1(1.2,11.0) *
24 週間	34.9(20.6)	32.5(19.8)	33.7(20.5)	32.1(20.5)	0.6(-1.0,2.3)	1.7(-3.4,6.9)	0.8(-4.4,6.0)	2.4(-2.9,7.6)
39 週間	36.2(21.0)	32.8(21.5)	30.2(21.7)	31.6(21.5)	1.6(-0.2,3.3)	2.8(-2.8,8.4)	5.8(0.5,11.2)	4.3(-1.2,9.9)
52 週間	36.5(21.8)	30.7(22.4)	31.9(22.5)	28.7(20.5)	2.2(0.4,4.0) *	5.4(-0.4,11.1)	4.6(-1.2,10.3)	7.6(2.0,13.2) *
6-52 週間†	-	-	-	-	1.7(0.4,3.0) *	3.1(-1.0,7.1)	5.3(1.2,9.3) *	5.0(0.8,9.2) *
機能的障害 (0～100 尺度)								
0 週間	45.2(21.8)	44.8(24.0)	46.1(23.4)	45.2(21.8)	-	-	-	-
6 週間	27.0(20.2)	28.5(20.3)	25.8(19.3)	30.1(20.9)	0.6(-2.3,1.0)	1.7(-6.8,3.4)	1.5(-3.7,6.6)	-3.1(-8.3,2.1)
12 週間	29.2(23.7)	24.8(18.6)	22.0(20.7)	23.4(20.5)	2.0(0.3,3.8) *	4.2(-1.0,9.4)	7.5(1.7,13.3) *	5.8(0.2,11.3)
18 週間	26.1(21.4)	23.5(19.4)	22.1(21.5)	22.4(19.2)	1.3(-0.4,2.9)	2.2(-2.8,7.2)	4.1(-1.5,9.7)	3.6(-1.5,8.7)
24 週間	27.1(25.2)	25.6(21.7)	24.0(20.4)	24.1(20.3)	1.1(-0.7,2.9)	1.4(-4.5,7.2)	3.4(-2.4,9.3)	2.9(-2.9,8.8)
39 週間	26.2(22.8)	24.5(22.6)	21.7(20.5)	24.1(22.7)	0.9(-0.9,2.8)	1.4(-4.3,7.1)	4.7(-0.7,10.2)	2.0(-3.9,7.9)
52 週間	28.0(23.7)	22.6(22.4)	22.4(21.2)	19.1(18.7)	2.7(1.0,4.4) *	5.2(-0.5,10.9)	5.9(0.1,11.8)	8.8(3.3,14.4) *
6-52 週間†	-	-	-	-	0.9(-0.3,2.2)	1.5(-2.1,5.1)	3.9(-0.0,7.7)	2.4(1.5,6.3)

SMT：脊椎徒手療法、SD：標準偏差、CI：信頼区間。  
 主要評価項目は、12 週間毎 24 週間後の疼痛強度と機能的障害として事前に指定した。未補正の集団平均は、データの補完を行っていないオリジナルデータから算出したものであり、傾きと集団差は、ベースラインの共変量で補正した、データの補完を行ったデータより計算した。傾きと平均差の正符号は、脊椎徒手整備をより多く実施することが好ましいという計算結果を示す。統計的有意性の両側検定は、2つの主要転帰を明らかにするために .025 と事前に規定し、全ての統計検定にこれを用いた。  
 \* p < .025.  
 † 一般化推定方程式を用いた長期プロフィール



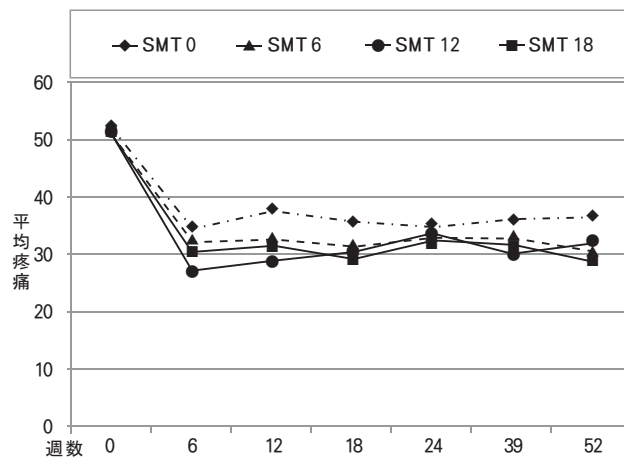


図2 疼痛の時間的プロファイル。疼痛強度は、0から100までの尺度で評価した。グラフは、各群におけるベースライン時と6回の追跡タイムポイントにおける疼痛の発生を示している。

主要評価項目は、12週間と24週間とした。SMT：脊椎徒手整復療法。

認められたが（6回の通院あたり2.2ポイント、 $p=0.007$ ）、24週目の主要評価項目では認められなかった（6回の通院あたり0.6ポイント）。他のタイムポイントでは、傾きは小さかった（1.6～2.0）。全体として、近接する用量群間では、全てのタイムポイントで差は最小限であった（図3）。

12週目では、12回のSMTの通院で治療と非SMTコントロール間で疼痛の最大の差が認められ（AMD=8.6、 $p=0.002$ ）、24週目では、コントロールとの意味ある差は認められなかった（AMD < 2.5）。副次タイムポイントについては、52週目に目立った効果が認められ、18回のSMTの通院がコントロールに対して最大の利益を示した（AMD=7.6、 $p=0.011$ ）。3種類の疼痛尺度要素

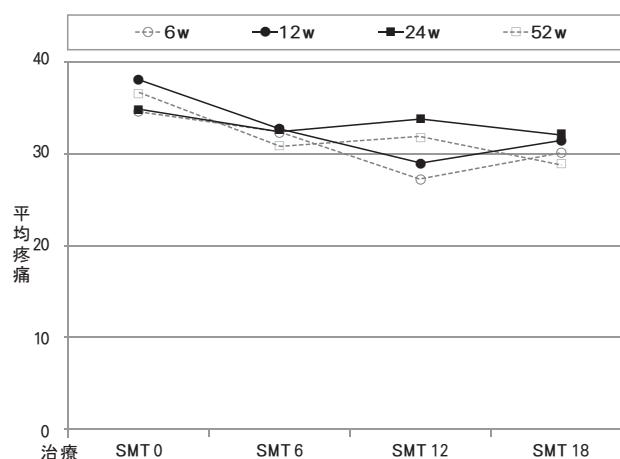


図3 疼痛の治療量反応曲線。治療量反応プロットは、各群を通じた疼痛強度（0ポイントから100ポイントの尺度）のわずかな傾きを、4点のタイムポイント（ケア終了(6週間)、主要評価項目(12週間、24週間)、最終フォローアップ(52週間)）について示している。線は、特定の治療量における時間の経過に伴う変化ではなく、特定のタイムポイントにおける各群にわたる差を表現していることに注意。SMT：脊椎徒手整復療法。

の解析も同様の結果を示した。この結果は著者らから入手可能である。

5 週目から 52 週目までの全疼痛プロファイルの反復評価解析により、コントロールを上回る治療の利点がわずかであることが示された。最大の効果は、12 回の SMT の通院対コントロールで認められた (AMD=5.3, p=.011)。

#### 機能的障害

機能的障害の平均的低下は、20 ポイントに達した。傾きと群比較における傾向は疼痛におけるそれと類似していたが、振幅はより小さく、統計的に有意な結果が少なかった (表 2)。12 週目の主要エンドポイントでは、12 回の SMT の通院でコントロールを上回る SMT の最大の利益も認められ (AMD=7.5, p=.011)、24 週目の主要エンドポイントでは、臨床的に意味のある効果はなかった (AMD < 3.4)。52 週目では、18 回の SMT の通院に最大の効果があることが認められた (AMD=8.8, p=.002)。疼痛と同様に、12 回および 18 回の SMT の通院の間に、障害プロファイルに臨床的に意味がある差はなかった。

#### 感度解析

補完されたデータを解析から除外した場合、疼痛転帰と障害転帰に関する結果に重大な変化はなかった。主要評価項目については、傾きでは変化は 0.3 未満であり、群間差では 1.4 ポイント未満 /100 ポイントであった。他のタイムポイントでの変化は、同様にごくわずかであった。ケア提供者またはクリニック別にクラスタ化した場合、効果量に実質的な差は生じなかった。

#### 応答者解析

図 4 および表 3 の応答者のプロファイルは、各グループの約 30%から 50%の個人が、各タイムポイントで 50%の疼痛改善を達成したことを示している。治療とコントロール間で統計

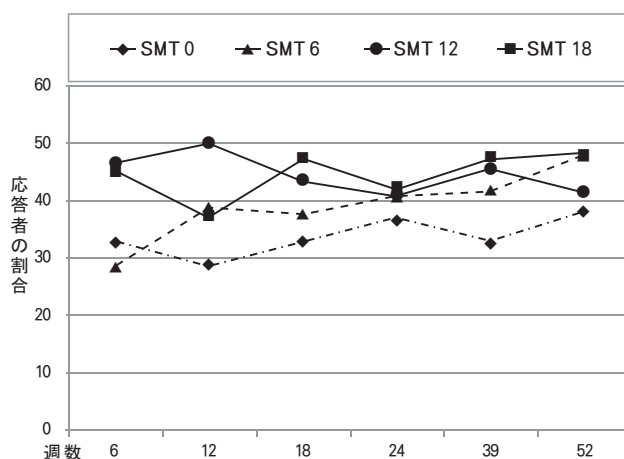


図 4 応答者：疼痛の50%改善を達成した個人の割合。  
SMT：脊椎徒手整復療法

表3 応答者（50%以上の個人的改善）

時間	観察された未補正の応答者 (SD)				傾き (95% CI)	応答者の補正差の割合 (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18		SMTセッション 6回あたり	SMT 0対6	SMT 0対12
疼痛強度								
6週間	32.6	28.3	46.4	45.0	5.5 (1.3, 9.8)*	-3.7 (-16.2, 8.8)	13.7 (0.6, 26.9)	13.0 (-0.4, 26.5)
12週間	28.4	38.4	49.5	37.0	3.8 (-0.5, 8.1)	10.0 (-3.2, 23.1)	21.1 (7.7, 34.6)*	8.9 (-4.2, 21.9)
18週間	32.6	37.4	43.3	47.0	4.8 (0.5, 9.2)	4.3 (-9.1, 17.6)	10.3 (-3.3, 23.9)	14.2 (0.5, 27.9)
24週間	36.8	40.4	40.2	42.0	1.4 (-2.9, 5.7)	3.7 (-10.0, 17.4)	3.2 (-10.5, 16.9)	4.9 (-8.7, 18.4)
39週間	32.6	41.4	45.4	47.0	4.8 (0.5, 9.1)	9.0 (-4.5, 22.5)	13.0 (-0.7, 26.6)	14.6 (1.1, 28.1)
52週間	37.9	47.5	41.2	48.0	2.5 (-1.9, 6.9)	10.2 (-3.5, 23.9)	3.9 (-9.8, 17.6)	10.6 (-3.2, 24.4)
機能的障害								
6週間	49.5	47.5	50.5	42.0	-0.1 (-0.5, 0.3)	-2.0 (-16.1, 12.0)	0.9 (-13.2, 15.1)	-7.5 (-21.4, 6.5)
12週間	43.2	49.5	59.8	55.0	4.4 (0.0, 8.8)	6.2 (-7.7, 20.0)	16.8 (2.9, 30.6)*	11.5 (-2.4, 25.4)
18週間	50.5	54.5	63.9	56.0	2.5 (-1.9, 6.9)	3.7 (-10.4, 17.9)	13.5 (-0.3, 27.2)	3.5 (-8.8, 19.3)
24週間	49.5	51.5	59.8	54.0	2.2 (-2.2, 6.6)	2.5 (-11.5, 16.5)	10.4 (-3.4, 24.3)	4.8 (-9.1, 18.6)
39週間	51.6	56.6	61.9	54.0	1.2 (-3.2, 5.6)	4.9 (-9.0, 18.8)	10.4 (-3.4, 24.2)	2.5 (-11.5, 16.4)
52週間	58.9	57.6	57.7	62.0	0.8 (-3.5, 5.2)	-1.1 (-14.8, 12.6)	-1.4 (-15.4, 12.6)	2.7 (-11.0, 16.5)

SMT：脊椎徒手整復療法、CI：信頼区間。  
 未補正の集団平均は、データの補完を行っていないオリジナルデータから算出したものであり、応答者の傾きと集団差は、ベースラインの共変量で補正した、データの補完を行ったデータより計算した。傾きと平均差の正符号は、SMTをより多く実施することが好ましいという計算結果を示す。  
 \* p < .025.

的に有意な差が認められたのは、1件のみであった。12週目では、ケアに対する応答者のかなり大きな割合が12回のSMTの通院のマニピュレーションに起因するものであった(21.1%、p=.002)。この差は、0.5相当を治療するために必要な数に相当する。機能的障害については、全てのグループとタイムポイントで、約40%から60%の個人が応答者であった。統計的に有意な群間差があったのは1件のみであった。

#### 副次転帰

一般に、ケアの終了時に記録された副次転帰に群内改善が認められ、これらの転帰は、疼痛転帰と同様の持続性を示していた。しかし、非SMTコントロール群における改善は、治療グループとコントロールグループ間に著しい統計的有意差があるものはほとんどないくらいの振幅であった(表4)。

疼痛および障害があった日数は、ベースラインから週に1日から2日ずつ減少した。感じられた疼痛および障害の改善は、一般に「より良い(better)」と評価された。標準化された12項目の身体的健康要素は、約7ポイントから10ポイント(最大1SD)改善し、3から6カ月のうちに米国の集団標準に戻った。精神的健康要素は、ベースライン時に集団標準から少し偏位し

表4 副次転帰

時間	観察された未補正平均値 (SD)				傾き (95% CI)	補正平均差 (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	SMTヒッショ ン 6回あたり	SMT 0対6	SMT 0対12	SMT 0対18
疼痛による不快感 (0～100尺度)								
0週間	41.7 (19.5)	41.1 (21.1)	40.3 (22.8)	42.4 (22.2)	-	-	-	-
6週間	24.7 (21.7)	20.3 (17.5)	15.6 (17.3)	18.4 (18.1)	2.4 (0.9, 4.0)*	4.2 (-0.8, 9.1)	8.7 (3.5, 14.0)*	6.6 (1.7, 11.6)*
12週間	29.7 (22.3)	24.5 (20.6)	22.2 (22.3)	23.9 (21.3)	2.0 (0.2, 3.7)	4.8 (-1.0, 10.5)	7.1 (1.1, 13.0)*	5.8 (0.1, 11.5)
18週間	26.8 (22.3)	21.9 (21.2)	20.4 (20.4)	21.3 (21.8)	1.8 (0.1, 3.6)	4.6 (-1.0, 10.3)	5.8 (0.4, 11.2)	5.8 (0.4, 11.2)
24週間	27.4 (21.7)	24.8 (21.8)	25.0 (21.9)	24.8 (22.4)	0.7 (-1.1, 2.5)	2.1 (-3.5, 7.8)	2.0 (-3.6, 7.5)	2.5 (-3.2, 8.3)
39週間	28.3 (21.6)	25.5 (21.3)	23.2 (20.8)	23.0 (21.1)	1.8 (0.1, 3.6)	2.4 (-3.3, 8.2)	4.8 (-0.6, 10.3)	5.4 (-0.2, 10.9)
52週間	27.8 (22.1)	22.1 (21.7)	24.0 (24.0)	21.5 (19.9)	1.7 (-0.1, 3.5)	5.4 (-0.3, 11.2)	3.7 (-2.4, 9.7)	6.4 (0.7, 12.1)
疼痛があった 日数 (過去4週間)								
0週間	24.8 (4.8)	24.1 (5.5)	23.3 (5.7)	24.1 (4.6)	-	-	-	-
6週間	18.9 (10.0)	17.6 (10.6)	15.0 (10.1)	16.3 (10.8)	0.8 (-0.0, 1.7)	0.9 (-1.8, 3.5)	3.0 (0.3, 5.6)	2.2 (-0.6, 4.9)
12週間	18.1 (9.5)	17.4 (9.2)	15.5 (9.7)	14.7 (10.2)	1.1 (0.2, 1.9)*	0.4 (-2.0, 2.8)	1.9 (-0.7, 4.5)	3.0 (0.5, 5.6)*
18週間	17.4 (10.2)	17.6 (10.3)	15.5 (10.3)	14.9 (10.6)	0.8 (-0.0, 1.7)	-0.6 (-3.2, 2.1)	1.2 (-1.5, 3.9)	2.2 (-0.6, 4.9)
24週間	16.9 (10.0)	16.5 (10.3)	15.4 (9.5)	13.5 (9.6)	1.0 (0.2, 1.8)*	0.1 (-2.5, 2.7)	0.7 (-1.8, 3.3)	3.1 (0.5, 5.7)*
39週間	17.7 (9.8)	15.6 (9.9)	15.1 (10.4)	14.3 (10.6)	0.9 (0.1, 1.7)	1.7 (-0.8, 4.2)	1.7 (-0.8, 4.3)	3.0 (0.4, 5.6)
52週間	17.0 (10.2)	15.3 (10.3)	14.0 (10.6)	13.6 (10.4)	1.0 (0.1, 1.8)*	1.3 (-1.4, 3.9)	2.1 (-0.7, 4.9)	3.0 (0.3, 5.7)
障害があった 日数 (過去4週間)								
0週間	7.4 (8.1)	6.7 (7.5)	6.8 (7.5)	6.5 (7.2)	-	-	-	-
6週間	2.4 (5.5)	1.6 (3.0)	1.3 (2.7)	1.2 (3.4)	0.3 (-0.0, 0.7)	0.8 (-0.4, 1.9)	1.0 (-0.1, 2.2)	1.1 (-0.1, 2.3)
12週間	3.4 (5.9)	2.6 (4.9)	2.0 (3.3)	2.3 (3.7)	0.4 (-0.0, 0.7)	0.6 (-0.7, 1.9)	1.3 (0.0, 2.5)	1.0 (-0.3, 2.2)
18週間	2.1 (5.3)	2.2 (5.0)	1.8 (4.1)	2.1 (4.9)	0.1 (-0.4, 0.5)	-0.1 (-1.4, 1.3)	0.3 (-0.9, 1.6)	0.0 (-1.3, 1.4)
24週間	3.5 (6.7)	2.7 (5.1)	2.8 (4.6)	2.8 (4.2)	0.1 (-0.3, 0.6)	0.7 (-0.7, 2.1)	0.6 (-0.8, 2.1)	0.5 (-0.9, 1.9)
39週間	2.8 (5.1)	3.0 (5.5)	2.4 (4.7)	2.5 (4.8)	0.1 (-0.3, 0.5)	-0.3 (-1.7, 1.1)	0.4 (-0.9, 1.6)	0.2 (-1.1, 1.5)
52週間	3.4 (6.0)	3.4 (6.9)	1.9 (3.7)	2.4 (4.7)	0.4 (-0.1, 0.8)	-0.1 (-1.7, 1.4)	1.5 (0.2, 2.8)	0.8 (-0.6, 2.1)
感知される 痛みの変化 (6段階リッカート)								
6週間	3.7 (0.9)	4.1 (0.9)	4.3 (1.0)	4.3 (1.0)	0.1 (0.1, 0.2)*	0.2 (-0.0, 0.5)	0.4 (0.2, 0.7)*	0.4 (0.1, 0.7)*
12週間	3.7 (0.9)	3.8 (0.9)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)	0.1 (-0.1, 0.4)	0.3 (-0.0, 0.5)	0.3 (0.0, 0.5)
18週間	3.7 (0.9)	4.0 (1.0)	4.1 (1.0)	4.0 (1.0)	0.1 (-0.0, 0.2)	0.2 (-0.0, 0.5)	0.4 (0.1, 0.7)*	0.2 (-0.0, 0.5)
24週間	3.8 (1.0)	3.9 (1.0)	3.9 (1.0)	3.9 (1.0)	0.0 (-0.1, 0.1)	0.1 (-0.2, 0.4)	0.1 (-0.2, 0.4)	0.1 (-0.2, 0.3)
39週間	3.7 (1.0)	3.8 (1.1)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)	0.1 (-0.2, 0.4)	0.3 (0.0, 0.6)	0.3 (0.0, 0.5)
52週間	3.7 (0.9)	4.0 (1.1)	3.9 (1.1)	4.0 (0.9)	0.1 (-0.0, 0.1)	0.3 (0.0, 0.6)	0.2 (-0.1, 0.5)	0.2 (-0.0, 0.5)
感知される 障害の変化 (6段階リッカート)								
6週間	3.6 (0.9)	3.9 (0.9)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)*	0.3 (0.0, 0.5)	0.4 (0.1, 0.6)*	0.4 (0.2, 0.6)*
12週間	3.5 (0.9)	3.8 (0.9)	3.8 (1.0)	3.8 (0.9)	0.1 (0.0, 0.2)*	0.2 (0.0, 0.5)	0.3 (0.0, 0.5)	0.3 (0.0, 0.5)
18週間	3.6 (0.8)	3.7 (1.0)	3.8 (1.0)	3.8 (0.9)	0.1 (-0.0, 0.2)	0.1 (-0.2, 0.3)	0.2 (-0.1, 0.4)	0.2 (-0.0, 0.5)
24週間	3.6 (0.9)	3.7 (0.9)	3.7 (0.9)	3.7 (1.0)	0.0 (-0.1, 0.1)	0.1 (-0.2, 0.3)	0.1 (-0.2, 0.3)	0.1 (-0.2, 0.4)
39週間	3.6 (0.9)	3.7 (1.0)	3.9 (1.0)	3.9 (0.9)	0.1 (0.0, 0.2)*	0.1 (-0.2, 0.4)	0.3 (0.1, 0.6)*	0.3 (0.0, 0.5)*
52週間	3.6 (0.9)	3.8 (1.0)	3.8 (1.1)	3.9 (0.9)	0.1 (-0.0, 0.2)	0.2 (-0.0, 0.5)	0.2 (-0.1, 0.5)	0.3 (0.0, 0.3)

SF-12 身体的健康要素	0週間	43.0 (9.5)	43.8 (8.9)	44.3 (8.4)	42.3 (8.8)	-	-	-	-
	12週間	45.5 (10.3)	47.1 (8.2)	49.6 (8.5)	47.5 (8.5)	1.0 (0.3, 1.6)*	1.2 (-0.8, 3.3)	3.5 (1.3, 5.7) *	2.4 (0.3, 4.6)
	24週間	50.0 (11.1)	50.5 (10.1)	51.4 (9.1)	50.9 (9.4)	0.5 (-0.3, 1.2)	0.0 (-2.3, 2.4)	0.8 (-1.6, 3.2)	1.3 (-1.1, 3.6)
	39週間	50.6 (11.5)	51.1 (10.3)	52.7 (9.6)	51.8 (9.3)	0.6 (-0.2, 1.4)	0.1 (-2.3, 2.4)	1.5 (-0.9, 4.0)	1.6 (-0.9, 4.0)
	52週間	50.7 (12.0)	50.8 (11.0)	52.6 (10.3)	52.5 (8.5)	0.8 (0.1, 1.6)	0.3 (-2.7, 2.1)	1.4 (-1.2, 4.0)	2.2 (-0.2, 4.5)
SF-12 精神的健康要素	0週間	50.2 (10.5)	48.6 (10.5)	47.6 (11.2)	49.4 (9.6)	-	-	-	-
	12週間	50.2 (10.8)	50.4 (9.4)	47.8 (11.0)	48.0 (9.8)	-0.7 (-1.3, -0.0)	1.3 (-0.7, 3.3)	-0.5 (-2.6, 1.5)	-1.5 (-3.5, 0.4)
	24週間	51.8 (10.9)	52.8 (10.2)	50.8 (11.8)	51.3 (11.2)	-0.1 (-0.8, 0.6)	2.1 (-0.0, 4.2)	0.7 (-1.3, 2.8)	0.1 (-2.1, 2.2)
	39週間	51.7 (11.3)	51.5 (11.6)	49.2 (13.6)	49.0 (11.7)	-0.8 (-1.5, -0.0)	1.1 (-1.3, 3.4)	-0.4 (-2.9, 2.1)	-2.0 (-4.3, 0.3)
	52週間	51.3 (12.0)	50.4 (11.4)	50.6 (12.7)	50.4 (11.7)	-0.0 (-0.8, 0.8)	0.2 (-2.3, 2.7)	1.1 (-1.6, 3.7)	-0.3 (-2.9, 2.3)
EuroQol 健康状態 (0~100尺度)	0週間	70.1 (17.2)	72.1 (13.8)	73.5 (14.4)	68.2 (17.4)	-	-	-	-
	12週間	73.5 (17.3)	78.4 (14.1)	77.9 (15.0)	75.7 (14.5)	0.8 (-0.4, 2.0)	3.9 (0.0, 7.7)	2.7 (-1.3, 6.7)	3.1 (-0.8, 6.9)
	24週間	73.9 (17.5)	77.8 (15.5)	77.0 (15.4)	74.5 (16.7)	0.3 (-1.0, 1.6)	2.9 (-1.0, 6.9)	1.4 (-2.6, 5.5)	1.5 (-2.7, 5.8)
	39週間	73.1 (20.0)	76.8 (17.2)	76.6 (15.6)	75.3 (16.8)	0.9 (-0.5, 2.3)	2.6 (-1.9, 7.0)	1.8 (-2.6, 6.2)	3.2 (-1.3, 7.6)
	52週間	74.8 (17.0)	77.1 (17.0)	77.3 (15.3)	77.2 (14.9)	1.0 (-0.3, 2.2)	1.3 (-2.7, 5.4)	0.9 (-3.1, 4.9)	3.3 (-0.5, 7.2)
非処方薬 (過去4週間の回数)	0週間	7.6 (10.0)	8.9 (10.8)	9.5 (10.0)	7.6 (9.4)	-	-	-	-
	6週間	4.3 (6.8)	4.1 (6.9)	4.0 (6.7)	4.0 (8.3)	0.1 (-0.5, 0.7)	0.6 (-1.2, 2.4)	0.9 (-0.8, 2.6)	0.3 (-1.5, 2.1)
	12週間	7.8 (11.0)	7.1 (23.0)	5.8 (7.6)	7.8 (25.8)	0.1 (-1.6, 1.8)	1.4 (-2.7, 5.6)	3.2 (0.4, 5.9) *	-0.1 (-5.5, 5.3)
	18週間	5.7 (7.4)	6.0 (10.4)	5.3 (9.1)	5.5 (9.1)	0.2 (-0.5, 0.8)	0.1 (-2.4, 2.6)	1.0 (-1.1, 3.2)	0.2 (-1.8, 2.2)
	24週間	8.1 (13.7)	7.7 (15.8)	6.5 (8.6)	6.4 (10.8)	0.7 (-0.4, 1.7)	0.9 (-3.3, 5.1)	2.4 (-0.7, 5.5)	1.8 (-1.5, 5.0)
	39週間	7.2 (14.2)	7.4 (12.2)	6.6 (8.9)	7.7 (13.5)	-0.0 (-1.1, 1.1)	0.3 (-3.4, 4.0)	1.4 (-1.9, 4.7)	-0.5 (-4.0, 3.1)
	52週間	6.5 (8.0)	6.6 (13.7)	6.7 (9.3)	6.8 (12.7)	-0.1 (-0.9, 0.8)	0.3 (-2.7, 3.4)	0.6 (-1.5, 2.6)	-0.3 (-2.9, 2.4)
<p>SMT : 脊椎徒手整復療法、SD : 標準偏差、CI : 信頼区間、SF-12 : 12項目健康関連 QOL 質問票。  未補正の集団平均は、データの補完を行っていないオリジナルデータから算出したものであり、傾きと集団差は、ベースラインの共変量で補正した、データの補完を行ったデータより計算した。傾きと平均差の正符号は、SMT をより多く実施することが好ましいという計算結果を示す。  * p &lt; .025.  † 6段階リッカート尺度:かなり悪くなった = 1、悪くなった = 2、ほぼ同じ = 3、良くなった = 4、かなり良くなった = 5、  および完全に回復した = 6。</p>									

ていた。EuroQol の健康状態視覚的アナログ尺度は、ベースラインからほとんど変化を示さなかった。ケア終了後の投薬使用には、わずかな減少があった。疼痛による不快感スコアの平均低下は、疼痛スコアの低下と同様であり、約 20 ポイントだった。

#### 有害事象

注目に値する有害事象は認められなかった。3 人が、本研究に関連した LBP 悪化の症状緩和のために治療を求めたことを報告した。不適格者 1 人が、その後、スクリーニング検査の後の痛みの増大を報告した。12-SMT グループの参加者 1 人は、完全解消または治療段階中の症状の発現により、数日間仕事を失った。12-SMT グループの参加者 1 人は、子どもを抱き上げたことに関連した悪化後に、途中で参加をやめた。

## 考察

SMTに関する初の大規模用量反応研究は、いくつかの重要な知見を得た。疼痛と機能的障害の主要転帰に基づいて、12回のSMTセッションで、軽度ではあるものの全体的に最良の治療効果（群間差）が得られた。このことは、特に短期で、12週間の主要評価項目で顕著だった。群間差は、24週目の主要評価項目ではごくわずかであり、52週目の長期間では、18回のSMTセッションがわずかながら支持された。一般に、データは用量反応関係と一致しており、12回のセッションで飽和し、追加のSMTのための通院による利益は52週目でもそれ以上ほとんど無かったか、全く無かった。フルタイム・プロファイルの解析は、12回のセッションを上回る、18回全体での追加の利益が皆無であることを裏付けた。また、短期間のみではあるが、12回の通院にいくらかの利点があることを応答者解析によりさらに裏付けられた。

主要転帰に関する直線用量反応の傾きは一般に小さく、12週目と52週目で6回のSMTセッションあたり、およそ2/100尺度ポイントに達した。短期の結果について最高用量のグループを除外したとしても、傾きは6回のSMTセッションあたり約4/100尺度ポイントに対してわずか2倍でしかない。近接する群間にほとんど差が無いという事実があるため、他を差し置いて1つの治療用量を推奨することが難しくなる。しかし、以下の2つのことが考えられる。第1に、各用量にわたる効果は、実地コントロールを上回るSMTのわずかな利益に対して蓄積する。第2に、本研究の目的は、今後の試験で用いる飽和用量レベルを見いだすことである。時間プロファイル、用量反応の傾き、およびコントロールグループとの比較は、全体として、特に短期的改善には、12回の通院が最良の選択であることを示唆している。

用量反応効果の解釈には、いくつかの要因を考慮する必要がある。本研究は、SMTの効果を探し出すようデザインされた、細心の注意を払った無作為化試験であった。我々は通院回数、患者と過ごす時間、患者への接触の効果、患者と施術者の相互関係、および介入の信頼性を制御した。これは、全てのグループで18回の実地両方と電氣的モダリティ（極微超音波）により達成された。非SMT通院での軽マッサージ、超音波、または医療サービス提供者への18回の通院の具体的かつ状況的な効果のため、医療現場で認められる可能性がある、より大きな用量反応の傾きが分かりにくくなる可能性がある。例えば、このようなより大きな効果は、我々の予備試験では、参加者が積極的介入のための通院のみに参加した場合に認められた<sup>[5]</sup>。

有効性に関して、軽マッサージ・コントロールは、技術的には真の疑似介入というより、むしろ比較介入である。最小限のマッサージであっても、期待以上の多数のセッションでより大きな効果が得られる可能性がある。そのため、SMTとコントロール間の差は、疑似マニピュレーションとの比較を行うには、若干小さいかも知れない。我々は、次の2つの理由から、疑似治療の使用を試みなかった。第1に、半数が治療とコントロールの両方に通院して処置を受けることから、参加者を盲検化することは事実上不可能であり、介入を比較してしまう可能性がある。第2に、我々は参加者が疑似介入を受けているかも知れないと考えた場合に生じる失望を回避

しなかった。

全患者を、割り当てられた SMT 用量による治療を受ける予定とした。治療期間中は、改善に基づく治療中止ルールは設けなかった。治療中止の影響は不明であり、短期および長期的に転帰に対して利益があるのか不利益であるのかのどちらかである可能性がある。

別の問題は、連続的に変化する主要転帰に関する、臨床的に重要な群間差の閾値である。患者が評価を行った場合の最低限の重要変化に関する研究は、一部の著者が 30%（約 15～20/100 ポイント）の改善を、これらの転帰に関する臨床的に重要な個人内の最小変化のロバストな指標と見なすことができると結論づけるとの結果を導き出している<sup>[35]</sup>。個人に関して、50%の改善が成功の閾値として推奨が行われてきた<sup>[36]</sup>。しかし、Dworkin ら<sup>[37]</sup>は、これらの数字は、群間の効果には適用せず、意味ある群間差の特定は、単純明快とはほど遠い多面的なプロセスである、と指摘している。主要評価項目で、疼痛スコアで 8.6、障害スコアで 7.5 だった群間差は、明らかに最低限度であったが、この効果が臨床的に関連性ある程度の振幅となるかどうかは、依然として不明である。疼痛に対する治療に必要なだった近接回数 (=5) は、実際には意味ある効果を示している可能性がある<sup>[12,38]</sup>。

## 結論

全体として、カイロプラクターによる 6 週間で 12 回の脊椎徒手整復セッションが、慢性非特異的 LBP に対して最も好ましい疼痛と機能的障害の改善をもたらした。このグループに関して、参加者の平均的改善は、ケア終了時で実質的なものであり、52 週間まで維持可能であった。およそ半数の患者が、疼痛/障害の 50%改善を達成すると期待される。従って、相対的有効性試験においては、12 回の SMT セッションを用いるのが、現在、最良と推定される。しかし、12 回の通院を近接する用量レベルと明確に区別するには、用量グループ間の治療効果の傾きが小さすぎることから、推奨は注意して行われる。12 回の通院でも、カイロプラクターによる集中的な軽マッサージ（実地コントロール）を上回る転帰に対する SMT の寄与は、12 週目の主要評価項目では最高に適しており、24 週目の主要評価項目ではごくわずかであった。

## 謝辞

本研究は、米国国立衛生研究所の国立補完代替医療センター（NCCAM）による資金提供（U01 AT001908）を受けた。この発表文献の内容は、著者らの単独責任であり、必ずしも NCCAM の公式見解を反映するものではない。著者らは、利害の対立がないことを宣言する。





Clinical Study

# Dose-response and efficacy of spinal manipulation for care of chronic low back pain: a randomized controlled trial

Mitchell Haas, DC<sup>a,\*</sup>, Darcy Vavrek, ND<sup>a</sup>, David Peterson, DC<sup>b</sup>, Nayak Polissar, PhD<sup>c</sup>,  
Moni B. Neradilek, MS<sup>c</sup>

<sup>a</sup>Center for Outcomes Studies, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA

<sup>b</sup>Division of Chiropractic Sciences, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA

<sup>c</sup>The Mountain-Whisper-Light Statistics, 1827 23rd Ave. E., Seattle, WA 98122, USA

Received 26 November 2012; revised 17 July 2013; accepted 22 July 2013

## Abstract

**BACKGROUND CONTEXT:** There have been no full-scale trials of the optimal number of visits for the care of any condition with spinal manipulation.

**PURPOSE:** To identify the dose-response relationship between visits to a chiropractor for spinal manipulation and chronic low back pain (cLBP) outcomes and to determine the efficacy of manipulation by comparison with a light massage control.

**STUDY DESIGN/SETTING:** Practice-based randomized controlled trial.

**PATIENT SAMPLE:** Four hundred participants with cLBP.

**OUTCOME MEASURES:** The primary cLBP outcomes were the 100-point modified Von Korff pain intensity and functional disability scales evaluated at the 12- and 24-week primary end points. Secondary outcomes included days with pain and functional disability, pain unpleasantness, global perceived improvement, medication use, and general health status.

**METHODS:** One hundred participants with cLBP were randomized to each of four dose levels of care: 0, 6, 12, or 18 sessions of spinal manipulation from a chiropractor. Participants were treated three times per week for 6 weeks. At sessions when manipulation was not assigned, they received a focused light massage control. Covariate-adjusted linear dose effects and comparisons with the no-manipulation control group were evaluated at 6, 12, 18, 24, 39, and 52 weeks.

**RESULTS:** For the primary outcomes, mean pain and disability improvement in the manipulation groups were 20 points by 12 weeks and sustainable to 52 weeks. Linear dose-response effects were small, reaching about two points per six manipulation sessions at 12 and 52 weeks for both variables ( $p < .025$ ). At 12 weeks, the greatest differences from the no-manipulation control were found for 12 sessions (8.6 pain and 7.6 disability points,  $p < .025$ ); at 24 weeks, differences were negligible; and at 52 weeks, the greatest group differences were seen for 18 visits (5.9 pain and 8.8 disability points,  $p < .025$ ).

**CONCLUSIONS:** The number of spinal manipulation visits had modest effects on cLBP outcomes above those of 18 hands-on visits to a chiropractor. Overall, 12 visits yielded the most favorable results but was not well distinguished from other dose levels.

© 2014 The Authors. Published by Elsevier Inc. Open access under CC BY-NC-SA license.

## Keywords:

Chronic low back pain; Dose-response; Spinal manipulation; Chiropractic; Randomized controlled trial

FDA device/drug status: Not applicable.

Author disclosures: **MH:** Grant: NIH: NCCAM (I, Paid directly to institution). **DV:** Grant: NIH: NCCAM (I, Paid directly to institution). **DP:** Grant: NIH: NCCAM (I, Paid directly to institution). **NP:** Fees for participation in review activities such as data monitoring boards, statistical analysis, end-point committees, and the like: University of Western States (B). **MBN:** Fees for participation in review activities such as data monitoring

boards, statistical analysis, end-point committees, and the like: University of Western States (B).

The disclosure key can be found on the Table of Contents and at [www.TheSpineJournalOnline.com](http://www.TheSpineJournalOnline.com).

The trial was registered at [ClinicalTrials.gov](http://ClinicalTrials.gov) NCT00376350.

\* Corresponding author. Center for Outcomes Studies, University of Western States, 2700 NE 132nd Ave., Portland, OR 97230, USA. Tel.: (503) 251-5728; fax: (503) 251-2832.

E-mail address: [haasmitch@comcast.net](mailto:haasmitch@comcast.net) (M. Haas)

## Introduction

It has long been known that low back pain (LBP) is a prevalent and costly condition [1,2] and that chiropractors provide the vast majority of spinal manipulation [3] and treat a large proportion of LBP in the United States [4]. It is therefore important to determine the optimal quantity of spinal manipulative therapy (SMT), particularly for chronic low back pain (cLBP) [5].

Recommendations for duration and frequency of SMT/chiropractic care for cLBP have varied widely and have long been based on the clinical experience and opinion [6]. In the early 1990s, a multidisciplinary RAND panel found that opinion was too varied to come to a formal evidence-based consensus (2–24 weeks of care, 1–5 visits per week), but on average, members expected the typical patient to improve in 4 to 6 weeks with three visits per week [7]. In contrast, an all-chiropractic RAND expert panel recommended 30 visits over 14 weeks [8]. Shekelle et al. [3] noted a range of 1 to 19 visits in the published studies of chiropractic care. Later, Nyiendo et al. [9] found a mean of 6.7 visits (standard deviation [SD]=7.5, range=1–56) in a practice-based cohort of 526 nonspecific cLBP patients.

To this day, there is no consensus on the efficacy of SMT and its role in the care of cLBP. Some systematic reviews have reported quality evidence in support of SMT [10,11], whereas others including the latest Cochrane review found SMT to be no better than other interventions [12]. Results of systematic reviews, whether meta-analysis or best-evidence synthesis, may depend on the quantity of care used in the trials included in the reviews. Investigators have had virtually no evidence from dose-response trials to inform the number of SMT sessions provided.

Because of the dearth of evidence for duration and frequency of care, we conducted the first pilot randomized trial evaluating dose-response of SMT (n=72) [5]. We found a clinically important association between number of visits to a chiropractor (1–4 weekly visits for 3 weeks) and short-term pain and disability relief showing that a higher number of visits yielded more favorable results. We have subsequently conducted the current 5-year study, the first full-scale dose-response trial with the aim of identifying optimal care of cLBP with SMT and informing the design of comparative effectiveness studies. We also evaluated the efficacy of the SMT dose levels by testing the hypothesis of no difference between SMT and a hands-on control. The trial evaluated the unique contribution of SMT to outcomes beyond the effects of a light massage to control attention (quantity of visits) and touching the patient, history, and context [13].

## Methods

### Design

In a prospective open-label randomized controlled trial, 400 participants with nonspecific cLBP were randomized to

## EVIDENCE & METHODS

### Context

The impact of spinal manipulation on chronic low back pain compared to light massage, serving as a control, is assessed by the authors.

### Contribution

In this well-performed RCT, spinal manipulation for 18 visits was found to result in modestly improved outcomes relative to light massage, but such improvements may not be clinically significant. They also found that the best treatment effects for manipulation were at 12 sessions versus the control and no additional benefit was afforded at 18 sessions.

### Implications

This study provides some guidance where there is currently little. Twelve chiropractic sessions are reasonable, manipulation may be modestly better than light massage at this endpoint, but not at 24 weeks. As important, the study serves as a reasonable model for the design of a practice-based randomized trial.

—The Editors

receive a dose of 0, 6, 12, or 18 SMT sessions from a chiropractor. All participants were assigned 18 treatment visits, 3 per week for 6 weeks. Spinal manipulative therapy was performed at the assigned number of visits, and a brief light massage control was performed at non-SMT visits to control provider attention and touching the participants [14]. For example, those receiving 12 visits for SMT received 6 visits for light massage from the chiropractor (Fig. 1). Follow-up evaluation was by mailed questionnaire or blinded phone interview at 6, 12, 18, 24, 39, and 52 weeks after randomization. The primary outcomes were prespecified as self-reported pain intensity and functional disability at the 12- and 24-week end points. The primary end points were chosen to emphasize a short- and a long-term post-treatment time point.

Randomization was conducted using computer-generated design-adaptive allocation [15,16] to balance six baseline variables across groups: pain and disability scores, age, gender, relative confidence in SMT and massage, and any previous SMT or massage care. Allocation to study groups was hence concealed from all study personnel and participants by requiring entry of data into the computer program collected immediately before randomization (pain, disability, and confidence in treatment success). Patient coordinators called in the allocation variables over the phone to research staff who entered the data into the allocation computer program. The patient coordinator then assigned the participant to group by placing an unmarked sealed envelope identifying care in the patient's clinic file. Participants and treating

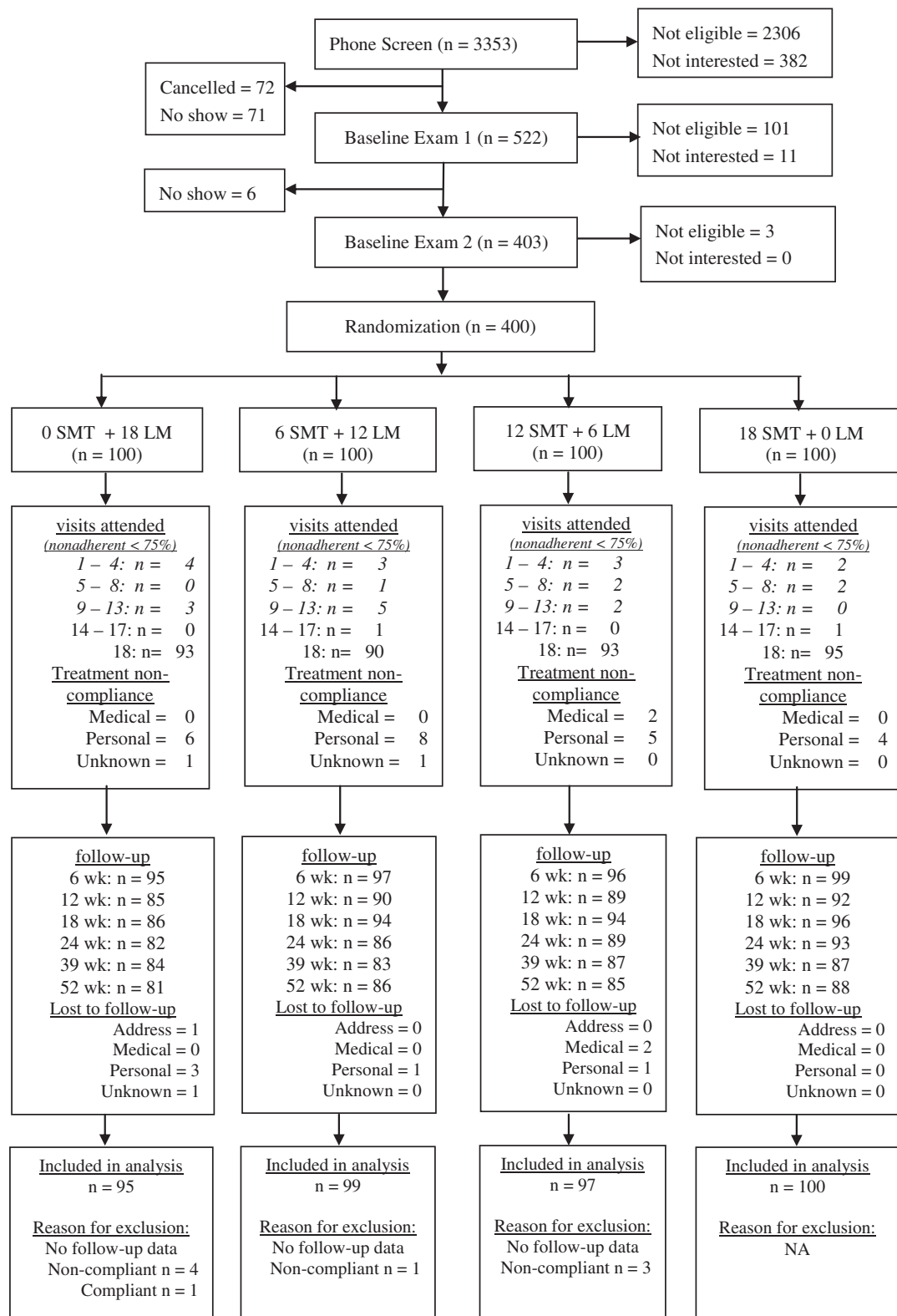


Fig. 1. Study flow diagram. All participants were assigned 18 treatment visits. They received either spinal manipulative therapy (SMT) or light massage control (LM) at any one visit.

clinicians were not blinded to intervention after randomization. However, patient coordinators, who collected some outcomes by phone interview, remained blinded to group assignment throughout the study. The study was approved by the University of Western States Institutional Review Board (FWA 851). The trial was registered at ClinicalTrials.gov, NCT00376350.

### *Protocol overview*

Participants were recruited through craigslist, mailers, and local newspapers. They were informed that the study was investigating 18 visits for different combinations of two manual therapies for cLBP. Preliminary screening was conducted through telephone interview by study staff. At the first baseline visit, participants signed an informed consent form and filled out a baseline survey. Eligibility screening examinations were then conducted at a central university clinic by one of the two licensed chiropractors (faculty with >20 years of experience) using history, physical examination, and lumbar X-rays [17].

Eligible participants selected a convenient clinic for study care. Care was provided by one of the 12 licensed chiropractors with 4 to 24 years of experience in 9 Portland-area clinics. The treating chiropractors were selected because their abilities were known to the authors, and some had previous experience on our trials. After a second baseline survey at the clinic, participants were given a brief confirmatory screening examination by the treating clinician. They were then randomized and received their first treatment. Participants were compensated for each treatment visit, mailed questionnaires, and phone interviews (\$10–\$20). Participants signed an informed consent form.

### *Participants*

Participants were required to be at least 18 years old and have a current episode of cLBP [18] of mechanical origin [19] of at least 3 months duration [3]. They were further required to have had some LBP on 30 days in the previous 6 weeks and a minimum LBP index of 25 on a 100-point scale to prevent floor effects. Participants were excluded if they received manual therapy within the previous 90 days or for contraindications to study interventions [17,20] and complicating conditions such as active cancer, spine pathology, inflammatory arthropathies, autoimmune disorders, anticoagulant conditions, neurodegenerative diseases, pain radiating below the knee, organic referred pain, pregnancy, and disability compensation.

### *Intervention*

Each visit was 15 minutes long with a treating chiropractor, consistent with chiropractic practice [21]. Participants received a hot pack for 5 minutes to relax spinal muscles followed by 5 minutes for the SMT or control intervention. The visit was completed with 5 minutes of very

low-dose pulsed ultrasound (20% duty cycle with 0.5 watts/cm<sup>2</sup>). This was used as a quasi-sham to enhance treatment credibility and adherence to care [13].

Spinal manipulative therapy consisted of manual thrust (high velocity, low amplitude) spinal manipulation in the lumbar and transition thoracic regions, predominantly in the side-posture position [22]. Specific manipulations to be performed were determined at each visit by the chiropractor through ongoing evaluation of the participants including patient progress, self-reported and provocative pain, spinal range of motion, and palpation of the spine and paraspinal soft tissue [17,22]. Manipulation was not performed at a visit, if the treating chiropractor failed to find any indication. Lighter thrust manipulation including the use of mechanical assistance of a spring-loaded table and segmental low-velocity mobilization were permitted in the case of acute exacerbation of the lumbar spine pain [22].

The light massage control consisted of 5 minutes of gentle effleurage and petrissage of the low back (lumbar and lower thoracic) paraspinal muscles [22,23], focused on the symptomatic areas. The massage used was gentler and of shorter duration than recommended for therapeutic massage practice [21,24]. As such, it was a minimalist intervention to control touching the patient; it was not a formal sham. The treating chiropractors were also asked to render SMT and control intervention with equal enthusiasm to help balance expectations of treatment success imparted by the practitioner. Protocol standardization and provider equipoise across treatment groups were maintained through quarterly training and monitored by office observation and patient phone interview [14,25].

### *Outcome and baseline variables*

This report emphasizes the prespecified primary outcomes, the self-reported modified Von Korff pain and disability scales validated by Underwood et al. [26]. The pain score is the average of three 11-point numeric rating scales converted to a 100-point scale: back pain today, worst back pain in the last 4 weeks, and average back pain in the last 4 weeks. The disability score is also the average of three scales: interference with daily activities, social and recreational activities, and the ability to work (outside or around the house).

Secondary outcomes included pain unpleasantness [27], Physical and Mental Component Summary Scales of the short-form 12 [28], Health State Visual Analog Scale from EuroQol [29], perceived pain and disability improvement, and the number of the following in the previous 4 weeks: days with pain and disability and medication use. Additional baseline variables included demographics, Fear-Avoidance Beliefs Questionnaire [30], confidence in treatment success [14], and any from a list of comorbid conditions (arthritis, asthma or allergies, gastrointestinal problems, gynecological problems, hypertension, or other chronic condition) [31].

### Statistical analysis

An intention-to-treat analysis was conducted with each participant included in the original allocation group; missing data were imputed using linear interpolation and then last datum was carried forward. Nine participants were omitted from the analysis because they had no follow-up data. A sensitivity analysis with all missing data excluded was conducted for the two primary outcomes.

The prespecified primary analysis consisted of regression models to identify the linear effect of SMT dose (slope=outcome increment/six SMT sessions) and to compare each SMT group to the no-SMT control group (adjusted mean differences). Seemingly unrelated (simultaneous) regression by Zellner [32] was used to model outcomes for the individual time points [33]. In addition, for the primary outcomes only, longitudinal effects across all follow-ups were modeled with generalized estimating equations using unstructured correlation to account for within-person correlation between time points [33]. Slopes and group

comparisons were adjusted for the six baseline balancing variables used to randomize the participants [15,16]: pain and disability scores, age, gender, relative confidence in SMT and massage, and any previous SMT or massage care. The baseline value of the outcome measure was added as a covariate if not already included among the six balancing variables.

In a prespecified secondary “responder” analysis, the two primary outcomes were dichotomized to show the proportion of participants with 50% improvement. The analysis mentioned previously was then repeated using binomial regression to identify slopes and group differences in proportion of responders [33].

The sample size was determined a priori to have at least 80% power to detect a between-group effect of 10 of 100 points in the two primary outcomes using a two-tailed test. It took into consideration a 10% dropout rate. The .025 level of significance was used to adjust for having two primary outcomes. Detection of a 10-point between-

Table 1  
Baseline characteristics

	SMT 0 (n=95)	SMT 6 (n=99)	SMT 12 (n=97)	SMT 18 (n=100)	All (n=391)
Sociodemographic information					
Age (y)	40.9 (14.1)	41.4 (14.8)	41.8 (14.0)	41.2 (13.8)	41.3 (14.1)
Female (%)	49	49	49	52	50
Non-white or Hispanic (%)	14	18	11	16	15
Married (%)	37	28	41	36	36
College degree (%)	58	63	51	53	56
Income <\$20,000/y (%)	31	27	19	28	26
Smoker (%)	17	13	6	8	11
Expectations					
Confidence in treatment success					
Spinal manipulation (1–6 scale)*	3.6 (1.2)	3.8 (1.1)	3.7 (1.2)	3.8 (1.1)	3.7 (1.2)
Light massage (1–6 scale)*	3.4 (1.2)	3.5 (1.2)	3.4 (1.2)	3.5 (1.2)	3.5 (1.2)
Previous treatment					
Spinal manipulation (%)	71	70	74	72	72
Light massage (%)	52	56	43	54	51
LBP complaint					
Pain intensity (0–100 scale)†	52.2 (16.3)	51.0 (18.2)	51.6 (17.5)	51.5 (16.8)	51.6 (17.2)
Functional disability (0–100 scale)†	45.2 (21.8)	44.8 (24.0)	46.1 (23.4)	45.2 (21.8)	45.3 (22.7)
Pain unpleasantness (0–100 scale)†	41.7 (19.5)	41.1 (21.1)	40.3 (22.8)	42.4 (22.2)	41.4 (21.4)
Days with pain (last 4 wk)	24.8 (4.8)	24.1 (5.5)	23.3 (5.7)	24.1 (4.6)	24.1 (5.2)
Days with disability (last 4 wk)	7.4 (8.1)	6.7 (7.5)	6.8 (7.5)	6.5 (7.2)	6.8 (7.6)
Duration (y)	11.6 (9.5)	11.2 (9.8)	11.7 (10.4)	12.5 (9.5)	11.8 (9.8)
Health status					
SF-12 physical health component‡	43.0 (9.5)	43.8 (8.9)	44.3 (8.4)	42.3 (8.8)	43.3 (8.9)
SF-12 mental health component‡	50.2 (10.5)	48.6 (10.5)	47.6 (11.2)	49.4 (9.6)	48.9 (10.5)
Health state (0–100 Visual Analog Scale)‡	70.1 (17.2)	72.1 (13.8)	73.5 (14.4)	68.2 (17.4)	70.9 (15.8)
Other comorbidity (%)	58	57	52	54	55
Fear-Avoidance Beliefs Questionnaire					
Work beliefs (0–100 scale)‡	36.6 (23.2)	32.0 (23.4)	31.0 (18.9)	32.2 (21.4)	32.9 (21.8)
Activity beliefs (0–100 scale)‡	55.0 (20.1)	53.8 (23.1)	56.4 (17.7)	58.8 (19.8)	56.0 (20.3)
Oral medication use (times in last 4 wk)					
Prescription	0.9 (3.6)	0.3 (1.1)	0.6 (2.6)	0.3 (1.5)	0.5 (2.4)
Nonprescription	7.6 (10.0)	8.9 (10.8)	9.5 (10.0)	7.6 (9.4)	8.4 (10.1)

SMT, spinal manipulation; SF-12, short-form 12; LBP, low back pain.

Values are mean (standard deviation) or percentage.

\* Six-point Likert scale with 1 indicating lowest and 6 indicating highest confidence.

† Lower scores favorable. Low back pain intensity and functional disability evaluated with modified Von Korff scales.

‡ Higher scores favorable. Short-form 12 health survey scores are standardized to the US general population (mean=50, standard deviation=10).

Table 2  
Primary outcomes

Time	Observed unadjusted mean (SD)				Slope (95% CI)	Adjusted mean difference (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	Per six SMT sessions	SMT 0 vs. 6	SMT 0 vs. 12	SMT 0 vs. 18
<b>Pain intensity (0–100 scale)</b>								
0 wk	52.2 (16.3)	51.0 (18.2)	51.6 (17.5)	51.5 (16.8)	—	—	—	—
6 wk	34.5 (18.4)	32.3 (15.8)	27.1 (14.7)	30.2 (19.0)	1.8 (0.3, 3.2)*	1.7 (–2.6, 5.9)	7.2 (2.8, 11.6)*	4.1 (–0.5, 8.6)
12 wk	37.9 (20.4)	32.7 (19.4)	29.0 (20.8)	31.4 (19.8)	2.2 (0.6, 3.8)*	4.5 (–0.6, 9.6)	8.6 (3.2, 14.0)*	6.1 (1.0, 11.2)*
18 wk	35.7 (19.5)	31.4 (18.4)	30.3 (19.3)	29.3 (19.7)	2.0 (0.4, 3.5)*	3.6 (–1.3, 8.5)	5.1 (0.2, 10.0)	6.1 (1.2, 11.0)*
24 wk	34.9 (20.6)	32.5 (19.8)	33.7 (20.5)	32.1 (20.5)	0.6 (–1.0, 2.3)	1.7 (–3.4, 6.9)	0.8 (–4.4, 6.0)	2.4 (–2.9, 7.6)
39 wk	36.2 (21.0)	32.8 (21.5)	30.2 (21.7)	31.6 (21.5)	1.6 (–0.2, 3.3)	2.8 (–2.8, 8.4)	5.8 (0.5, 11.2)	4.3 (–1.2, 9.9)
52 wk	36.5 (21.8)	30.7 (22.4)	31.9 (22.5)	28.7 (20.5)	2.2 (0.4, 4.0)*	5.4 (–0.4, 11.1)	4.6 (–1.2, 10.3)	7.6 (2.0, 13.2)*
6–52 wk <sup>†</sup>	—	—	—	—	1.7 (0.4, 3.0)*	3.1 (–1.0, 7.1)	5.3 (1.2, 9.3)*	5.0 (0.8, 9.2)*
<b>Functional disability (0–100 scale)</b>								
0 wk	45.2 (21.8)	44.8 (24.0)	46.1 (23.4)	45.2 (21.8)	—	—	—	—
6 wk	27.0 (20.2)	28.5 (20.3)	25.8 (19.3)	30.1 (20.9)	0.6 (–2.3, 1.0)	1.7 (–6.8, 3.4)	1.5 (–3.7, 6.6)	–3.1 (–8.3, 2.1)
12 wk	29.2 (23.7)	24.8 (18.6)	22.0 (20.7)	23.4 (20.5)	2.0 (0.3, 3.8)*	4.2 (–1.0, 9.4)	7.5 (1.7, 13.3)*	5.8 (0.2, 11.3)
18 wk	26.1 (21.4)	23.5 (19.4)	22.1 (21.5)	22.4 (19.2)	1.3 (–0.4, 2.9)	2.2 (–2.8, 7.2)	4.1 (–1.5, 9.7)	3.6 (–1.5, 8.7)
24 wk	27.1 (25.2)	25.6 (21.7)	24.0 (20.4)	24.1 (20.3)	1.1 (–0.7, 2.9)	1.4 (–4.5, 7.2)	3.4 (–2.4, 9.3)	2.9 (–2.9, 8.8)
39 wk	26.2 (22.8)	24.5 (22.6)	21.7 (20.5)	24.1 (22.7)	0.9 (–0.9, 2.8)	1.4 (–4.3, 7.1)	4.7 (–0.7, 10.2)	2.0 (–3.9, 7.9)
52 wk	28.0 (23.7)	22.6 (22.4)	22.4 (21.2)	19.1 (18.7)	2.7 (1.0, 4.4)*	5.2 (–0.5, 10.9)	5.9 (0.1, 11.8)	8.8 (3.3, 14.4)*
6–52 wk <sup>†</sup>	—	—	—	—	0.9 (–0.3, 2.2)	1.5 (–2.1, 5.1)	3.9 (–0.0, 7.7)	2.4 (1.5, 6.3)

SMT, spinal manipulative therapy; SD, standard deviation; CI, confidence interval.

Primary end points were prespecified as pain intensity and functional disability at 12 and 24 weeks. Unadjusted group means are from original data without imputation; slopes and group differences are computed from imputed data adjusted for the baseline covariates. Positive signs of slopes and mean differences were computed to favor higher doses of manipulation. A two-tailed test of statistical significance was prespecified at the .025 to account for two primary outcomes and used for all statistical tests.

\*  $p < .025$ .

<sup>†</sup> Longitudinal profile using generalized estimating equations.

group difference was chosen to be consistent with our past studies [31,34]. All analyses were conducted with Stata 11.2 (StataCorp, College Station, TX, USA).

## Results

Participants were enrolled from March 2007 to May 2010 and followed for 1 year with the last follow-up ending May 2011. Allocation was equally spread out across clinics and providers with group assignments averaging 25%

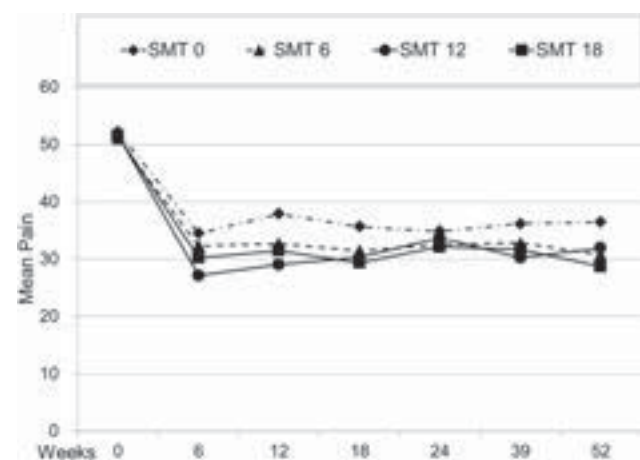


Fig. 2. Pain time profile. Pain intensity was evaluated on a 0 to 100 scale. The graphs show pain development for each group at baseline and the six follow-up time points. The primary end points were 12 and 24 weeks. SMT, spinal manipulative therapy.

(SD=6%) per group per clinic and 25% (SD=12%) per group per treating physician.

The study flowchart in Fig. 1 shows strong adherence to care with 90% to 95% of participants attending all 18 study visits. Four participants, allocated to 18 SMT sessions and who attended all 18 visits, collectively had 5 treatment visits where SMT was deemed inappropriate and withheld per protocol. Three had SMT withheld at one visit because

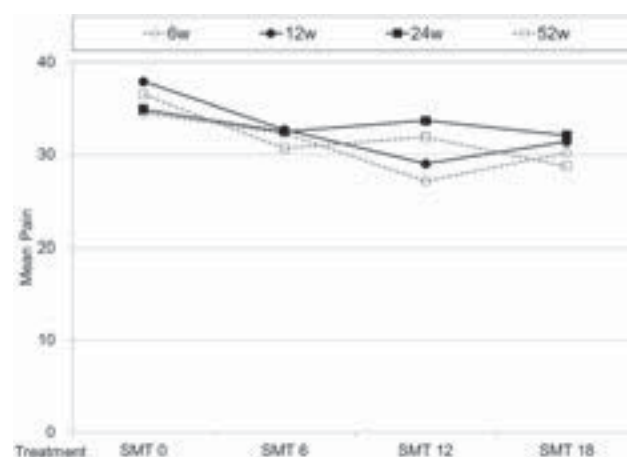


Fig. 3. Pain dose-response curves. The dose-response plots demonstrate small gradients in pain intensity (0- to 100-point scale) across dose groups for four time points: end of care (6 weeks), primary end points (12 and 24 weeks), and the final follow-up (52 weeks). Note that a line illustrates differences across the dose groups at a particular time point, rather than change over time for a particular dose. SMT, spinal manipulative therapy.

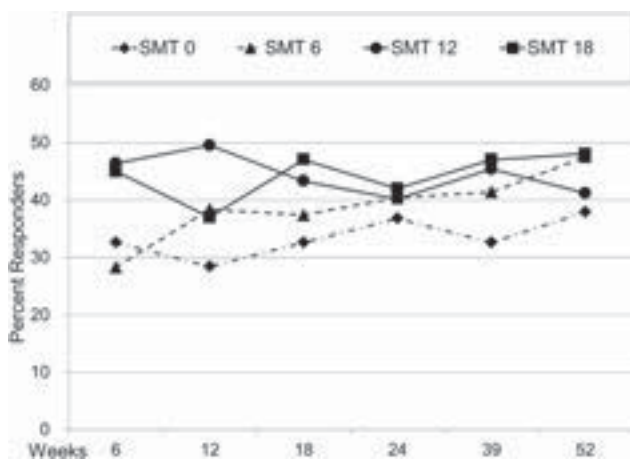


Fig. 4. Responders: percentage of individuals attaining 50% pain improvement. SMT, spinal manipulative therapy.

it was not indicated, and two received mobilization at one visit because of acute exacerbation. There was one violation of protocol where a patient accidentally received 13 SMT visits instead of 12. Compliance with data collection was greater than 80% for all follow-up time points. Nine participants were completely lost to follow-up.

Medication use and care from a nonstudy provider for cLBP were balanced across groups at each time point. During the treatment phase, 93% to 97% of participants in each treatment arm refrained from professional care outside the study and 94% to 95% abstained from prescription medication. Thereafter, approximately three-fourths reported no outside professional care at each follow-up; the maximum difference between groups ranged from 4% to 11% of participants. Also, 90% refrained from prescription medication at each follow-up with maximum group differences ranging from 1% to 8%. Nonprescription analgesics were balanced across groups. Confidence in the success of the two

interventions was approximately equal and balanced across groups.

Baseline characteristics (Table 1) were balanced across groups with the exception of smoking; inclusion of smoking in the analysis produced no substantive changes in effect sizes. The mean age was 41.3 years, and most participants were white non-Hispanic. Half of the participants reported the following characteristics: female, college degree, comorbidity, and experience with a study intervention. The mean duration of LBP was 11.8 years. The average participant experienced LBP 6 days per week and took medication for it twice per week.

### Pain

Pain improved by the end of treatment and was durable up to 52 weeks after randomization for all groups (Table 2 and Fig. 2). Mean pain reduction within groups reached more than 20 points for SMT treatment arms.

Adjusted slopes and mean differences (AMD) with confidence intervals are presented in Table 2 for the primary analysis. A small statistically significant linear dose-response effect in pain intensity across the treatment levels was observed at the 12-week primary end point (2.2 points per six visits,  $p=.007$ ) but not at the 24-week primary end point (0.6 points per six visits). Slopes were small at the other time points (1.6–2.0). Overall, there were minimal differences between adjacent dose groups at all time points (Fig. 3).

At 12 weeks, the maximum pain difference between treatment and no-SMT control was observed for 12 SMT visits (AMD=8.6,  $p=.002$ ); at 24 weeks, there were no meaningful differences from the control (AMD <2.5). For the secondary time points, a notable effect was observed at 52 weeks; here, 18 SMT visits showed the greatest advantage over the control (AMD=7.6,  $p=.011$ ).

Table 3

Responders ( $\geq 50\%$  individual improvement)

Time	Observed unadjusted % of responders				Slope (95% CI)	Adjusted difference in percentage of responders (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	Per six SMT sessions	SMT 0 vs. 6	SMT 0 vs. 12	SMT 0 vs. 18
<b>Pain intensity</b>								
6 wk	32.6	28.3	46.4	45.0	5.5 (1.3, 9.8)*	-3.7 (-16.2, 8.8)	13.7 (0.6, 26.9)	13.0 (-0.4, 26.5)
12 wk	28.4	38.4	49.5	37.0	3.8 (-0.5, 8.1)	10.0 (-3.2, 23.1)	21.1 (7.7, 34.6)*	8.9 (-4.2, 21.9)
18 wk	32.6	37.4	43.3	47.0	4.8 (0.5, 9.2)	4.3 (-9.1, 17.6)	10.3 (-3.3, 23.9)	14.2 (0.5, 27.9)
24 wk	36.8	40.4	40.2	42.0	1.4 (-2.9, 5.7)	3.7 (-10.0, 17.4)	3.2 (-10.5, 16.9)	4.9 (-8.7, 18.4)
39 wk	32.6	41.4	45.4	47.0	4.8 (0.5, 9.1)	9.0 (-4.5, 22.5)	13.0 (-0.7, 26.6)	14.6 (1.1, 28.1)
52 wk	37.9	47.5	41.2	48.0	2.5 (-1.9, 6.9)	10.2 (-3.5, 23.9)	3.9 (-9.8, 17.6)	10.6 (-3.2, 24.4)
<b>Functional disability</b>								
6 wk	49.5	47.5	50.5	42.0	-0.1 (-0.5, 0.3)	-2.0 (-16.1, 12.0)	0.9 (-13.2, 15.1)	-7.5 (-21.4, 6.5)
12 wk	43.2	49.5	59.8	55.0	4.4 (0.0, 8.8)	6.2 (-7.7, 20.0)	16.8 (2.9, 30.6)*	11.5 (-2.4, 25.4)
18 wk	50.5	54.5	63.9	56.0	2.5 (-1.9, 6.9)	3.7 (-10.4, 17.9)	13.5 (-0.3, 27.2)	3.5 (-8.8, 19.3)
24 wk	49.5	51.5	59.8	54.0	2.2 (-2.2, 6.6)	2.5 (-11.5, 16.5)	10.4 (-3.4, 24.3)	4.8 (-9.1, 18.6)
39 wk	51.6	56.6	61.9	54.0	1.2 (-3.2, 5.6)	4.9 (-9.0, 18.8)	10.4 (-3.4, 24.2)	2.5 (-11.5, 16.4)
52 wk	58.9	57.6	57.7	62.0	0.8 (-3.5, 5.2)	-1.1 (-14.8, 12.6)	-1.4 (-15.4, 12.6)	2.7 (-11.0, 16.5)

SMT, spinal manipulative therapy; CI, confidence interval.

Unadjusted group percentages are from original data without imputation; slopes and group differences in responders are computed from imputed data adjusted for the baseline covariates. Positive signs of slopes and mean differences were computed to favor higher doses of SMT.

\*  $p<.025$ .

Table 4  
Secondary outcomes

Time	Observed unadjusted mean (SD)				Slope (95% CI)	Adjusted mean difference (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	Per six SMT sessions	SMT 0 vs. 6	SMT 0 vs. 12	SMT 0 vs. 18
Pain unpleasantness (0–100 scale)								
0 wk	41.7 (19.5)	41.1 (21.1)	40.3 (22.8)	42.4 (22.2)	—	—	—	—
6 wk	24.7 (21.7)	20.3 (17.5)	15.6 (17.3)	18.4 (18.1)	2.4 (0.9, 4.0)*	4.2 (−0.8, 9.1)	8.7 (3.5, 14.0)*	6.6 (1.7, 11.6)*
12 wk	29.7 (22.3)	24.5 (20.6)	22.2 (22.3)	23.9 (21.3)	2.0 (0.2, 3.7)	4.8 (−1.0, 10.5)	7.1 (1.1, 13.0)*	5.8 (0.1, 11.5)
18 wk	26.8 (22.3)	21.9 (21.2)	20.4 (20.4)	21.3 (21.8)	1.8 (0.1, 3.6)	4.6 (−1.0, 10.3)	5.8 (0.4, 11.2)	5.8 (0.4, 11.2)
24 wk	27.4 (21.7)	24.8 (21.8)	25.0 (21.9)	24.8 (22.4)	0.7 (−1.1, 2.5)	2.1 (−3.5, 7.8)	2.0 (−3.6, 7.5)	2.5 (−3.2, 8.3)
39 wk	28.3 (21.6)	25.5 (21.3)	23.2 (20.8)	23.0 (21.1)	1.8 (0.1, 3.6)	2.4 (−3.3, 8.2)	4.8 (−0.6, 10.3)	5.4 (−0.2, 10.9)
52 wk	27.8 (22.1)	22.1 (21.7)	24.0 (24.0)	21.5 (19.9)	1.7 (−0.1, 3.5)	5.4 (−0.3, 11.2)	3.7 (−2.4, 9.7)	6.4 (0.7, 12.1)
Days with pain (last 4 wk)								
0 wk	24.8 (4.8)	24.1 (5.5)	23.3 (5.7)	24.1 (4.6)	—	—	—	—
6 wk	18.9 (10.0)	17.6 (10.6)	15.0 (10.1)	16.3 (10.8)	0.8 (−0.0, 1.7)	0.9 (−1.8, 3.5)	3.0 (0.3, 5.6)	2.2 (−0.6, 4.9)
12 wk	18.1 (9.5)	17.4 (9.2)	15.5 (9.7)	14.7 (10.2)	1.1 (0.2, 1.9)*	0.4 (−2.0, 2.8)	1.9 (−0.7, 4.5)	3.0 (0.5, 5.6)*
18 wk	17.4 (10.2)	17.6 (10.3)	15.5 (10.3)	14.9 (10.6)	0.8 (−0.0, 1.7)	−0.6 (−3.2, 2.1)	1.2 (−1.5, 3.9)	2.2 (−0.6, 4.9)
24 wk	16.9 (10.0)	16.5 (10.3)	15.4 (9.5)	13.5 (9.6)	1.0 (0.2, 1.8)*	0.1 (−2.5, 2.7)	0.7 (−1.8, 3.3)	3.1 (0.5, 5.7)*
39 wk	17.7 (9.8)	15.6 (9.9)	15.1 (10.4)	14.3 (10.6)	0.9 (0.1, 1.7)	1.7 (−0.8, 4.2)	1.7 (−0.8, 4.3)	3.0 (0.4, 5.6)
52 wk	17.0 (10.2)	15.3 (10.3)	14.0 (10.6)	13.6 (10.4)	1.0 (0.1, 1.8)*	1.3 (−1.4, 3.9)	2.1 (−0.7, 4.9)	3.0 (0.3, 5.7)
Days with disability (last 4 wk)								
0 wk	7.4 (8.1)	6.7 (7.5)	6.8 (7.5)	6.5 (7.2)	—	—	—	—
6 wk	2.4 (5.5)	1.6 (3.0)	1.3 (2.7)	1.2 (3.4)	0.3 (−0.0, 0.7)	0.8 (−0.4, 1.9)	1.0 (−0.1, 2.2)	1.1 (−0.1, 2.3)
12 wk	3.4 (5.9)	2.6 (4.9)	2.0 (3.3)	2.3 (3.7)	0.4 (−0.0, 0.7)	0.6 (−0.7, 1.9)	1.3 (0.0, 2.5)	1.0 (−0.3, 2.2)
18 wk	2.1 (5.3)	2.2 (5.0)	1.8 (4.1)	2.1 (4.9)	0.1 (−0.4, 0.5)	−0.1 (−1.4, 1.3)	0.3 (−0.9, 1.6)	0.0 (−1.3, 1.4)
24 wk	3.5 (6.7)	2.7 (5.1)	2.8 (4.6)	2.8 (4.2)	0.1 (−0.3, 0.6)	0.7 (−0.7, 2.1)	0.6 (−0.8, 2.1)	0.5 (−0.9, 1.9)
39 wk	2.8 (5.1)	3.0 (5.5)	2.4 (4.7)	2.5 (4.8)	0.1 (−0.3, 0.5)	−0.3 (−1.7, 1.1)	0.4 (−0.9, 1.6)	0.2 (−1.1, 1.5)
52 wk	3.4 (6.0)	3.4 (6.9)	1.9 (3.7)	2.4 (4.7)	0.4 (−0.1, 0.8)	−0.1 (−1.7, 1.4)	1.5 (0.2, 2.8)	0.8 (−0.6, 2.1)
Perceived pain change (six-point Likert) <sup>†</sup>								
6 wk	3.7 (0.9)	4.1 (0.9)	4.3 (1.0)	4.3 (1.0)	0.1 (0.1, 0.2)*	0.2 (−0.0, 0.5)	0.4 (0.2, 0.7)*	0.4 (0.1, 0.7)*
12 wk	3.7 (0.9)	3.8 (0.9)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)	0.1 (−0.1, 0.4)	0.3 (−0.0, 0.5)	0.3 (0.0, 0.5)
18 wk	3.7 (0.9)	4.0 (1.0)	4.1 (1.0)	4.0 (1.0)	0.1 (−0.0, 0.2)	0.2 (−0.0, 0.5)	0.4 (0.1, 0.7)*	0.2 (−0.0, 0.5)
24 wk	3.8 (1.0)	3.9 (1.0)	3.9 (1.0)	3.9 (1.0)	0.0 (−0.1, 0.1)	0.1 (−0.2, 0.4)	0.1 (−0.2, 0.4)	0.1 (−0.2, 0.3)
39 wk	3.7 (1.0)	3.8 (1.1)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)	0.1 (−0.2, 0.4)	0.3 (0.0, 0.6)	0.3 (0.0, 0.5)
52 wk	3.7 (0.9)	4.0 (1.1)	3.9 (1.1)	4.0 (0.9)	0.1 (−0.0, 0.1)	0.3 (0.0, 0.6)	0.2 (−0.1, 0.5)	0.2 (−0.0, 0.5)
Perceived disability change (six-point Likert) <sup>†</sup>								
6 wk	3.6 (0.9)	3.9 (0.9)	4.0 (1.0)	4.0 (0.9)	0.1 (0.0, 0.2)*	0.3 (0.0, 0.5)	0.4 (0.1, 0.6)*	0.4 (0.2, 0.6)*
12 wk	3.5 (0.9)	3.8 (0.9)	3.8 (1.0)	3.8 (0.9)	0.1 (0.0, 0.2)*	0.2 (0.0, 0.5)	0.3 (0.0, 0.5)	0.3 (0.0, 0.5)
18 wk	3.6 (0.8)	3.7 (1.0)	3.8 (1.0)	3.8 (0.9)	0.1 (−0.0, 0.2)	0.1 (−0.2, 0.3)	0.2 (−0.1, 0.4)	0.2 (−0.0, 0.5)
24 wk	3.6 (0.9)	3.7 (0.9)	3.7 (0.9)	3.7 (1.0)	0.0 (−0.1, 0.1)	0.1 (−0.2, 0.3)	0.1 (−0.2, 0.3)	0.1 (−0.2, 0.4)
39 wk	3.6 (0.9)	3.7 (1.0)	3.9 (1.0)	3.9 (0.9)	0.1 (0.0, 0.2)*	0.1 (−0.2, 0.4)	0.3 (0.1, 0.6)*	0.3 (0.0, 0.5)*
52 wk	3.6 (0.9)	3.8 (1.0)	3.8 (1.1)	3.9 (0.9)	0.1 (−0.0, 0.2)	0.2 (−0.0, 0.5)	0.2 (−0.1, 0.5)	0.3 (0.0, 0.3)
SF-12 physical health component								
0 wk	43.0 (9.5)	43.8 (8.9)	44.3 (8.4)	42.3 (8.8)	—	—	—	—
12 wk	45.5 (10.3)	47.1 (8.2)	49.6 (8.5)	47.5 (8.5)	1.0 (0.3, 1.6)*	1.2 (−0.8, 3.3)	3.5 (1.3, 5.7)*	2.4 (0.3, 4.6)
24 wk	50.0 (11.1)	50.5 (10.1)	51.4 (9.1)	50.9 (9.4)	0.5 (−0.3, 1.2)	0.0 (−2.3, 2.4)	0.8 (−1.6, 3.2)	1.3 (−1.1, 3.6)
39 wk	50.6 (11.5)	51.1 (10.3)	52.7 (9.6)	51.8 (9.3)	0.6 (−0.2, 1.4)	0.1 (−2.3, 2.4)	1.5 (−0.9, 4.0)	1.6 (−0.9, 4.0)
52 wk	50.7 (12.0)	50.8 (11.0)	52.6 (10.3)	52.5 (8.5)	0.8 (0.1, 1.6)	0.3 (−2.7, 2.1)	1.4 (−1.2, 4.0)	2.2 (−0.2, 4.5)
SF-12 mental health component								
0 wk	50.2 (10.5)	48.6 (10.5)	47.6 (11.2)	49.4 (9.6)	—	—	—	—
12 wk	50.2 (10.8)	50.4 (9.4)	47.8 (11.0)	48.0 (9.8)	−0.7 (−1.3, −0.0)	1.3 (−0.7, 3.3)	−0.5 (−2.6, 1.5)	−1.5 (−3.5, 0.4)
24 wk	51.8 (10.9)	52.8 (10.2)	50.8 (11.8)	51.3 (11.2)	−0.1 (−0.8, 0.6)	2.1 (−0.0, 4.2)	0.7 (−1.3, 2.8)	0.1 (−2.1, 2.2)
39 wk	51.7 (11.3)	51.5 (11.6)	49.2 (13.6)	49.0 (11.7)	−0.8 (−1.5, −0.0)	1.1 (−1.3, 3.4)	−0.4 (−2.9, 2.1)	−2.0 (−4.3, 0.3)
52 wk	51.3 (12.0)	50.4 (11.4)	50.6 (12.7)	50.4 (11.7)	−0.0 (−0.8, 0.8)	0.2 (−2.3, 2.7)	1.1 (−1.6, 3.7)	−0.3 (−2.9, 2.3)
EuroQol Health State (0–100 scale)								
0 wk	70.1 (17.2)	72.1 (13.8)	73.5 (14.4)	68.2 (17.4)	—	—	—	—
12 wk	73.5 (17.3)	78.4 (14.1)	77.9 (15.0)	75.7 (14.5)	0.8 (−0.4, 2.0)	3.9 (0.0, 7.7)	2.7 (−1.3, 6.7)	3.1 (−0.8, 6.9)
24 wk	73.9 (17.5)	77.8 (15.5)	77.0 (15.4)	74.5 (16.7)	0.3 (−1.0, 1.6)	2.9 (−1.0, 6.9)	1.4 (−2.6, 5.5)	1.5 (−2.7, 5.8)
39 wk	73.1 (20.0)	76.8 (17.2)	76.6 (15.6)	75.3 (16.8)	0.9 (−0.5, 2.3)	2.6 (−1.9, 7.0)	1.8 (−2.6, 6.2)	3.2 (−1.3, 7.6)
52 wk	74.8 (17.0)	77.1 (17.0)	77.3 (15.3)	77.2 (14.9)	1.0 (−0.3, 2.2)	1.3 (−2.7, 5.4)	0.9 (−3.1, 4.9)	3.3 (−0.5, 7.2)
Nonprescription medication (times in last 4 wk)								
0 wk	7.6 (10.0)	8.9 (10.8)	9.5 (10.0)	7.6 (9.4)	—	—	—	—
6 wk	4.3 (6.8)	4.1 (6.9)	4.0 (6.7)	4.0 (8.3)	0.1 (−0.5, 0.7)	0.6 (−1.2, 2.4)	0.9 (−0.8, 2.6)	0.3 (−1.5, 2.1)
12 wk	7.8 (11.0)	7.1 (23.0)	5.8 (7.6)	7.8 (25.8)	0.1 (−1.6, 1.8)	1.4 (−2.7, 5.6)	3.2 (0.4, 5.9)*	−0.1 (−5.5, 5.3)

(Continued)



Table 4  
(Continued)

Time	Observed unadjusted mean (SD)				Slope (95% CI)	Adjusted mean difference (95% CI)		
	SMT 0	SMT 6	SMT 12	SMT 18	Per six SMT sessions	SMT 0 vs. 6	SMT 0 vs. 12	SMT 0 vs. 18
18 wk	5.7 (7.4)	6.0 (10.4)	5.3 (9.1)	5.5 (9.1)	0.2 (−0.5, 0.8)	0.1 (−2.4, 2.6)	1.0 (−1.1, 3.2)	0.2 (−1.8, 2.2)
24 wk	8.1 (13.7)	7.7 (15.8)	6.5 (8.6)	6.4 (10.8)	0.7 (−0.4, 1.7)	0.9 (−3.3, 5.1)	2.4 (−0.7, 5.5)	1.8 (−1.5, 5.0)
39 wk	7.2 (14.2)	7.4 (12.2)	6.6 (8.9)	7.7 (13.5)	−0.0 (−1.1, 1.1)	0.3 (−3.4, 4.0)	1.4 (−1.9, 4.7)	−0.5 (−4.0, 3.1)
52 wk	6.5 (8.0)	6.6 (13.7)	6.7 (9.3)	6.8 (12.7)	−0.1 (−0.9, 0.8)	0.3 (−2.7, 3.4)	0.6 (−1.5, 2.6)	−0.3 (−2.9, 2.4)

SMT, spinal manipulative therapy; SD, standard deviation; CI, confidence interval; SF-12, short-form 12.

Unadjusted group means are from original data without imputation; slopes and group differences are computed from imputed data adjusted for the baseline covariates. Positive signs of slopes and mean differences were computed to favor higher doses of SMT.

\*  $p < .025$ .

† Likert scale: much worse=1, worse=2, about the same=3, better=4, much better=5, and completely recovered=6.

Analyses of the three pain score components gave similar results and are available from the authors.

Repeated-measures analysis of the full 6- to 52-week pain profile demonstrated a small advantage of treatment over control. The largest effect was observed for 12 SMT visits versus control (AMD=5.3,  $p=.011$ ).

#### Functional disability

Mean functional disability reduction reached 20 points. Trends in slopes and group comparisons were similar to those for pain but smaller in magnitude with fewer statistically significant results (Table 2). At the 12-week primary end point, the greatest advantage for SMT over control was also found for 12 SMT visits (AMD=7.5,  $p=.011$ ), and at the 24-week primary end point, there were no clinically meaningful effects (AMD <3.4). At 52 weeks, 18 SMT visits were observed to have the greatest effect (AMD=8.8,  $p=.002$ ). As for pain, there were no clinically meaningful differences in disability profiles between 12 and 18 SMT visits.

#### Sensitivity analysis

There were no material changes in the results for pain and disability outcomes when imputed data were excluded from the analysis. For the primary end points, changes were 0.3 or less in slope and 1.4 or less/100 points in group differences. The changes at other time points were similarly negligible. Clustering by care provider or clinic produced no substantive changes in effect sizes.

#### Responder analysis

The responder profile in Fig. 4 and Table 3 shows that about 30% to 50% of individuals in each group achieved 50% pain improvement at each time point. Only one statistically significant difference between treatment and control was found. At 12 weeks, a substantial proportion of response to care was attributable to manipulation for 12 SMT visits (21.1%,  $p=.002$ ). This difference corresponds to a number needed to treat equal to 5. For functional disability, about 40% to 60% of individuals were responders

for all groups and time points. There was only one statistically significant group difference.

#### Secondary outcomes

Generally, there was within-group improvement in secondary outcomes recorded at the end of care showing the same durability for these outcomes as for pain outcomes. However, the improvement in the no-SMT control group was of such magnitude that there were few sizable statistically significant differences between treatment and control groups (Table 4).

Days with pain and disability were reduced from baseline by 1 to 2 per week. Perceived pain and disability improvement were typically rated as “better.” The standardized short-form 12 physical health component improved about 7 to 10 points (up to 1 SD), returning to US population norms in 3 to 6 months. The mental health component deviated little from population norms at baseline. EuroQol Health State Visual Analog Scale showed little change from baseline. There was a small decrease in medication use after end of care. The mean reductions in pain unpleasantness scores were similar to pain score reduction, about 20 points.

#### Adverse events

There were no notable adverse events. Three persons reported seeking care for symptomatic relief of LBP exacerbation related to the study. One ineligible person subsequently reported increase of pain after the screening examination. One participant in the 12-SMT group lost several days of work followed by complete resolution or the episode during the treatment phase. One participant in the 12-SMT group dropped out after an exacerbation associated with lifting a child.

#### Discussion

This first full-scale dose-response study of SMT had several notable findings. Based on the pain and functional disability primary outcomes, 12 sessions of SMT yielded the overall best, albeit modest, treatment effects (group

differences). This was particularly noted in the short term at the 12-week primary end point. Group differences were negligible at the 24-week primary end point and favored 18 SMT sessions to a small degree in the long term, at 52 weeks. In general, the data were consistent with a dose-response relationship being saturated at 12 sessions with little or no additional benefit attributable to additional SMT visits, even at 52 weeks. Analysis of the full-time profile supported no additional benefit overall of 18 over 12 sessions. In addition, responder analysis gave additional support for some advantage of 12 visits but only in the short term.

The linear dose-response gradients for the primary outcomes were small in general, reaching approximately 2/100 scale points per six sessions of SMT at 12 and 52 weeks. Even excluding the highest dose group for short-term results, the gradient would only double to about 4/100 scale points per six SMT sessions. The fact that there was little difference between adjacent dose groups makes it difficult to recommend one treatment dose over another. However, two considerations come into play. First, the effects across dose accumulate to modest benefit of SMT above the hands-on control. Second, an aim of the study was to find a saturation dose level for use in future studies. The time profiles, dose-response gradients, and comparisons with the control group suggest in aggregate that 12 visits would be the best choice, particularly for short-term improvement.

Interpretation of the dose-response effects requires consideration of several factors. This was a fastidious randomized trial designed to isolate the effects of SMT. We controlled number of visits, time with the participant, effects of touching the patient, patient-provider interaction, and intervention credibility. This was accomplished with 18 visits of hands-on therapy and electronic modality (minimal ultrasound) for all groups. The specific and contextual effects of light massage at non-SMT visits, ultrasound, or simply 18 visits to a health-care provider potentially obfuscated a larger dose-response gradient that might be found in clinical practice. For example, such larger effects were seen in our pilot study where participants attended only visits for the active intervention [5].

In terms of efficacy, the light massage control is technically a comparison intervention rather than a true sham. Many sessions with even a minimal massage may have more effect than one might expect. As such, the differences between SMT and the control may be somewhat smaller than for a comparison with a sham manipulation. We did not attempt to use a sham for two reasons. First, it would be virtually impossible to blind participants because half received visits for both treatment and control and could compare interventions. Second, we wanted to avoid some disappointment that can arise when participants think they may be receiving sham intervention.

All participants were scheduled to receive their assigned dose of SMT. There were no treatment stoppage rules based on improvement during the care period. The effects of care

stoppage are unknown and could be either beneficial or detrimental to outcomes in the short and long terms.

Another issue is the threshold of a clinically important difference between groups for the continuous variable primary outcomes. Studies on patient-rated minimal important change have led some authors to conclude that 30% improvement (about 15–20/100 points) can be considered a robust indicator of within-person minimal clinically important change for these outcomes [35]. A 50% improvement has been recommended as a success threshold for the individual [36]. However, Dworkin et al. [37] point out that these numbers do not apply to between-group effects, and identifying meaningful group differences is a multifactorial process that is far from straightforward. The between-group differences of 8.6 in pain and 7.5 in disability scores at a primary end point are certainly marginal, but it is not clear yet whether effects of this magnitude constitute a degree of clinical relevance. The associated number needed to treat for pain (=5) may actually indicate a meaningful effect [12,38].

## Conclusions

Overall, 12 sessions of spinal manipulation in 6 weeks from a chiropractor yielded the most favorable pain and functional disability improvement for chronic nonspecific LBP. Mean participant improvement for this group was substantial at the end of care and sustainable to 52 weeks. Approximately, half of patients would be expected to achieve 50% improvement in pain/disability. Therefore, 12 sessions of SMT is the current best estimate for use in comparative effectiveness trials. However, the recommendation is made with caution because the gradient of treatment effects across dose groups was too small to clearly distinguish 12 visits from adjacent dose levels. Even with 12 visits, the contribution of SMT to outcomes beyond that of a focused light massage delivered by a chiropractor (hands-on control) was at best modest at the 12-week primary end point and negligible at the 24-week primary end point.

## Acknowledgments

This study was funded by the National Center for Complementary and Alternative Medicine (NCCAM), National Institutes of Health (U01 AT001908). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the official views of NCCAM. The authors declare no conflicts of interest.

## References

- [1] Andersson GBJ. Epidemiological features of chronic low-back pain. *Lancet* 1999;354:581–5.
- [2] Leigh JP, Markowitz SB, Fahs M, et al. Occupational injury and illness in the United States. *Arch Intern Med* 1997;157:1557–68.
- [3] Shekelle P, Adams A, Chassin M, et al. Spinal manipulation for low-back pain. *Ann Intern Med* 1992;117:590–8.

- [4] Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21:2820–5.
- [5] Haas M, Group E, Kraemer DF. Dose-response for chiropractic care of chronic low back pain. *Spine J* 2004;4:574–83.
- [6] Shekelle PG, Adams AH, Chassin MR, et al. The appropriateness of spinal manipulation for low-back pain: project overview and literature review. Santa Monica, CA: RAND, 1991.
- [7] Shekelle PG, Adams AH, Chassin MR, et al. The appropriateness of spinal manipulation for low-back pain: indications and ratings by a multidisciplinary expert panel. Santa Monica, CA: RAND, 1991.
- [8] Shekelle PG, Adams AH, Chassin MR, et al. The appropriateness of spinal manipulation for low-back pain: indications and ratings by an all-chiropractic expert panel. Santa Monica, CA: RAND, 1992.
- [9] Nyiendo J, Haas M, Goldberg B, Sexton G. Patient characteristics and physicians' practice activities for chronic low back pain patients: a practice-based study of primary care and chiropractic physicians. *J Manipulative Physiol Ther* 2001;24:92–100.
- [10] Bronfort G, Haas M, Evans R, et al. Evidence-informed management of chronic low back pain with spinal manipulation and mobilization. *Spine J* 2008;8:213–25.
- [11] Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians Clinical Practice Guideline. *Ann Intern Med* 2007;147:492–504.
- [12] Rubinstein SM, van Middelkoop M, Assendelft WJ, et al. Spinal manipulative therapy for chronic low-back pain: an update of a Cochrane review. *Spine* 2011;36:E825–46.
- [13] Kaptchuk TJ. The placebo effect in alternative medicine: can the performance of a healing ritual have clinical significance? *Ann Intern Med* 2002;136:817–25.
- [14] Haas M, Spegman A, Peterson DH, et al. Dose-response and efficacy of spinal manipulation for chronic cervicogenic headache: a pilot randomized controlled trial. *Spine J* 2010;10:117–28.
- [15] Begg CB, Iglewicz B. A treatment allocation procedure for sequential clinical trials. *Biometrics* 1980;36:81–90.
- [16] Aickin M. Randomization, balance, and the validity and efficiency of design—adaptive allocation methods. *J Stat Plan Inference* 2001;94: 97–119.
- [17] Gatterman MI, Panzer DM. Disorders of the lumbar spine. In: Gatterman MI, ed. *Chiropractic management of spine related disorders*. Baltimore, MD: Williams & Wilkins, 1990:129–75.
- [18] Anderson JAD. Problems of classification of low-back pain. *Rheumatol Rehabil* 1977;16:34–6.
- [19] DiFabio RP, Mackey G, Holte JB. Physical therapy outcomes for patients receiving workers' compensation following treatment for herniated lumbar disc and mechanical low back pain syndrome. *J Orthop Sports Phys Ther* 1996;23:180–7.
- [20] Belanger A. Ultrasound. In: Belanger A, ed. *Evidence-based guide to therapeutic physical agents*. Philadelphia, PA: Lippincott Williams & Wilkins, 2002:223–61.
- [21] Cherkin DC, Deyo RA, Sherman KJ, et al. Characteristics of licensed acupuncturists, chiropractors, massage therapists and naturopathic physicians. *J Am Board Fam Pract* 2002;15:463–72.
- [22] Peterson DH, Bergmann TF. *Chiropractic technique: principles and practice*. 2nd ed. St Louis, MO: Mosby, 2002.
- [23] Nicholson GG, Clendaniel RA. Manual techniques. In: Scully RM, Barnes MR, eds. *Physical therapy*. Philadelphia, PA: J.B. Lippincott Company, 1989:926–85.
- [24] Furlan AD, Brosseau L, Imamura M, Irvin E. Massage for low back pain: a systematic review within the framework of the Cochrane collaboration back review group. *Spine* 2002;27:1896–910.
- [25] Haas M, Aickin M, Vavrek D. A path analysis of expectancy and patient-provider encounter in an open-label randomized controlled trial of spinal manipulation for cervicogenic headache. *J Manipulative Physiol Ther* 2010;33:5–13.
- [26] Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. *Spine* 1999;24:1104–12.
- [27] Price DD. *Psychological mechanisms of pain and analgesia*. Seattle, WA: IASP Press, 1999.
- [28] Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.
- [29] Luo N, Johnson JA, Shaw JW, et al. Self-reported health status of the general adult U.S. population as assessed by the EQ-5D and Health Utilities Index. *Med Care* 2005;43:1078–86.
- [30] Waddell G, Newton M, Henderson I, et al. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157–68.
- [31] Nyiendo J, Haas M, Goldberg B, Sexton G. Pain, disability and satisfaction outcomes and predictors of outcomes: a practice-based study of chronic low back pain patients attending primary care and chiropractic physicians. *J Manipulative Physiol Ther* 2001;24: 433–9.
- [32] Zellner A. An efficient method of estimating seeming unrelated regressions and tests for aggregation bias. *J Am Stat Assoc* 1962;57: 348–68.
- [33] StataCorp. *Stata: release 11.2 statistical software*. College Station, TX: Stata Corporation, 2009.
- [34] Haas M, Group E, Panzer D, et al. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. *Spine* 2003;28: 1091–6.
- [35] Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine* 2008;33: 90–4.
- [36] Fritz JM, Hebert J, Koppenhaver S, Parent E. Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain. *Spine* 2009;34:2803–9.
- [37] Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009;146:238–44.
- [38] Froud R, Eldridge S, Lall R, Underwood M. 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* 2009;9:35.