Title

Do Short-Term Effects Predict Long-Term Improvements in Women who Receive Manual Therapy or Surgery for Carpal Tunnel Syndrome? A Bayesian Network Analysis of a Randomized Clinical Trial

Authors

Bernard X W Liew¹ PT, PhD; Ana I. de-la-Llave-Rincón² PT, PhD; Marco Scutari PhD³; José L. Arias-Buría² PT, PhD; Chad E. Cook⁴ PT, PhD; Joshua Cleland⁵ PT, PhD; César Fernández-de-las-Peñas² PT, PhD

Affiliations

(1) School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, United Kingdom.
(2) Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Universidad Rey Juan Carlos, Alcorcón, Spain
Istituto Dalle Molle di Studi Sull’Intelligenza Artificiale (IDSIA), Lugano, Switzerland
(3)(4) Doctor of Physical Therapy Program, Duke University, Durham, NC, USA.
(4)(5) Doctor of Physical Therapy Program, Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, Mass, USA.

Address for reprint requests / corresponding author:
César Fernández de las Peñas
Facultad de Ciencias de la Salud
Universidad Rey Juan Carlos
Avenida de Atenas s/n
28922 Alcorcón, Madrid, SPAIN
Telephone number: + 34 91 488 88 84
Fax number: + 34 91 488 89 57
E-mail address: cesar.fernandez@urjc.es

Conflicts of interest: None declared

Sources of funding: No funds were received for this study

Trial registration: ClinicalTrials.gov: NCT01789645
Abstract

**Background:** We present a data-driven Bayesian Network approach to understand the potential multivariate pathways of the effect of manual physical therapy in carpal tunnel syndrome (CTS).

**Methods:** We analysed data from a randomised clinical trial (n=104) comparing manual physical therapy including desensitization manoeuvres of the central nervous system versus surgery in women with CTS. A range of potential variables was included in a Bayesian Network to explore its multivariate relationship. The model was used to quantify the direct and indirect pathways of the effect of physical therapy and surgery on short-, mid- and long-term changes in the clinical variables of pain, related-function, and symptoms’ severity.

**Results:** Manual physical therapy reduced related-disability and pain in women with CTS. The BN showed that early improvements (one-month after) in function and symptom severity led to long-term (twelve-months after) changes in related-disability both directly and via complex pathways involving baseline pain intensity and depression levels. Additionally, the severity of the condition also had a direct influence on long-term changes of related-disability.

**Conclusions:** Current findings suggest that short-term benefits in function and symptoms’ severity are associated with long-term improvement in related-disability but mechanisms driving these effects interact with depressive levels and EMG severity. Further data-driven analyses involving a broad range of biopsychosocial variables are recommended to fully understand the pathways underpinning CTS treatment effects.

**Keywords:** Carpal tunnel syndrome, manual therapy, Bayesian Networks, Machine learning.
Do Short-Term Effects Predict Long-Term Improvements in Women who Receive Manual Therapy or Surgery for Carpal Tunnel Syndrome? A Bayesian Network Analysis of a Randomized Clinical Trial

Introduction

Carpal tunnel syndrome (CTS) is the most prevalent entrapment neuropathy of the upper extremity. Although there is no consensus on which treatment option should be applied as first-line management, surgery and conservative treatment are approaches commonly recommended by clinical practice guidelines (1, 2). This lack of consensus reflects the finding that differences between conservative and surgical treatment are smaller than expected and most patients want to avoid surgery (3).

Recent theories support that CTS involves peripheral and central sensitization processes, suggesting that conservative treatment should also include interventions targeting the central nervous system, in addition to the peripheral nervous system (4). A randomized clinical trial conducted by Fernández-de-las-Peñas et al showed that a manual physical therapy intervention including desensitization maneuvers of the central nervous system obtained better short-term and similar long-term effects on pain intensity and related-function compared with carpal tunnel surgery in women with CTS (5). An economic analysis of the same trial revealed that this manual physical therapy approach was equally effective clinically but less costly compared to surgery (6).

Whereas both treatments are effective, understanding the potential pathways, or “why” a treatment is effective, could be of benefit to clinicians and patients. Where treatment is ineffective, a pathway analysis would reveal where a hypothesized variable broke down, and potentially where the intervention needs to be strengthened or where clinicians should intercede with more personalized management. To the best of the authors’ knowledge, there is only one path analysis in a cross-sectional cohort including people with CTS (7).
study reported that function partially mediated the effects of depressive levels and symptoms’ severity relationships to pain (7). No previous studies have investigated the mediating mechanisms by which different treatments influence short- to long-term clinical outcomes in individuals with CTS.

Fernández-de-las-Peñas et al. (7) used structural equations modelling (SEM) for path analysis. SEM is ideal for investigating the validity of a hypothesized structural path model of disease mechanism. However, SEM does not consider whether there could be competing path models that better explain the disorder. For example, it is well-established that experimentally inducing pain can result in altered motor function (8). In addition, although depression can drive pain (9), pain may also drive the presence of depressive levels, as has been shown in individuals with whiplash disorders (10). Understanding all competing paths will improve our understanding of how interventions provide an effect.

Bayesian Networks (11) are a class of probabilistic models that provides a data-driven approach to derive complex pathways of effects, that may or may not include structural assumptions in the form of prior knowledge. This approach has previously been used to investigate the pathways of effect for some musculoskeletal disorders including whiplash (12) and postoperative cervical radiculopathy (13). Since several pathways of effect could have underpinned the benefits of physical therapy and surgery in individuals with CTS, this study aimed to use BN to explore the pathways of the effect of manual physical therapy and surgery for women with CTS treated in the previous randomized clinical trial (5). We hypothesize that changes in function would lie on the path of symptoms’ severity to pain intensity (7).
Methods

Study Design
A preplanned secondary analysis was conducted alongside a randomized clinical trial with a one-year evaluation primary end-point performed in an Urban Hospital in Madrid (Spain) (5). Full details of the clinical trial, participants, interventions, and primary results of the clinical outcomes have been previously reported (5). The design was approved by the Hospital Universitario Fundación Alcorcón Ethics Committee (PI01223-HUFA12/14) and the clinical trial was prospectively registered (ClinicalTrials.gov: NCT01789645).

Participants
Full details of participant selection are published in the original trial (5). Briefly, women with pain and/or paresthesia in the median nerve distribution, positive Tinel or Phalen signs, and deficits of sensory or motor median nerve conduction on electro-diagnostic examination (14) were included. They were excluded if patients had any of the following: 1) motor or sensory deficits in the ulnar/radial nerves; 2) hand surgery; 3) use of steroid injections for CTS; 4) multiple diagnoses on the upper extremity such as cervical radiculopathy; 5) previous neck, shoulder, or upper extremity trauma; 6) systemic diseases causing CTS such as diabetes mellitus; 7) underlying medical conditions altering pain perception such as fibromyalgia; or 8) pregnancy. Participants signed the written informed consent form before their inclusion.

Randomization and Interventions
Participants were randomly assigned to receive either manual physical therapy or surgery as previously described (5). Those allocated to the manual therapy group received 3 treatment sessions of 30-min duration which included desensitization maneuvers of the central nervous system, once/week. Briefly, the desensitization maneuvers included soft
tissue mobilization techniques targeting anatomical sites of potential entrapment of the median nerve such as the scalene, pectoralis minor, biceps brachii, bicipital aponeurosis, pronator teres, wrist flexors, transverse carpal ligament, palmar aponeurosis, or lumbrical muscles. Additionally, lateral glides were applied to the cervical spine, and tendon and nerve gliding exercises targeting the median nerve were also applied (5). Finally, all participants received an educational session on performing the tendon/nerve gliding exercises as home exercises. A complete description of the intervention can be found elsewhere (5). Those randomly allocated to the surgery group underwent open or endoscopic release of the carpal tunnel, pragmatically applied based on the surgeon and patient preferences (15). Patients allocated to this group also received the same educational session for performing the tendon/nerve gliding exercises as the manual physical therapy group (5).

**Variables included in the Bayesian Network**

Clinical outcomes on the original trial were assessed at baseline, and 1, 3, 6, and 12 months after the intervention (5). The primary outcome was pain intensity assessed with an 11-point Numerical Pain Rating Scale (NPRS, 0: no pain; 10: maximum pain). In the original trial, the mean intensity of pain and the worst level of pain experienced in the preceding week were assessed independently, but in the current Bayesian network, the worst pain score was used. Secondary outcomes included the functional status and symptoms’ severity subscales of the Boston Carpal Tunnel Questionnaire (BCTQ) (16). Each scale is scored from 1 to 5, with higher scores indicating poor function or more symptoms’ severity.

Accordingly, we included 21 variables in the BN. The 21 variables were categorized into: 1) baseline patient characteristics, 2) treatment group, and 3) clinical outcomes. Baseline patient characteristics included age (years), duration of the symptoms (years with pain), area (pain extent assessed with digital pain draws), electromyography (EMG, classified as minimal, moderate, or severe), depressive levels (assessed with The Beck Depression
Inventory -BDI II, score ranging from 0 to 21 which higher scores suggestive of higher depressive levels). Treatment group included the randomized allocation to treatment (1-manual therapy, 2 -surgery) from the original clinical trial. Clinical Outcomes included pain intensity, function, and symptoms’ severity collected at baseline, 3, 6, and 12 months follow-up (3 outcomes * 4 times = 12 variables).

**Approach to Data Analysis**

Bayesian network analysis

All analyses were performed in R software (17) using the “bnlearn” package (18), with codes and results included in a public online repository [https://bernard-liew.github.io/2020_cts_bn/2-bn_analysis.html](https://bernard-liew.github.io/2020_cts_bn/2-bn_analysis.html). BN is a graphical modelling technique (11) that can leverage either data alone, or data combined with expert prior knowledge to learn multivariate pathway models. Building a BN model using a data-driven approach involves two stages: 1) structure learning - identifying which arcs are present in the graphical model; and 2) parameter learning - estimating the parameters that regulate the strength and the sign of the corresponding relationships.

As previously mentioned, BN can easily include prior knowledge, sourced from the literature and experts, during the model building process. In the BN framework, prior knowledge can be included in the model as blacklist and whitelist arcs. Blacklist arcs are those that contravene known biological/physical mechanisms. In the current study, we imposed the following blacklist:

- Arcs cannot point backward in time (e.g. 12\textsuperscript{th} month pain cannot influence 6\textsuperscript{th} month pain);
- No variables can influence group, since group allocation was random;
- No other variables can influence baseline patient characteristics.
We made use of model averaging to reduce the potential of including spurious relationships in the BN, using bootstrap resampling \((B = 200)\) and performing structure learning on each of the resulting samples using the hill-climbing (HC) algorithm. We computed an “average” consensus DAG by selecting those arcs that have a frequency of \(>50\%\) in the bootstrapped samples, to create a sparse and interpretable network (17).

To determine the validity of the trained model, validation was performed using nested 10-fold cross-validation (CV). This splits the training set into 10 approximately equal folds, trains the model on 9 folds using bootstrap resampling (as described above), and evaluates the model’s performance on the 10th fold. Model performance was defined as the correlation coefficient between the predicted and observed values of each continuous variable. The strength of correlation was categorized as negligible \((|r| \leq 0.30)\), low \((|r| = 0.31 \text{ to } 0.50)\), moderate \((|r| = 0.51 \text{ to } 0.70)\), high \((|r| = 0.71 \text{ to } 0.90)\) and very high \((|r| = 0.91 \text{ to } 1.0)\) (19).

The greater the model predictive performance, the greater the correlation between predicted and observed values of the modelled variables. A nested 10-fold CV reduces validation optimism, since a model would perform well on the data it was exactly trained on.

**Conditional Probability Queries**

The derived averaged BN model is considered an “expert system”, which means that we can elicit a sample of realisations of the modelled variables under specific conditions. For each conditional probability query, we sampled \(10^4\) realisations of the variables of interest to obtain precise probability estimates. We used a technique known as belief updating, which estimates the posterior probability of an event happening based on the available evidence on the values of certain variables. We adopted a specific method of belief updating known as logic sampling (11).
Results

From the baseline, 120 women were included in the original trial, and after trimming missing values, 104 women (52 in each group) were included in the current BN. Of the 52 participants allocated to manual physical therapy, 15 were classified as minimal, 20 as moderate, and 17 as severe. For participants allocated to surgery, nine were classified as minimal, 26 as moderate, and 17 as severe (Figure 1).

Figure 2 shows the averaged BN consensus model learned from 200 networks constructed from the data, with arcs appearing in >50% of the networks kept. The predictive correlations for all variables are included in Table 1. An advantage of BNs is that the model enables the reader to query different elements of the system to fully understand the interaction between the variables. By systematically removing individual variables from the model, the impact of that variable on the remaining variables in the model becomes clearer. For simplicity, the magnitude of the relationship between variables is reported using $\beta$ coefficients, which are interpreted as a one-unit change in the independent variable resulting in a $\beta$ unit change in the dependent variable. As the BN model was complex with multiple interconnected variables, we explore in more detail below four findings that were of most relevance to understanding the mechanisms of the interventions. It can be observed that four independent variables influenced 12th month function: 1) treatment group, 2) baseline pain intensity, 3) 1st month symptoms’ severity, and 4) EMG classification (Figure 2).

The Effect of Treatment Group on 12th Month Function

Based on simulated data from the model, 12th month function was 0.09 points lower in the manual therapy (suggesting better function) than in the surgery group (t=8.44, P<0.001, Figure 3). From Figure 2, group had a direct effect on 12th month function, but also an indirect effect via its influence on 1st month function. When we simulated a scenario where the group to 1st month function arc was removed by fixing the value of the 1st month function
regression coefficient in the local distributions to zero, 12th month function was 0.03 points lower in the manual therapy group (suggesting better function) than in the surgery group (t=3.38, P=0.001). This suggests that almost one-third of the influence of treatment group on 12th month function was attributed to its direct effect, and the remaining two-thirds of the influence was attributed to the treatment group effect obtained at 1st month for function.

**Influence of Baseline Pain on 12th Month Function**

Baseline pain intensity was significantly associated with 12th month function. A one-point increase in pain worsened 12th month function by 0.008 points (t = 2.39, P=0.017, Fig. 4). From Figure 2, baseline pain intensity influenced 12th month function via its impact on baseline depressive levels. When we simulated a scenario where baseline depression was made independent from baseline pain intensity, the latter no longer had a significant association with 12th month function (t=1.30, P=0.195).

**Influence of 1st Month Symptoms’ Severity on 12th Month Function**

The 1st month symptoms’ severity was significantly associated with 12th month function. A one-point increase in 1st month symptoms’ severity worsened 12th month function by 0.15 points (t = 14.57, P< 0.001, Figure 5). Although 1st month symptoms’ severity influenced 12th month function via multiple pathways, a common point of effect was its effect on 6th month function. When we simulated a scenario where 1st month symptoms’ severity was made independent from 6th month symptoms’ severity, there was no longer a significant association between 1st month symptoms’ severity and 12th month function (t = -0.79, P=0.429).

**Influence of EMG classification on 12th Month Function**

Based on simulated data from the model, there was a significant effect of EMG classification with 12th month function (F=70.86, P<0.001, Figure 6). Post-hoc Tukey Honest Significant Differences test found that people with moderate CTS had 0.14 points (P<0.001)
poorer function than those with mild CTS and 0.12 points (P<0.001) poorer function than those with severe CTS. There was no difference in 12th month function between severe and mild CTS (P=0.39).

Discussion

This study presents the first longitudinal investigation to adopt a data-driven modelling approach to quantify the probabilistic pathways of effect underpinning manual physical therapy or surgery for individuals with CTS. We found that allocation to a manual physical therapy group directly improved long-term function, in comparison to surgery. We also observed that baseline pain intensity influenced the effects of treatment and was mediated by depressive symptoms. Lastly, we found that early improvements (~1 month) in function and symptom severity influenced the effect of physiotherapy and surgery on pain and related-disability in women with CTS. These findings are worth discussing in further detail.

The Effect of Treatment Group on 12th Month Function

Allocation to the manual physical therapy treatment group had a direct effect on 12th month function, but also an indirect effect via its influence on 1st month function. The indirect effect may be associated with the recovery time associated with CTS surgery. On average, it takes 4 to 6 weeks before the surgery heals to a point where individuals may report improvement of symptoms. During this time, individuals are asked to avoid heavy lifting, repeated movements of their wrist, and are placed in a protected protocol to allow the surgical approach to healing. In fact, although there are different types of surgeries, some more invasive than others, the time required to heal will likely always result in a delayed recovery in comparison to a conservative approach. The slower improvement in function observed in
the surgery group mostly at 1st and 3rd months seems to be related to the tissue healing recovery process needed after the surgical intervention.

The Influence of Baseline Pain on 12th Month Function

We feel there is an intuitive relationship between pain-related symptoms and disability and that the higher level of baseline symptoms is likely related to the severity of the pathology. This relationship has been explored and validated in numerous studies including individuals with musculoskeletal pain conditions. Higher baseline pain intensity is a predictor for low back pain (20) and neck pain (21) at 12 months and also at 12 months after distal radius fracture (22). In people with neuropathic pain, baseline leg pain intensity has been also associated with surgery in individuals with sciatica (23). The review by Mallen et al. indicated that higher baseline pain intensity is a negative prognostic factor for a number of musculoskeletal pain conditions, including the spine, shoulder, neck, hip, knee, and elbow (24). To our knowledge, we are the first to report on the predictive capacity of baseline pain intensity for individuals with CTS.

Influence of 1st Month Symptoms’ Severity on 12th Month Function

Previous studies have evaluated whether early change in symptoms are related to long-term outcomes. Rundell et al investigated if 3-month outcomes including back and leg pain could predict 12-month back and pain, disability, and patient satisfaction; they reported that the 3-month data were stronger predictors than the baseline counterparts (25). Walston and McLester reported that early changes in reports of low back pain within the first 23 to 30 days predicted long-term changes and were reflective of those who benefited from physical therapy (26). Others studies examining changes in pain during much shorter terms found benefits in a sustained between-session treatment. Early responses from the second to fourth treatment visits were able to correctly predict 80.4% of the discharge outcomes in patients with chronic low back pain (27). A better short-term response that occurs from the first to
second treatment is associated with better effects of physical therapy interventions at 6-month follow-up periods in people with low back pain (28). It is worth noting that all the aforementioned studies lacked a comparative group and used simple correlational statistical measures. To our knowledge, no previous studies have explored early change of pain intensity in individuals with CTS.

**Influence of EMG classification on 12th Month Function**

We also observed that women with mild or severe EMG findings experienced between long-term function than those with moderate findings. This was an unexpected result since it is suggested that more severe EMG findings would lead to worse clinical outcomes, at least within the manual physical therapy group. In fact, clinical guidelines recommend conservative treatment mostly for mild to moderate, and, sometimes severe, cases of CTS (1, 2). This premise is based on trials examining therapeutic approaches using the traditional clinical reasoning that CTS is just a localized pathology associated with a peripheral lesion at the carpal tunnel and, therefore, local interventions are applied. Since the current clinical trial used manual physical therapy interventions considering current nociceptive understanding of CTS, it is possible that this approach would lead to better outcomes albeit in severe CTS. Future studies should investigate this hypothesis.

**Limitations**

This study had several potential strengths and limitations. First, the analysis was performed on a robustly conducted, randomized clinical trial, with 87% of the original cohort included in the current analysis. Secondly, we uncovered surprising pathways of treatment effect, such as long-term effects on related-function being dependent on short-term changes, which have not been previously tested in this pain condition. Despite these strengths, some weaknesses are also present. We note that methods to quantify mechanisms of change is not a singular method, but actually consists of three progressive levels (29):
association (i.e., “how are the variables related?”), intervention (i.e., “what would Y be if I do X?”), and lastly counterfactuals (i.e., “what if I had acted differently?”). Each ladder provides a potential incremental amount of evidence towards causal inference. The current BN analysis can only act on the first and second rung (via our simulated intervention analysis) of the causation ladder, meaning that we cannot definitively conclude our reported pathways as really causal. Despite an inability to perform counterfactual analysis, the current analysis could be said to provide competing, and potentially more probable, pathways of an effect than those presented in the literature, that can be confirmed by future research. Another limitation of this study was that psychophysical and psychological variables were not included in the BN since the original clinical trial did not consider them. Newer insights into the mechanisms of effect of treatments offered to manage CTS could be revealed by the inclusion of biopsychosocial variables in future studies.

Conclusions

A data-driven BN modelling approach showed that manual physical therapy reduced related-disability in women with CTS. Early improvement in function and symptoms’ severity led to long-term improvement in function both directly and via more complex pathways involving baseline pain intensity and depressive levels. In addition, the severity of the condition, expressed by EMG data, had also a direct influence on long-term effects on function. Current findings as a whole suggest that short-term benefits in function and symptoms’ severity were associated to long-term improvement in related-disability; however, the mechanisms driving the effects interacted with depressive levels and EMG severity.
Legend of Figures

**Figure 1:** Mean with error bars as one standard deviation of the values of each modelled variable in the Bayesian Network. Abbreviations: _base: values at baseline; _1m: values at 1-month follow-up; _3m: values at 3-months follow-up; _6m: values at 6-months follow-up; _12m: values at 12-months follow-up; dep: total score of the Beck Depression Inventory (BDI-II) for depressive symptoms; severe: symptoms’ severity subscale of the Boston Carpal Tunnel Questionnaire; func: function subscale of the Boston Carpal Tunnel Questionnaire

**Figure 2:** The directed acyclic graph (DAG) underlying the consensus Bayesian Network of learned from the variables. The colour of the arcs reflects the sign (positive [blue] or negative [red]) of the $\beta$ coefficient value relating the “parent” to “child” variables. The thickness of the arcs reflects the proportion of times each arc was found in the 200 Bayesian Network models built; only arcs with a proportion greater than 0.5 are included in the final averaged consensus network. Abbreviations: _base: values at baseline; _1m: values at 1-month follow-up; _3m: values at 3-months follow-up; _6m: values at 6-months follow-up; _12m: values at 12-months follow-up; dep: total score of the Beck Depression Inventory (BDI-II) for depressive symptoms; severe: symptoms’ severity subscale of the Boston Carpal Tunnel Questionnaire; func: function subscale of the Boston Carpal Tunnel Questionnaire

**Figure 3:** Mean with error bars as one standard deviation of the posterior samples ($10^4$) of the relationship between group (manual therapy -MT- versus surgery) and 12th month function.

**Figure 4:** Posterior samples ($10^4$) of the relationship between baseline pain intensity (pain_base) and 12th month function (func_12m).

**Figure 5:** Posterior samples ($10^4$) of the relationship between 1st month severity (severe_1m) and 12th month function (func_12m).

**Figure 6:** Mean with error bars as one standard deviation of the posterior samples ($10^4$) of the relationship between EMG classification and 12th month function.
References


7. Fernández-de-las-Peñas C, Fernández-Muñoz JJ, Palacios-Ceña M, Navarro-Pardo E, Ambite-Quesada S, Salom-Moreno J. Direct and Indirect Effects of Function in


carpal tunnel release for idiopathic carpal tunnel syndrome: a meta-analysis of

16. Leite JC, Jerosch-Herold C, Song F. A systematic review of the psychometric
properties of the Boston Carpal Tunnel Questionnaire. *BMC Musculoskelet Disord*
2006;7: 78.

17. Scutari M, Nagarajan R. Identifying significant edges in graphical models of

18. Scutari M. Learning Bayesian Networks with the bnlearn R Package. *J Stat Softw*


20. Garcia AN, Costa LOP, Costa L, Hancock M, Cook C. Do prognostic variables
predict a set of outcomes for patients with chronic low back pain: a long-term follow-
up secondary analysis of a randomized control trial. *J Man Manip Ther* 2019;27(4):
197-207.

Trajectories of Pain Intensity Over 1 Year in Adults With Disabling Subacute or

22. Mehta SP, MacDermid JC, Richardson J, MacIntyre NJ, Grewal R. Baseline pain
intensity is a predictor of chronic pain in individuals with distal radius fracture. *J

Systematic review of prognostic factors predicting outcome in non-surgically treated


