The Validity of Physical Therapy Assessment of Low Back Pain via Telerehabilitation in a Clinical Setting

Piers Truter, BPhty (Hons),¹,² Trevor Russell, PhD, BPhty,³ and Robyn Fary, PhD, Grad Dip Manip Therapy, BAppSc (Phty)¹,⁴

¹School of Physiotherapy and ⁴Curtin Health Innovation Research Institute, Curtin University of Technology, Bentley, Western Australia, Australia.
²Bowen Hospital, Queensland Health, Bowen, Queensland, Australia.
³Department of Physiotherapy, University of Queensland, St. Lucia, Queensland, Australia.

Introduction

Nearly 3 million Australians report living with chronic back pain and disc disorders, with peak prevalence in the prime working ages of 35–64 years.¹ There is a disparity in burden of musculoskeletal disease² and disease chronicity in Australia that is proportionally linked to rurality. Lack of access to healthcare facilities and professionals is a critical factor in the variation in health outcomes between rural/remote and urban populations³,⁴ and a contributing factor to higher prevalence of chronic musculoskeletal conditions in rural areas.⁵,⁶ Access to physical therapy services in rural areas follows the pattern of other medical and health-related services. In rural areas the number of physical therapists (PTs) per capita is less than in urban areas,⁷ the workforce is aging,⁸ and the trend is for declining numbers⁸ and clinical experience of those PTs.⁹ Rural populations have reduced access to physical therapy services and often have to travel to larger urban centers for treatment.¹⁰

One proposed solution to the physical therapy service delivery problem in rural and remote Australia is the development and evaluation of telerehabilitation (TR) models of service delivery.¹¹–¹³ Remote assessment of musculoskeletal conditions via videoconference is considered acceptable by patients in the absence of local service.¹⁴–¹⁶ The issue remains that this mode of service delivery is perceived as lacking depth of evidence in regard to implementation, clinical efficacy, clinical ethicality,¹⁷ and cost-effectiveness.¹⁸

There is currently only a small body of research in TR for physical therapy management.¹⁶,¹⁹–²⁴ Functional measures and validated pathology-specific questionnaires are the most commonly used outcome measures. Although these components can inform management of musculoskeletal conditions, diagnosis requires physical measures and special tests. Previous research has validated some components of a musculoskeletal assessment, including observation,²⁵ postural analysis, muscular strength testing,²⁶ joint range of movement,²⁷–²⁹ and functional task evaluation.²⁵,²⁶,³⁰ Building the link between clinimetrics and practical clinical application, remote assessment, and diagnosis of lower limb musculoskeletal injuries by a PT has been validated.¹⁴,¹⁵

The purpose of this study is to further establish the validity of performing an effective physical examination of the lumbar spine via TR in a clinical setting. Specifically, this study aims to establish the concurrent validity of remote assessment of spinal posture, active movements of the lumbar spine, and the passive straight leg raise (SLR) test. A secondary aim is to evaluate satisfaction with the TR assessment in a rural patient population.

Abstract

Background: Back pain is a common and disabling condition for people in rural and remote areas. In these areas, access to rehabilitation services is limited by service availability. Telerehabilitation is suggested as a solution for providing physical therapy services; however, the validity of clinical assessment is largely unproven. The aim of this study was to establish the validity of clinically pragmatic remote assessment of spinal posture, active movements of the lumbar spine, and the passive straight leg raise (SLR) test. Subjects and Methods: Face-to-face physical therapist assessment was compared with telerehabilitation assessment of spinal posture, active movements of the lumbar spine, and the SLR test. Twenty-six participants recruited from a rural population with current or recent low back pain (LBP) were assessed by a face-to-face physical therapist and a remote physical therapist. Pain, disability, and clinical measurements were assessed. Outcomes were compared to establish agreement. Results: High levels of agreement were found with detecting pain with specific lumbar movements, eliciting symptoms, and sensitizing the SLR test. Moderate agreement occurred with identifying the worst lumbar spine movement direction, SLR range of motion, and active lumbar spine range of motion. Poor agreement occurred with postural analysis and identifying reasons for limitations to lumbar movements. Conclusions: Conducted in a rural clinical setting, this study validates elements of the physical assessment of the lumbar spine and identifies technical and clinical issues to be addressed by future research. Important components of the standard musculoskeletal assessment of LBP are valid via telerehabilitation in a clinical setting.

Key words: e-health, telehealth, teledmedicine, rehabilitation

DOI: 10.1089/tmj.2013.0088
© MARY ANN LIEBERT, INC. • VOL. 20 • NO. 2 • FEBRUARY 2014

TELEMEDICINE and e-HEALTH
Subjects and Methods

STUDY DESIGN

A single-blinded validation study comparing face-to-face (F2F) assessment with TR assessment of patients with low back pain (LBP) by a PT was conducted. The study was approved by the relevant human research ethics committees (EC00183 and EC00262).

PARTICIPANTS

Twenty-six participants with LBP from a small town in regional Queensland, Australia (population 10,260)11 were recruited by local public advertising. Participants were included if they had current or recent (in the last 2 years) LBP. Participants were excluded if they were minors, had medical conditions precluding a safe physical examination, lacked adequate cognition or communication to allow the use of the TR system, could not mobilize independently, or currently had severe irritable LBP and/or severe neurological symptoms. Participants gave written consent before entering the study. Participants were asked to bring a friend to the appointment to assist with the assessment. Where they were unable to do this, an untrained assistant was drafted from the nonclinical staff at the hospital.

RANDOMIZATION AND BLINDING

The order of physical examinations (F2F and TR) and the PT assigned to each examination (two assessing PTs) were randomly assigned using a computer-generated code. There were four possible combinations of first assessment and therapist. Participants were sequentially allocated as they were recruited using a balanced block design of 4. Each assessor was blinded to the other assessor’s findings throughout data collection.

MATERIALS

General. The F2F and TR assessments were conducted in the same physical therapy treatment room. The treatment room housed a videoconferencing unit that was remotely controlled by the TR assessor using a second unit from a separate room in the hospital.

Videoconferencing equipment. TR assessments were performed using the eHAB® TR system (eHAB v2; NeoRehab, Brisbane, QLD, Australia). eHAB is a computer-based videoconferencing system that enables remote consultations between patients and health professionals via a wireless 3G Internet connection (Telstra Next G™; Telstra, Melbourne, VIC, Australia). The system includes a battery of measurement tools that objectively quantify the participant’s physical performance (such as balance, joint range of motion, muscle strength, and assessment of gait) across the Internet link. This system has been described elsewhere.14,15,25,28–30 The remote assessment used various features of the eHAB system, including real-time videoconferencing, store-and-forward still images (640 × 480 pixel resolution) of the participant, and remote camera pan and zoom control.

PROCEDURE

On arrival, participants were asked to rate LBP intensity and complete the Oswestry Disability Index (ODI) questionnaire.32 Depending on randomization, the participant was then either assessed by the F2F PT or positioned in front of the TR unit and assessed by the TR PT. Once the first assessment was complete, participants re-assessed their LBP intensity. Any participant who demonstrated a significant increase in LBP was offered the option to end his or her involvement in the study. After a short interval (5–10 min) the second assessment was performed, being the assessment type that had not already occurred.

At the conclusion of both assessments each participant was again asked to re-assess his or her LBP and to rate satisfaction with the assessment procedures.

The F2F assessment was conducted according to normal physical therapy clinical practice. The TR assessment was conducted with the participant standing on a reference line on the floor of the clinic and next to a 300-mm black and white calibration index. The participant moved relative to the TR unit under the PT’s instruction to provide the appropriate views to the camera on the unit. Movements were recorded on video (640 × 480 pixels), and posture was recorded by still images (640 × 480 pixels) using features of the eHAB system. Clinical measurements such as SLR leg angle were extracted from the recorded video once the participant had left, by the TR PT using the in-built software tools in the eHAB units. The TR assessment was pragmatically designed to require the minimum amount of equipment and set-up at the remote end.

OUTCOME MEASURES

Disability. Disability was measured using the ODI, a valid32 and reliable13 outcome measure in patients with LBP. In this study it was used as a demographic tool to characterize the participant population in this rural setting.

Pain. Pain was measured using a 100-mm visual analog scale (VAS), bounded by the descriptors “no pain” scoring 0 and “worst possible pain” scoring 100. This measure was included for participant safety and to screen for possible bias caused by a change in pain between the first and second assessments.14

Posture. Postural analysis is a standard element of a typical physical therapy clinical examination of the lumbar spine34 and is routinely analyzed from a photographic image.35,36 Standing posture was assessed in the coronal and sagittal planes. In the coronal plane assessment the PT identified the presence of postural symmetry and scoliosis and categorically rated pelvic tilt and spinal rotation. In the sagittal plane assessment the PT categorized rated lumbar lordosis, thoracic kyphosis, pelvic tilt, and the relative positions of the shoulders, pelvis, and lower limbs.

Active movement. Active movements of the lumbar spine form a standard part of a typical physical therapy clinical examination of the lumbar spine. While standing, participants were asked to move to the end of their range of movement in flexion, extension, lateral flexion (left and right), and rotation (left and right). Flexion, extension, and lateral flexion range of motion was measured between the
floor and the tip of the third finger(s) with a tape measure (F2F) or using the built-in measuring tools in the eHAB system. The finger-to-floor method has demonstrated good correlation with lumbar spine movements and is considered a valid and reliable measure. Range of motion for rotation was rated as either full or restricted owing to difficulties obtaining an objective measure. With each movement participants were asked to describe changes in pain levels (100-mm VAS) and to provide a description of why range of motion was restricted. Participants’ responses were categorically rated by the PT. If a movement caused an increase in pain, the PT coded this in a separate binary variable.

**SLR test.** SLR is commonly used to evaluate LBP. The small number of reliability studies suggest that it has good intra-rater reliability (0.95 for intra-session and 0.88 for inter-session [Cronbach’s alpha] and intraclass coefficient of 0.95–0.98) and good criterion validity with a hand-held inclinometer. SLR is routinely used by PTs to assess mechanosensitivity of the sciatic nerve as part of the clinical picture of LBP. Cadaveric studies have informed clinical practice, and diagnostically it is suggested that SLR should be sensitized with ankle dorsiflexion, hip internal rotation, and passive neck flexion. SLR was measured with the participant lying in supine on a treatment plinth. The F2F PT passively flexed the hip of the test leg (with full knee extension) until the participant reported a symptomatic response. This symptomatic response-limiting movement was categorically rated. The range of motion was measured with a hand-held fluid-filled inclinometer. The PT then conducted three nerve tensioning-sensitizing movements. The nerve tensioning-qualifying movements were rated positive or negative depending on the symptomatic responses. In the TR assessment the passive leg movement was facilitated by the participant’s attending friend or an untrained nonclinical assistant. The measurement of hip flexion was taken from a still image using the built-in software goniometer in the eHAB system.

**Participant satisfaction.** At the conclusion of the assessment, participants rated their satisfaction with TR on a series of six 100-mm VAS instruments: (1) how confident they were with the Internet practice; (2) whether they would recommend this to a friend who is unable to travel; (3) whether they thought this method of assessment was as good as traditional F2F assessment; (4) whether they could see the PT clearly at all times; (5) whether they could hear the PT at all times; and (6) what their overall satisfaction was. The rating scale was bounded by the descriptors “complete dissatisfaction” scoring 0 at the left side anchor and “extremely satisfied” scoring 100 at the ride side anchor.

**STATISTICAL ANALYSIS**

All data were analyzed using the SPSS software package version 18.0 (SPSS, Inc., Chicago, IL). The agreement between F2F and TR assessments was analyzed by chi-squared tests (dichotomous variables), percentage of exact agreement, and kappa statistics (categorical data) and Pearson’s correlation (r) (continuous data).

The strength of the Pearson’s correlation can be assessed as follows: 0.00–0.35, weak correlation; 0.36–0.67, moderate correlation; and 0.68–1.00, strong correlation. Kappa statistics (κ) are commonly used in validity articles because they effectively discount the proportion of agreement that is expected by chance. Strength of agreement between 0 and 1 can be judged as follows: 0.00–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.8, substantial; and 0.81–1.00, almost perfect. Values of κ over 0.40 were considered to be clinically acceptable.

A paired t test was used to compare resting pain (VAS) levels. All analyses set the level of significance at p < 0.05.

**Results**

**PARTICIPANTS**

Thirty-two people were recruited to the study. Four participants were excluded, and two did not attend their appointment. Twenty-six participants (11 men, 15 women; mean age, 43 years; mean ODI score, 18 ± 16%) were assessed. Most participants reported minimal (0–20%) and moderate (20–40%) disability on the ODI, although four participants reported severe disability (40–60%).

**POSTURAL ANALYSIS**

There was slight agreement between F2F and TR postural assessment (Table 1). Categorical data produced kappa statistics below the clinically acceptable threshold, and binary data were not found to significantly agree.

<table>
<thead>
<tr>
<th>Table 1. Agreement Analysis Between Face-to-Face and Telerehabilitation Assessment of Posture</th>
<th>% EXACT AGREEMENT</th>
<th>CHI-SQUARED (P VALUE)</th>
<th>KAPPA STATISTIC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coronal postural assessment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symmetry</td>
<td>56%</td>
<td>0.43 (0.51)</td>
<td></td>
</tr>
<tr>
<td>Presence of scoliosis</td>
<td>72%</td>
<td>2.21 (0.14)</td>
<td>0.17</td>
</tr>
<tr>
<td>Pelvic tilt (LR)</td>
<td>52%</td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td>Spinal asymmetry classification</td>
<td>36%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sagittal postural assessment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelvic tilt (AP)</td>
<td>75%</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Pelvic position: forward/back</td>
<td>71%</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Lumbar lordosis</td>
<td>25%</td>
<td>–0.20</td>
<td></td>
</tr>
<tr>
<td>Thoracic kyphosis</td>
<td>50%</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Thoracic position</td>
<td>67%</td>
<td>0.19</td>
<td></td>
</tr>
</tbody>
</table>

*Unable to calculate because of non-symmetrical table.
AP, anterior/posterior; LR, left/right.
ACTIVE RANGE OF MOTION

The means of measures for all active movements are shown in Table 2. There was strong correlation in the measurement of flexion \((r=0.89)\) and extension \((r=0.83)\).

Agreement that a movement was painful. There was excellent agreement between the F2F and TR assessments in identifying if a particular movement was painful. All movements exceeded 80% exact agreement and were found to be in significant agreement (chi-squared test) (Table 3).

Reason for limit to movement. The pooled categorical data (all directions of movement) for the reason that movements were limited had fair agreement \((\kappa=0.37;\) percentage exact agreement = 55%).

Worst direction of movement. The assessment of the worst direction of movement had moderate agreement \((\kappa=0.56;\) percentage exact agreement = 65%).

SLR TEST

Range of motion to symptom. The mean range of motion for the pooled data of both legs for the F2F assessment was 52° (standard deviation \([SD]=19°\)). The mean for the TR was 59° \((SD=15°)\). Mean difference between F2F and TR was –6° \((SD=15°)\), and standard error of the mean for the difference was 2.23°. There was moderate correlation between F2F and eHAB measurements \((r=0.64; p<0.001)\).

Symptoms and sensitization. There was substantial agreement in the categorical rating of symptoms with SLR. There was significant agreement in detecting pain as a symptom, sensitizing the test with dorsiflexion or hip internal rotation (Table 4).

PARTICIPANT SATISFACTION

Overall, the participants indicated a good level of satisfaction with all questions, except Question 3, where they clearly preferred F2F assessment (Fig. 1).

---

<table>
<thead>
<tr>
<th>Table 2. Range of Lumbar Spine Motion Data for Face-to-Face and Telerehabilitation Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIRECTION OF MOVEMENT</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Flexion</td>
</tr>
<tr>
<td>Extension</td>
</tr>
<tr>
<td>Lateral flexion</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

All measurements are in centimeters.
EHAB, eHAB telerehabilitation; F2F, face to face; SD, standard deviation.

---

<table>
<thead>
<tr>
<th>Table 3. Levels of Agreement in Identifying If a Lumbar Spine Movement Direction Was Painful</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIRECTION OF MOVEMENT</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Flexion</td>
</tr>
<tr>
<td>Extension</td>
</tr>
<tr>
<td>Lateral flexion</td>
</tr>
<tr>
<td>Right</td>
</tr>
<tr>
<td>Left</td>
</tr>
<tr>
<td>Rotation</td>
</tr>
<tr>
<td>Right</td>
</tr>
<tr>
<td>Left</td>
</tr>
</tbody>
</table>

---

REPEATED-MEASURES EFFECT ON PARTICIPANT PAIN LEVELS

There was no significant difference in mean pain ratings among the initial rating, the rating after the first assessment, and the rating after the second assessment.

Discussion

Using F2F physical therapy assessment as a gold standard, we have compared the TR assessment of spinal posture, active movements of the lumbar spine, and SLR. We found high levels of agreement with establishing if a lumbar spine movement was painful, detecting pain, eliciting symptoms, and sensitizing the SLR. Moderate agreement was found in identifying the limitation to an active lumbar spine movement, identifying the worst lumbar spine movement direction, SLR range of motion, and active lumbar spine range of motion. Poor agreement was found in all elements of the postural analysis.

Postural analysis is a routine part of a physical therapy assessment of the lumbar spine, and certain postures have been correlated with LBP.36,49 This section of the TR assessment was particularly difficult for the PTs because of several factors. There appeared to be a keystoning effect from the position of the wide angle camera lens that made it difficult to analyze coronal posture. Furthermore, the resolution of the images \((640 \times 480)\) was insufficient to discriminate physical landmarks and hence allow postural assessment. Additionally, four female participants were unwilling to disrobe for the postural analysis, and a compromise was reached by tucking their shirts into their bra straps. This is not an unusual occurrence in this physiotherapy clinic, and participants requested this for both the F2F and eHAB assessments. This compounded
In this study the TR SLR examination was facilitated by a friend accompanying the participant or by an untrained nonclinical assistant. We found moderate correlation for the SLR range of motion measure, where this has been found to be strongly reliable (left leg, $r = 0.81$; right leg, $r = 0.79$) when conducted in a F2F examination by a trained and qualified assessor. The sensitizing movements were included to aid diagnosis, although there is debate about the necessity of this. We found excellent agreement with these dichotomous data, which is consistent with other reliability study literature. Overall our results suggest that this pragmatic method of assessment may be clinically appropriate.

The participant satisfaction was similar to ratings taken during earlier urban studies, with a minor (but not significant) increase in rural participant likelihood to recommend and overall satisfaction. There was moderate satisfaction with the quality of the audio and vision in the assessment, despite two afternoon sessions where there was a significant degradation to the quality of the audio that was noted by both the PTs and the participants.

A strength of this study is that it was conducted in a rural hospital by rural clinicians, with a group of local participants who have experience with limited access to health services. In setting up the study the clinicians had to address many of the issues that would occur in implementing a TR service in a rural hospital (e.g., finding an appropriate and available clinical space). The study was designed with outcome measures that emphasized easy set-up and recording (clinical pragmatism), as these would be easier to implement than protocols that required elaborate or technical set-up (e.g., postural analysis without photoreflective markers). Other than the TR equipment, all the equipment should be easily available in a rural hospital or even a community dwelling. The use of untrained nonclinical assistants was also aimed at ease of implementation in these environments. This is important for facilitating uptake of the technology into normal clinical practice. From a participant perspective the assessments were simple and nonintimidating, which is important for a technology that could deliver care into the family home.

The repeated-measures study design may have affected the comparability of the two assessments. A potential participant learning effect may have occurred as the second assessment was often quicker because the participant had experienced the assessments and instructions previously. Other studies investigating LBP have used a dual rater/single assessment design, but this is not without bias as the dual rating may artificially boost the agreement between raters. We also noted that the effect of repeated movement is not consistent in a LBP population, with some participants getting relief from repeated movement and others finding repetition provocative. In this study there was also a degree of heterogeneity in the population, with most patients (50%) stating that they did not have LBP on the day and a few who had moderate pain levels at rest and found some movement pain provocative. These results are presented from a sample where 76% demonstrated minimal disability scores on the ODI. It would be interesting to see if the validity of range of motion and SLR assessments remains in a sample where patients have higher ODI scores. It may be more important for these people to be able to access...
TR assessment as their high disability levels may preclude them from traveling for services.

Only a small percentage of the people attended with a friend to assist in the SLR assessment. The assistant was typically a nonclinical member of the hospital staff recruited on the day of the assessment. During sessions where multiple participants were evaluated, there may have been a learning effect, improving the assistant’s skills with each subsequent assessment.

We have shown that several important elements of the standard musculoskeletal assessment of LBP are valid in a sample of people with minimal disability in a rural context using clinically pragmatic assessment methodology via TR. This study has highlighted that some areas of the assessment require further testing and development (e.g., postural analysis), but this is consistent with other studies of assessment and diagnosis that have observed that the physical therapy assessment has to be adapted to the TR environment.14 Further research is needed into developing and testing a clinically robust TR assessment for LBP, with a view to facilitating diagnosis and ultimately treatment. There may be particular value in investigating multidisciplinary treatment for chronic and persistent LBP, which is highly prevalent in rural areas and requires specialist care.

Acknowledgments

The authors thank the staff of Bowen Hospital, especially Pauline Maude, Michael Young, Annie Barton, and Cheryl Young. The study could not have happened without the assistance of Chelsey Carr, the second assessing PT.

Disclosure Statement

T.R. has a material interest in the eHAB telerehabilitation system. To prevent bias this author was not involved in data collection or analysis aspects of the study. P.T. and R.F. declare no competing interests exist.

REFERENCES


Address correspondence to:
Piers Truter, BPhty (Hons)
School of Physiotherapy
Curtin University of Technology
G.P.O. Box U1987
Perth, WA 6845
Australia

E-mail: pierstruter@gmail.com

Received: March 24, 2013
Revised: May 29, 2013
Accepted: May 30, 2013