Development of an Imaging Evaluation System for Low Back Pain

Ping-En Sun^{1,4} · Bing-Shiang Yang^{1,2,3,4}

Received: 10 September 2020 / Accepted: 12 January 2021 / Published online: 20 February 2021
© The Author(s) 2021

Abstract

Purpose To design a low back pain (LBP) imaging evaluation system for non-professional operators, quantifying risk of LBP in patients and asymptomatic individuals.

Methods Twenty-one previously asymptomatic subjects and five LBP patients diagnosed by a physician performed a series of test movements under a fixed camera, including hip abduction test, forward bending test, side bending test, double legs lowering test (DLLT), and modified Thomas test. The video clips were analyzed by a system program interface and were classified into a score from 1 to 4. The average total scores of the two groups were compared. Five intact subjects were retested to verify reliability. Twelve intact subjects and three patients' clips were viewed by an experienced therapist to verify system validity.

Results The average total scores of two groups were significantly different ($p=0.0004$). The results of hip abduction test, forward bending test, side bending test and DLLT showed significant differences between the two groups. The total score of two trials in the retest experiment exhibited a similar result ($p=0.058$), with good linear correlation ($r=0.98$). The total scores of the evaluations by the experienced therapist and the system program interface agreed with each other ($p=0.141$; linear correlation $r=0.90$).

Conclusion The evaluation system showed potential utility in screening LBP risk, with acceptable test–retest reliability and expert validity.

1 Introduction

Low back pain (LBP) is a common illness in modern society [1]. LBP can be divided into specific and non-specific LBP. The reason for specific LBP includes tumor, infection or radicular problem, while the non-specific LBP usually does not have a clear cause. Currently the diagnosis of non-specific LBP is conducted by a professional in hospital outpatient clinics through medical history inquiry and physical examinations [2].

Multiple studies have proved that people with LBP have altered movement patterns when conducting physical examinations, such as hip abduction test, forward bending test, side bending test, double legs lowering test (DLLT) and modified Thomas test (MTT). The performance of these five physical examinations test certain muscle groups around low back region and can indicate the potential risk of LBP. The performance of hip abduction test relates to the condition of gluteus medius muscle. Studies have suggested that if the leg deviates from the frontal plane, then the subject will likely develop LBP in the future [3, 4]. The performance of forward bending test (also known as instability catch sign) relates to the stability of lumbar muscle on the back side. LBP patients tend to exhibit deviation from the sagittal plane [5]. In the side bending test, a smaller side bending angle results from the unbalance of lumbar muscle on both sides. Studies show that LBP patients tend to have smaller bending angle [6]. DLLT is used to measure the strength of core muscles. There is evidence indicating that LBP patients tend to have a large hip flexion angle [7, 8]. MTT is used for measuring the flexibility of hip flexor muscles. Some research suggest that LBP patients tend to have smaller hip extension

✉ Bing-Shiang Yang
bsyang@nctu.edu.tw

¹ Department of Mechanical Engineering, National Chiao Tung University, Hsinchu City, Taiwan

² Graduate Degree Program of Robotics, National Chiao Tung University, Hsinchu City, Taiwan

³ Institute of Biomedical Engineering, National Chiao Tung University, Hsinchu City, Taiwan

⁴ Department of Mechanical Engineering, National Yang Ming Chiao Tung University, Hsinchu City, Taiwan

angle and knee flexion angle in the modified Thomas test [9–11].

Early intervention prevents worsening of severe non-specific LBP, but some patients may ignore the risk of LBP and delay medical intervention. In the health belief model, a higher perceived threat leads to a higher likelihood of engagement in health-promoting behaviors such as preventive care [12]. However, to increase the perceived threat for LBP patients, physical examinations must be operated by physicians in hospitals. Consequently, there is a lack of accessibility to increase the perceived threat for patients at risk of LBP. Little research has been done on developing an LBP evaluation system which can be operated in a home environment. The purpose of this study was to develop a LBP evaluation system which could be operated by non-professional operators to raise the awareness of a perceived threat of LBP in patients.

2 Methods

2.1 Participants

Approval of research was granted by the Research Ethics Committee for Human Subject Protection, National Chiao Tung University. Each subject signed an informed-consent form prior to participation in this study. Subjects ranged from 20 to 45 years old and were divided into a patient group and an intact group. Five patients were recruited from the orthopedics clinic of Mackay Memorial Hospital after being diagnosed with non-specific LBP by a physician. Twenty-one subjects in the intact group were students or employees recruited from National Chiao Tung University. Subject information is shown in Table 1.

2.2 Equipment and Setting

The evaluation system consisted of equipment, an imaging analysis method, a grading method, and an analysis interface. Equipment included an HD camera (HDR-CX405, Sony, Tokyo, Japan), a pressure biofeedback unit (PBU, Chattanooga Group Inc., Austin, United States), two

standing tripods, and several color markers. In each examination, researchers set up the camera on a tripod at a specific distance in front of the subject. The PBU was placed on the other tripod at a specific location in front of the subject in the DLLT and MTT tests. The system required several marks being attached to subject's limbs to track the movement of body segments. In the hip abduction test, one mark was placed on the center of foot sole. In the forward bending and side bending tests, two marks were placed on the first lumbar vertebra (L1) and the first sacral vertebra (S1). The system developed the following method to locate L1 and S1: researchers located the lowest ribs with palpation, aligned horizontally to the spine, and then moved downward vertically one vertebra to locate L1. To find S1, researchers located the iliac crest with palpation, aligned horizontally to the spine, and then moved downward vertically one vertebra to locate S1. In the DLLT, two marks were placed on one side of the greater trochanter and the outside of the knee joint. In the MTT, four marks were placed on one side of the greater trochanter, the outside of the knee joint, the upper end of fibula, and the outside of ankle joint. The locations of marks in each test are depicted in Fig. 1.

To make sure the evaluation system could be operated by non-professional operators, an instructional video was recorded that would be played before the procedure. The detailed guideline of the equipment setting and the marks locating method would be demonstrated in the video so that the subjects and operators could follow the instruction and performed tests.

2.3 Physical Examinations and Imaging Analysis

At the beginning of each test, the researcher gave an initial signal to the subject while the camera starting filming at the same time. Each test was repeated five times to obtain an average performance. Video clips would be analyzed by a MATLAB program to acquire the physical examination results.

In the hip abduction test, the subject was side-lying on the floor with a mark attached to the upper foot. The camera shot from the direction of the bottom of subject's feet. The subject then steadily raised the upper leg toward the ceiling along the frontal plane until reaching the highest point. After reaching the highest point, the subject steadily lowered the upper leg to the original state. When analyzing the image, the MATLAB program captured centroid of the color mark in every frame, and then calculated the horizontal offset distance between the coordinate of the current centroid and the centroid in the first frame, as shown in Fig. 1a. An average of the absolute value of offset distance in every frame was then calculated.

In the forward bending test, the subject stood naturally with their back facing the camera. Two color markers were

Table 1 Subject information table

Group	Intact group	Patients group
Number	21	5
Gender	14M, 7F	1M, 4F
Age (y/o)	25.7 ± 5.6	36.6 ± 7.2
Weight (Kg)	66.5 ± 17.5	63.0 ± 7.2
Height (cm)	168.2 ± 10.0	167.8 ± 6.5

M = male, F = female, y/o = years old

Fig. 1 **a** hip abduction test, d_1 was the offset distance, **b** forward bending test, α_2 was the offset angle, **c** side bending test, α_3 was the maximum bending angle on left side, **d** DLLT, α_4 was the hip flexion

angle, **e** MTT, α_5 was the hip extension angle and α_6 was the knee flexion angle, **f** design of the analysis interface program

attached to subject's L1 and S1 point on the midline of spine. The subject then bent forward steadily with the knee straightened, and then returned to the original state after reaching the lowest point. Subject was told to keep the spine moving along the sagittal plane throughout the process. The custom-made MATLAB computer program captured the line segment connecting the centroid of two color marks, and then calculated the offset angle between the segment and the vertical direction, as shown in Fig. 1b. The maximum of the absolute value of offset angle in every frame was then obtained.

In the side bending test, subject sat on a stool with the upper body straightened and foot on the ground naturally. Color markers remained at the same position of previous test and camera faced to the subject's back. At first the subject performed a lumbar side-bending toward the right-side, and then reverted to the left-side after reaching the limit. The line segment between the centroid of the color marks was captured, and the maximum bending angle between the line segment and the vertical direction was then calculated, as shown in Fig. 1c.

In DLLT, the subject lay on the floor with color markers attached to the left greater trochanter and outside of

the left knee joint. The subject raised legs until the legs were perpendicular to the ground, with the flattened PBU placed under their lumbar. Subject then exerted abdominal muscles to press the PBU, while the researcher pumped the PBU to 40 mmHg. Keeping their abdominal muscles tight, the subject steadily lowered their legs to the ground. The MATLAB computer program captured the line segment between centroid of the color marks and calculated the hip flexion angle from the line segment, as shown in Fig. 1d. The program also captured the reading of PBU pressure gauge. The hip flexion angle was calculated when the PBU pressure dropped to 30 mmHg.

In MTT, the subject lay on a table with their entire legs extending off of the edge of table. Four color markers were attached to the left greater trochanter, the outside of left knee joint, the upper end of left fibula, and the outside of left ankle joint. The subject braced their right knee against their chest while their contralateral leg hung freely. The PBU was placed under the subject's lumbar to ensure that anterior pelvic tilt did not occur. The researcher then shot the scene, then relocated the color marks to the contralateral side and repeated the process. The MATLAB program captured the line segment representing the thigh and the calf, then calculated the hip extension angle and knee flexion angle, as shown in Fig. 1e.

2.4 Grading Method

The evaluation system used the data of all intact group subjects as a control reference for the grading method. The mean value, the mean value plus one standard deviation, and the mean value minus one standard deviation were calculated to construct four levels of the grade. In the hip abduction test, forward bending test and DLLT, the four levels, ranked in ascending order, were given a score from 1 to 4 respectively. In the side bending test, MTT-hip angle and MTT-knee angle, the four levels were given from a score from 4 to 1 respectively. If the score was higher, the subject exhibited a higher risk of LBP. The summation of scores from the hip abduction test, forward bending test, side bending test, DLLT, MTT-hip angle and MTT-knee angle was therefore between six and 24. We used this summation score as an indicator for the risk of LBP.

2.5 Analysis Interface

To simplify the analysis process, the evaluation system included an analysis interface program compiled by MATLAB, as shown in the Fig. 1f. A user imported the video clip of each test into analysis interface by placing it in the same folder with the program. The calculation process of test result was designed in advance. User only needed to assign analyzing scope and distinguish threshold to make the image clear.

2.6 Experimental Design

Twenty-six subjects were evaluated by our developed evaluation system and the results were compared to determine the difference between intact group and patient group. To verify the test–retest reliability, five randomly-selected intact subjects also performed the same experiment again after 24 hours. To verify the expert validity, an experienced physical therapist was invited to view the video clips of experiments. Due to the limited of duration of the experiment, fifteen subjects (from the intact group and patients group) were selected to be viewed by the experienced physical therapist and classified into four levels of grade.

2.7 Data Analysis

One tailed independent t-tests were used to compare results between the two groups. The significant level (α) was 0.05. If the p value was smaller than 0.05, then the statistical power ($1-\beta$) would be calculated. A repeated measure two-way ANOVA was applied on the score of each test to analyze the variability of within-subject and between-subject variations. In the retest experiment, the five subjects' total scores of the first and second tests were compared by the two-tailed

paired t-test and were used to draw a two-dimensional distribution diagram. In the expert validation, the 15 subjects' total scores of the evaluation system and the experienced physical therapist evaluating results were compared by a two-tailed independent t-test and were used to draw a two-dimensional distribution diagram.

3 Results

3.1 Difference Between Groups

The differences of total scores and each test results between two groups is shown in Table 2. The average total score of the intact group was 14.9 ± 2.2 and the patient group was 19.0 ± 2.1 , showing a significant difference between the two groups ($p = 0.0004$). Comparing the test results respectively, hip abduction test ($p = 0.002$), forward bending test ($p = 0.029$), side bending test ($p = 0.001$) and DLLT ($p = 0.046$) also showed significant differences between the two groups. The results of hip angle ($p = 0.45$) and knee angle ($p = 0.45$) in MTT showed no significant differences between two groups. The statistical power of the t-test between the two groups in hip abduction test, side bending test, and total score was above 0.9.

The results of two-way ANOVA are shown in Fig. 2 and Table 3. The Mauchly spherical test showed no significant difference ($p = 0.057$). This meant that we could accept the null hypothesis that the covariance matrix is spherical. Table 3 presents the variability of the source of variations. The variability of different tests (SS_B) showed no significant difference ($p = 0.056$). The variability of interaction between test and groups (SS_{AB}) also showed no significant difference ($p = 0.184$). The variability of groups (SS_A) had significant difference ($p = 0.001$). The score of each test is shown in Fig. 2. The score of the patient group was higher than the intact group in every test. The difference of score

Table 2 Results of difference between groups

Group	IG	PG	P	Power
Hip abduction test (pixel length)	18.6 ± 4.3	27.6 ± 7.6	.002*	0.93
Forward bending test (degree)	5.8 ± 5.1	13.9 ± 16.5	.029*	0.61
Side bending test (degree)	17.6 ± 6.1	7.7 ± 3.6	.001*	0.96
DLLT (degree)	78.7 ± 7.7	84.9 ± 3.7	.046*	0.52
MTT hip angle (degree)	5.1 ± 8.7	4.5 ± 8.8	45	N/A
MTT Knee angle (degree)	71.7 ± 12.6	70.1 ± 10.3	45	N/A
Total score (score)	14.9 ± 2.2	19.0 ± 2.1	.0004*	0.98

Mean $\pm$ standard deviation, independent t-test, one-tailed, * $p < 0.05$
IG intact group, PG patient group

Fig. 2 Score of each test in two groups

Table 3 Result of two-way ANOVA

Source of variation	DOF	F	P
Within-subject			
Text SS _B	5	2.226	0.056
Interaction SS _{AB}	5	1.535	0.184
Between-subject			
Group SS _A	1	14.519	0.001

was high in the hip abduction test and side bending test; medium in forward bending test and DLLT; and similar in MTT-thigh and MTT-knee.

3.2 Test-Retest Reliability and Expert Validity

The results of the retest experiment is shown in Fig. 3. There was no significant difference between total score of the first test and second test ($p=0.058$). In the two-dimensional distribution diagram, the correlation coefficient (Pearson's r) was 0.98, the coefficient of determination (R^2) was 0.99, and the linear regression slope was 0.45. This showed that there was a moderate linear correlation between the two tests. The results of the expert validation are shown in Fig. 4. There was no significant difference between the expert and the system score ($p=0.141$). In the two-dimensional distribution diagram, the correlation coefficient was 0.90, the coefficient of determination was 0.82, and the linear regression slope was 0.77. This showed that there was a moderate linear correlation between the two scores.

4 Discussion

The main propose of this study was to establish an LBP evaluation system for non-professional operators. Our hypothesis was that the results had significant difference between the intact and patient groups. In addition, the two set of results in the retest experiment and the expert validation should show no significant difference and show a linear correlation.

The evaluation system improves the convenience of conducting examinations. The original hip abduction test and forward bending test required a therapist judging the deviation of foot and spine. In contrast, our system used a camera

Fig. 3 Result of retest experiment. **a** Average total score of first test was 15.0 ± 4.3 (score). Second test was 12.2 ± 1.9 (score). No significant difference ($p=0.058$, paired t-test, two-tailed). **b** Two dimensional distribution diagram: Pearson's $r=0.98$, $R^2=0.99$, linear regression slope = 0.45

Fig. 4 Result of expert validation. **a** Average total score of expert was 13.6 ± 4.0 (score). System evaluating was 15.7 ± 3.4 (score). No significant difference ($p=0.141$, Independent t-test, two-tailed). **b** Two

dimensional distribution diagram: Pearson's $r=0.90$, $R^2=0.82$, linear regression slope = 0.77

and a custom-made image analysis program so that users could determine outcome by operating the program on a computer. In the original DLLT, a tester had to reach one hand under the subject's lumbar region to detect the moment of lumbar departing from the floor and measured angle by halting the subject's legs. In contrast, our system used a PBU to detect this moment so that there was a clear standard for non-professional operators.

4.1 Difference Between Groups

From the comparison of separate tests, the hip abduction test and the side bending test showed significant difference between groups and had high statistical power, suggesting that they are capable of identifying LBP risk. Although the p values reached significant level in forward bending test and DLLT, the statistical powers were relatively low, suggesting that they only had mediocre capability. The result of MTT-hip angle and MTT-knee angle showed no significant difference between groups. There were two possible reasons: (i) the current study recruited LBP patients from an orthopedic clinic which may have not included subjects with tight hip flexors; and (ii) when performing MTT, the current study asked the subject read the pressure of PBU to adjust the location of his knee. However, this may have distracted the subject and preventing the leg from totally relaxing. The average total score of patient group was significantly higher than intact group with good statistical power.

From the two-way ANOVA, the variance between tests was not significant ($p=0.056$), but the p value approached

0.05. This might result from the fact that the score of MTT-thigh and MTT-knee were close (Fig. 2). The variance of interaction was not significant ($p=0.184$) and the variance between groups was significant ($p=0.001$). This demonstrated that the group of subjects was the major factor of the variation of score in most of the examinations. In summary, it could be seen that despite MTT been not effective, the total score of system was capable of distinguishing between the patients and the intact group.

4.2 Test-Retest Reliability and Expert Validity

In the retest experiment, total score of the first test and the second had no significant difference ($p=0.058$), but the p value approached 0.05. This might result from the fact that the interval time between two trials was not long enough. The subject had learning effect, so the second test performed better than the first test. Extending the interval time between tests is a possible way in the future to limit this potential learning effect. The two-dimensional distribution diagram showed high correlation coefficient and R^2 , but the linear regression slope was low (0.45). This might also result from the difference between two trials. In summary, the test-retest reliability of the system was acceptable.

In the expert validation, total score of expert and system evaluating had no significant difference ($p=0.141$). The two-dimensional distribution diagram showed high correlation coefficient and R^2 , but the linear regression slope was rather low (0.77). This might result from the fact that the therapist used to rehabilitate patients having much worse

LBP than our subjects. Therefore, the expert evaluating score was lower than that from our system. In summary, the expert validity of the system was also acceptable.

4.3 Limitations

The devices of current system, including camera, PBU and tripods, were separated. The relative position among them had to be adjusted between examinations, lowering the convenience. This could be improved by integrating the devices into one in the future. In terms of image analysis, the MTT showed no significant difference between groups. The process of MTT should be redesigned or excluded from system, depending on results from additional studies. In addition, we summed the score of different examinations with same weighting. Setting different weighting factor for each examination can be examined in future research. Besides, the recruited subjects were not gender and age matched between groups. The intact group was younger and with higher ratio of males, resulting in enlargement of the difference between groups. This may require more subjects with matched gender and age in the future work.

5 Conclusion

We have successfully developed a LBP image evaluation system for distinguishing non-specific LBP patients and asymptomatic individuals. The system also showed acceptable test–retest reliability and expert validity. The system appears to have potential utility as a screening tool to determine the severity of LBP for non-professional operators. Future work is needed to improve the design and effectiveness of the system with a wider spectrum of subjects.

Funding Partially supported by Taiwan Ministry of Science and Technology Grant # MOST 108-2221-E-009-056.

Compliance with Ethical Standards

Ethics Approval Protocol approved by the Research Ethics Committee for Human Subject Protection, National Chiao Tung University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long

as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Hoy, D., et al. (2012). A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism*, 64(6), 2028–2037.
2. Chou, R., et al. (2007). Diagnosis and treatment of low back pain: A joint clinical practice guideline from the American college of physicians and the American pain society. *Annals of Internal Medicine*, 147(7), 478–491.
3. Nelson-Wong, E., Flynn, T., & Callaghan, J. P. (2009). Development of active hip abduction as a screening test for identifying occupational low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 39(9), 649–657.
4. Nelson-Wong, E., & Callaghan, J. P. (2014). Transient low back pain development during standing predicts future clinical low back pain in previously asymptomatic individuals. *Spine*, 39(6), E379–E383.
5. Biely, S. A., et al. (2014). Clinical observation of standing trunk movements: What do the aberrant movement patterns tell us? *Journal of Orthopaedic & Sports Physical Therapy*, 44(4), 262–272.
6. Wong, T. K. T., & Lee, R. Y. W. (2004). Effects of low back pain on the relationship between the movements of the lumbar spine and hip. *Human Movement Science*, 23(1), 21–34.
7. Youdas, J. W., et al. (2000). Lumbar lordosis and pelvic inclination in adults with chronic low back pain. *Physical Therapy*, 80(3), 261–275.
8. Krause, D. A., et al. (2005). Abdominal muscle performance as measured by the double leg-lowering test. *Archives of Physical Medicine and Rehabilitation*, 86(7), 1345–1348.
9. Van Dillen, L. R., et al. (2000). Effect of knee and hip position on hip extension range of motion in individuals with and without low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 30(6), 307–316.
10. Peeler, J., & Leiter, J. (2013). Using digital photography to document rectus femoris flexibility: A reliability study of the modified Thomas test. *Physiotherapy Theory and Practice*, 29(4), 319–327.
11. Clapis, P. A., Davis, S. M., & Davis, R. O. (2008). Reliability of inclinometer and goniometric measurements of hip extension flexibility using the modified Thomas test. *Physiotherapy Theory and Practice*, 24(2), 135–141.
12. Janz, N. K., & Becker, M. H. (1984). The health belief model—A decade later. *Health Education Quarterly*, 11(1), 1–47.