

The inter-observer reproducibility of Lasègue's sign in patients with low back pain in general practice

HANS J M VAN DEN HOOGEN

BART W KOES

WALTER DEVILLÉ

JACQUES TH M VAN EIJK

LEX M BOUTER

SUMMARY

Background. *The spectrum of low back pain patients in general practice differs significantly from that in an orthopaedic clinic. The most frequent specific cause of low back pain is nerve-root irritation or compression caused by intervertebral protrusion, and the diagnosis is still problematic. Testing for Lasègue's sign could be a useful way of detecting high-risk patients, but so far the reproducibility of the test has been measured only in hospital-based studies.*

Aim. *To assess the inter-observer reproducibility of Lasègue's sign in general practice.*

Method. *Fifteen General practitioners from Amsterdam and the surrounding areas tested all consecutive low back pain patients who visited them during a period of two years for Lasègue's sign. The test was repeated within two weeks in two samples: sample I consisted of 50 consecutive low back pain patients; sample II consisted of all patients who had pelvic tilt, scoliosis, or positive Lasègue's sign.*

Results. *In sample I, the observation was repeated in 49 patients. The Kappa coefficient was 0.33, and the proportions of positive and negative agreement were 33% and 96%, respectively. In sample II, the observation was repeated in 48 patients. The Kappa coefficient was 0.56, whereas the proportion of positive agreement was 67% and the proportion of negative agreement was 91%.*

Conclusions. *The reproducibility of Lasègue's sign in routine general practice seems to be low, but may be similar to the reproducibility observed in hospital settings in selected patients who have a high chance of low back pain owing to a specific disease.*

Keywords: *Backache; Lasègue's sign; Netherlands.*

Introduction

THE most frequent specific cause of low back pain is nerve-root irritation or compression caused by intervertebral disk protrusion.¹⁻⁵ Diagnosing nerve-root irritation or compression

Hans J M van den Hoogen, MD, general practitioner; Bart W Koes, PhD, epidemiologist; Walter Devillé, MD, biostatistician; Jacques Th M van Eijk, PhD, professor of general practice; and Lex M Bouter, PhD, professor of epidemiology. All five authors are affiliated to The Institute for Research in Extramural Medicine; the Department of General Practice, Nursing Home Medicine and Social Medicine; and the Department of Epidemiology and Biostatistics, Faculty of Medicine, Vrije Universiteit, Amsterdam, The Netherlands.

Submitted: 26 October 1995; accepted: 27 July 1996.

© British Journal of General Practice, 1996, 46, 727-730.

still constitutes a problem. Anatomic evidence of disk protrusion is found in about 20–30% of imaging tests among people without low back pain.^{1,6} Therefore, signs and symptoms are of major importance in determining a diagnosis, and also in the choice of therapeutic management.¹⁻³ For theoretical reasons, and also on the basis of empirical evidence, testing for Lasègue's sign (i.e. sciatic pain experienced by a patient during the straight-leg raising test) could be a suitable method of testing for nerve-root irritation.⁷⁻¹⁰ In the straight-leg raising test, the straightened leg of the supine patient is raised until it cannot be raised any further. The test is interpreted in various ways. If the leg cannot be raised to approximately 90 degrees, movement is considered to be limited. If the test causes pain radiating into the raised leg, Lasègue's sign is considered to be present, and if it causes pain radiating into the leg which is not being raised, the crossed Lasègue's sign is considered to be present.

Limited straight-leg raising and Lasègue's sign both appeared to be highly sensitive signs for surgically proven disk protrusion. Moreover, Lasègue's sign appeared to combine this high sensitivity with a fairly high specificity.¹¹ The potential role of this test in diagnosing low back pain gives rise to questions regarding its reproducibility.^{12,13} The reproducibility of the straight-leg raising test has been tested in two previous hospital-based studies. The first study, by Waddell *et al*, in a group of low back pain patients of an orthopaedic outpatient clinic ($n = 30$), resulted in agreement between the two observers on 86% of the observations and a 0.45 Kappa coefficient.¹⁴ The most recent study, by McCombe *et al*, in two groups of patients from two orthopaedic practices ($n = 50$ and $n = 33$), resulted in 0.66 and 0.44 Kappa coefficients, respectively.¹⁵

The spectrum of low back pain patients in general practice differs significantly from that in an orthopaedic clinic. In general practice, Lasègue's sign is expected to be absent in almost all patients with low back pain. Furthermore, the general practitioner will be less experienced in applying the straight-leg raising test, since it is applied only a few times each week. Consequently, the reproducibility of Lasègue's sign in general practice may well be less than that found in these orthopaedic practices.^{12,13,16,17} Yet, it is in general practice that patients with low back pain will first be encountered, and the choice of therapeutic management should be based on the determined diagnosis.

Methods

Design of the study

An inter-observer reproducibility study was carried out in which a group of general practitioners tested consecutive low back pain patients for Lasègue's sign. The test was repeated within two weeks, in two samples of patients, by another general practitioner.

Study sample

The study was set in 11 general practices attended by 15 general practitioners from the Dutch capital of Amsterdam and the surrounding areas, with a catchment population of about 26 000 persons. Eligible for inclusion in this study were patients who consulted these 11 practices for low back pain of any duration

between May 1990 and May 1992. Additional criteria were age (over 16 years) and complaints of pain in the back (or radiating from the back) in the area between Th12 and the gluteal fold. Because of the low prevalence of Lasègue's sign, reproducibility was tested in two different samples. Sample I consisted of 50 consecutive patients who were included at the end of the first year of the inclusion period. The prevalence of Lasègue's sign in this sample, as in the general population, was expected to be low. Sample II consisted of all patients with pelvic tilt, scoliosis, or Lasègue's sign assessed by the general practitioner at the initial visit. The prevalence of Lasègue's sign in this second sample was artificially higher than in the first sample.

Measurements

At the initial visit, the general practitioner completed a form on the history of the complaints and carried out a physical examination including the straight-leg raising test. This test was repeated shortly after the initial visit by one of the authors (HJMvdH), a general practitioner.

The straight-leg raising test

The physician, supporting the heel, raised the relaxed and straightened leg of the supine patient until it could not be raised any further. Lasègue's sign was considered to be present if the raising provoked sciatic pain beyond the knee.^{7,9} The participating general practitioners were trained to carry out and interpret the straight-leg raising test in a two-hour training session prior to and shortly after the start of the study. Pelvic tilt was tested and measured by levelling the pelvis by means of wooden boards of different thickness. Scoliosis was tested by examining, from behind, for evidence of a hump in a patient bending forwards.

The 50 consecutive low back pain patients of the first sample were identified to the researcher by the research assistant, who received the mailed data of patients included by the general practitioners. In the second sample, the patients were identified to the researcher by the practice nurses of the general practices involved. The researcher made appointments to visit the patients in their homes as soon as possible, in order to repeat the straight-leg raising test.

The different ways of identifying patients in the two samples reduced the blindness of the researcher to the results of the straight-leg raising test performed by the general practitioner. The prevalence of Lasègue's sign was expected to be low in patients identified by the research assistant, and higher in the other patients. Apart from this, the blindness of the researcher to further information was maintained as far as possible. The researcher was not informed about the results of the assessment of the patients at the initial visit.

The presence of radiating pain in the straight-leg raising test will, in turn, raise the expectation of a positive result for Lasègue's sign.^{12,13} In order to assess the potential bias resulting from this effect, the presence of radiating pain (up to or beyond the knee) was recorded as a part of the history-taking at the initial visit. Furthermore, low back pain, including the accompanying signs and symptoms, tends to resolve in time.^{1,3,18} Therefore, the results of the repeated observations may be affected by the length of time between the two observations. In order to assess any bias arising from the time lag between the two measurements, the number of days between the two observations was recorded.

Analysis

The inter-observer reproducibility of Lasègue's sign in the two samples was assessed using Cohen's Kappa, a chance-corrected

coefficient for the agreement between two observers.¹⁹ No clear-cut interpretation of the Kappa coefficient can be given, although some authors suggest a grading in which Kappa values from 1 to 0.75, from 0.75 to 0.4, and from 0.4 to -1, would indicate an excellent, a fair to good, and a poor agreement, respectively.²⁰ A low prevalence of one of the test results, despite high percentages of agreement, may result in a paradoxically low Kappa coefficient.²¹ If both observers only have to confer on a small number of positive findings, any single agreement or disagreement in these positive cases will have a relatively large influence on the calculated Kappa. According to Cicchetti and Feinstein, in order to obtain a better understanding of the results, Kappa coefficients should be accompanied by separate values for the observed proportions of positive and negative agreement.²² Consequently, in our study the observed proportions of positive and negative agreement were calculated and presented in addition to the Kappa coefficient. The two samples were analysed separately because of the different ways the patients were identified and the reduced blindness of the researcher.

The presence of radiating pain and the length of time between the two observations may both have influenced the results of the reproducibility of the repeated observations. In both samples, the reproducibility was calculated for patients with and without radiating pain, and also for patients for whom the number of days between the two observations was (in turn) less than or equal to the median, and more than the median. In the calculation of all indices, the group of participating general practitioners was taken to represent one observer, and the researcher represented the other observer (see Appendix).

Results

Sample I. Consecutive low back pain patients

In this sample of consecutive low back pain patients, the mean age was 46 years; 50% were men. The median duration of the low back pain at the initial visit was 5 days (interquartile range 3–20). Pain radiating into the leg up to the knee was reported in 24% of the patients, and pain radiating beyond the knee was reported in 26% of the patients. Sixty-four per cent of patients reported a sudden onset in low back pain, and 6% of the patients had a history of low back surgery. At the initial visit, four patients were diagnosed as possibly having specific low back pain (i.e. low back pain due to nerve-root irritation, neoplastic disease, or ankylosing spondylitis). In only one patient was this confirmed by a consultant to be low back pain due to disc protrusion. In one case the test was not repeated because the researcher did not succeed in making an appointment with the patient.

The results of the repeated straight-leg raising test are presented in Table 1. The Kappa coefficient was calculated to be 0.33 (CI -0.31–0.97), whereas the proportion of positive agreement was 33% and the proportion of negative agreement 96%.

Table 1. Reproducibility of Lasègue's sign in 49 consecutive low back pain patients.

Repeated SLR*	SLR at initial visit* Lasègue's sign	
	Present	Not present
Lasègue's sign Present	1	3
Not present	1	44
Total	2	47

*SLR = straight-leg raising test. Kappa = 0.33; $P_{\text{pos}} = 33\%$; $P_{\text{neg}} = 96\%$.

Kappa coefficient in patients with radiating pain was 0.33 ($n = 26$), whereas the proportion of positive agreement was 40% and the proportion of negative agreement 94%. Of the patients without radiating pain ($n = 23$), none were found to have a positive test when the test was repeated. Consequently, only the proportion of negative agreement could be calculated, which was found to be 98%. The median time between the initial visit and the repeated assessment was 8 days (range 0–30, interquartile range 4–10). Of the patients with fewer than 8 days between the two observations, none were found to have any positive tests for the two measurements. Consequently, no reproducibility coefficients could be calculated. For the patients with at least 8 days between the two observations, the Kappa coefficient was 0.50, whereas the proportion of positive agreement was 50% and the proportion of negative agreement 95%, which is even higher than in the complete sample.

Sample II. Patients with pelvic tilt, scoliosis or Lasègue's sign

At the initial visit, Lasègue's sign appeared to be present in 31 of the 605 patients, whereas pelvic tilt or scoliosis was indicated in 37 patients. Of these patients, two had already been included in the sample of 50 consecutive patients. Consequently, sample II consisted of 66 patients. Straight-leg raising (SLR) was repeated in 48 of these patients. In the remaining 18 patients the test was not repeated, either because the patient was not identified to the researcher ($n = 12$) or because the researcher did not succeed in making an appointment ($n = 6$). At the initial visit, 30 of the 66 patients were diagnosed as possibly having specific low back pain. This was confirmed by a consultant to be low back pain due to disc protrusion in eight of these patients and to be low back pain due to ankylosing spondylitis in one patient. The mean age was 40 years, and 53% of the patients were men. The median duration of the low back pain at the initial visit was 14 days (interquartile range 5–49). Pain was reported radiating into the leg up to the knee in 29% of the patients, and beyond the knee in 32% of the patients. Thirty-six percent of patients reported a sudden onset of low back pain, and 14% of the patients had a history of low back surgery. The results of the repeated straight-leg raising test are presented in Table 2. The Kappa coefficient was 0.56 (CI 0.26–0.86), whereas the proportion of positive agreement was 67% and the proportion of negative agreement 91%. The Kappa coefficient in patients with radiating pain ($n = 25$) was 0.44, whereas the proportion of positive agreement was 70% and the proportion of negative agreement 74%. Of the patients without radiating pain ($n = 23$) none were found to have any positive tests for the two measurements. Consequently, no reproducibility coefficients could be calculated. For the patients with fewer than 8 days between the two observations, the Kappa coefficient was 0.57, whereas the proportion of positive agreement was 67% and the proportion of negative agreement was 89%. For the patients

with at least 8 days between the two observations, the Kappa coefficient was 0.60, whereas the proportion of positive agreement was 67% and the proportion of negative agreement 93%. Therefore, the delay between the two measurements was again of no significant importance.

Discussion

The results of studies in hospital-based settings may be of limited value in judging reproducibility in general practice. However, routine general practice is clearly not the most suitable setting in which to study the optimal reproducibility of a test. Most pathological signs, including those in low back pain, have a low prevalence in general practice. Reproducing the tests for these signs for all patients would be ineffective. Moreover, because of the low prevalence, disagreements on the small number of positive findings may have a large influence on the Kappa coefficient.^{19,20}

In the present study, we have tried to counteract this problem by creating a separate sample with an artificially high prevalence of Lasègue's sign. In addition, the presence of radiating pain may influence the interpretation of the straight-leg raising test, since it increases the expectation of nerve-root irritation or compression.^{12,13} However, in the present study, the Kappa coefficient, and also the proportion of positive and negative agreement in patients with radiating pain, appeared to be quite similar to the values found in the complete samples. Finally, because of its natural course, Lasègue's sign may disappear in the interval between the two observations.^{1,3} In the present study, the delay between the two measurements appeared to have been of no significant importance.

The design of the study, using two different samples, and the difference in the results of both samples do not enable a straightforward conclusion to be drawn on inter-observer reproducibility. In sample I, consisting of consecutive low back pain patients, both the Kappa coefficient and the proportion of positive agreement appeared to be much lower than in sample II, which consisted of selected patients with a higher prevalence of Lasègue's sign. The Kappa coefficient in sample II appeared to be quite similar to the coefficients found in two previous hospital-based studies.^{12,13} The proportion of negative agreement in both samples appeared to be high.

The results in sample I probably reflect the reproducibility of the test being applied to consecutive patients in routine daily practice, whereas the results in sample II may reflect the reproducibility for selected patients with a high chance of low back pain owing to a specific disease. The selected patients in sample II are probably similar to the patients seen in hospital settings. Consequently, we argue that the results indicate that the reproducibility of Lasègue's sign in routine general practice seems to be low, whereas it may be similar to the reproducibility in hospital settings for selected patients with a high chance of low back pain owing to a specific disease. The high proportion of negative agreement indicates that, if it was found to be absent in the first straight-leg raising test, Lasègue's sign will also be absent if the test is repeated. Consequently, if Lasègue's sign is found to be absent, repeating the test (and carrying out further assessment, followed eventually by imaging tests) would not be warranted.

The lack of positive agreement in both samples indicates that a positive finding may well be followed by a negative finding if the straight-leg raising test is repeated. Consequently, if Lasègue's sign is found to be present, the general practitioner should consider repeating the straight-leg raising test. Only a duplicated positive finding, or additional other findings, would warrant further assessment and subsequent imaging tests.

Table 2. Reproducibility of Lasègue's sign in all patients with pelvic shift, scoliosis or Lasègue's sign present at the initial visit ($n = 48$).

Repeated SLR	SLR at initial visit Lasègue's sign	
	Present	Not present
Lasègue's sign		
Present	7	2
Not present	5	34
Total	12	36

Kappa = 0.56; $P_{\text{pos}} = 67\%$; $P_{\text{neg}} = 91\%$.

Appendix

When two observers (I and II) perform a test to determine the presence or absence of a particular entity, the results can be arranged in a fourfold table, as shown.

Observer II Entity	Observer I Entity		Total
	Present	Not present	
Present	a	b	g_1
Not present	c	d	g_2
Total	f_1	f_2	N

The Kappa, the proportion of positive agreement and the proportion of negative agreement can be calculated from this table as follows:

$$P_o \text{ Observed total agreement} = (a + d) / N$$

$$P_e \text{ Expected total agreement by chance} = (f_1 g_1 + f_2 g_2) / N^2$$

$$\text{Kappa} = \frac{P_o - P_e}{1 - P_e}$$

$$P_{\text{pos}} \text{ Proportion of positive agreement} = \frac{2a}{f_1 + g_1}$$

$$P_{\text{neg}} \text{ Proportion of negative agreement} = \frac{2d}{f_2 + g_2}$$

References

- Spitzer WO, LeBlanc FE, Dupuis M, *et al*. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. *Spine* 1987; **12**: S1-S9.
- Deyo RA, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992; **268**: 760-765.
- Nachemson AL. Newest knowledge of low back pain. A critical look. *Clinical Orthopaedics and Related Research* 1992; **279**: 8-20.
- Dillane JB, Fry J, Kalton G. Acute back syndrome. A study from general practice. *BMJ* 1966; **268**: 82-84.
- Heliövaara M, Knekt P, Aromaa A. Incidence and risk factors of herniated lumbar intervertebral disc or sciatica leading to hospitalization. *J Chronic Diseases* 1987; **40**: 251-285.
- Jensen MC, Brant-Zawadzki MN, Obuchowski N, *et al*. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994; **331**: 69-73.
- Forst JJ. Contribution à l'étude clinique de la sciatique. Lyon, France: Thesis, 1881.
- Sugar O. Charles Lasègue and his 'considerations on sciatica'. *JAMA* 1985; **253**: 1767-1768.
- Sjöqvist O. The mechanism of origin of Lasègue's sign. *Acta Psychiatr Neurol* 1947; **46**: 290-297.
- O'Connell JEA. Protrusions of the lumbar intervertebral discs. A clinical review based on five hundred cases treated by excision of the protrusion. *Journal of Bone and Joint Surgery* 1951; **33B**: 8-30.
- Hoogen HMM van den, Koes BW, Eijk JThM van, Bouter LM. On the accuracy of history, physical examination, and erythrocyte sedimentation rate in diagnosing low back pain in general practice. A criteria-based review of the literature. *Spine* 1995; **20**: 318-327.
- Koran LM. The reliability of clinical methods, data and judgments. (First of two parts). *N Engl J Med* 1975; **293**: 642-646.
- Koran LM. The reliability of clinical methods, data and judgments. (Second of two parts). *N Engl J Med* 1975; **293**: 695-701.
- Waddell G, Main CJ, Morris EW, *et al*. Normality and reliability in the clinical assessment of backache. *BMJ* 1982; **284**: 1519-1523.
- McCombe PF, Fairbank JCT, Cockersole BC, Pynsent PB. Reproducibility of physical signs in low-back pain. *Spine* 1989; **14**: 908-918.
- Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978; **299**: 926-930.
- Knottnerus JA, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol* 1992; **45**: 1143-1154.
- Porter RW, Trailescu IF. Diurnal changes in straight leg raising. *Spine* 1990; **15**: 103-106.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; **20**: 37-46.
- Seigel DG, Podgor MJ, Remaley NA. Acceptable values of Kappa for comparison of two groups. *Am J Epidemiol* 1992; **135**: 571-578.
- Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; **43**: 543-549.
- Cicchetti DV, Feinstein AR. High agreement but low Kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990; **43**: 551-558.

Do you have patients between 10 and 16 years who are frequently hospitalised and missing schooling through asthma, eczema or other chronic medical conditions. We CAN help them to lead normal lives.



Pilgrims is Britain's only specialist school providing residential education for youngsters with asthma and eczema. We can also help children with other chronic medical conditions. All pupils have a statement of Special Educational Needs and funding is provided by the LEA, often in conjunction with Health or Social Services. **Pilgrims**, with 24 hours a day, qualified medical cover, enables youngsters to reduce their medication, minimise hospital admissions and control their condition.

Further information from the Head teacher, Janice Richardson, MA BEd, **Pilgrims**, (An I CAN School) Firle Rd, Seaford, East Sussex BN25 2HX Telephone 01323 892697 Fax 01323 892503.