Are Relieving Maneuvers Useful in Diagnosis of Carpal Tunnel Syndrome?

Rahatlatıcı Manevralar Karpal Tünel Sendromu Tanısında Yararlı mı?

Haydar Gök¹, Saime Ay², Sehim Kutlay¹
¹Ankara Üniversitesi Tıp Fakültesi, Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, Ankara, Türkiye
²Ufuk Üniversitesi Tıp Fakültesi, Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, Ankara, Türkiye

Abstract

Objective: Our aim was to assess and compare the diagnostic accuracy of the carpal tunnel syndrome relief maneuver (CTS-RM) and the Flick sign for the diagnosis of carpal tunnel syndrome (CTS) in female patients.

Materials and Methods: This is a diagnostic test study with blind comparison to a reference criterion. A total of 87 consecutive female patients with typical symptoms for CTS referred for electrophysiological examination were included in the study. Normal limits of nerve conduction were obtained from 50 healthy female subjects. After the electrodiagnostic assessment clinical evaluation was performed by a physician and it included testing of all patients for the CTS-RM and Flick sign. The diagnostic accuracy was evaluated for each test alone and in combination and sensitivity was correlated with the electrophysiological severity of CTS. Main outcome measures included the estimates of sensitivity, specificity, positive and negative likelihood ratios (PLR, NLR).

Results: After electrophysiological assessment, 58 patients (pre-test probability, 67%) have been diagnosed as CTS. The sensitivity and specificity estimates were %81,86 for the CTS-RM and %69,79 for the Flick maneuver. Combining a positive CTS-RM and Flick sign improved the specificity to 93%. The PLRs of the CTS-RM and Flick sign were 3.3 and 5.9 and the NLRs were 0.39 and 0.22 respectively. Combining a positive CTS-RM and Flick sign had the PLR of 9.5 and the NLR of 0.37. When evaluating the subjects with CTS, the CTS-RM detected significantly more subjects compared to the Flick sign.

Conclusions: Our study reveals that the accuracy of Flick sign is low in the diagnosis of CTS. While the CTS-RM alone is helpful in confirming the diagnosis in patients with typical symptoms, combination with the Flick sign further improves its predictive accuracy.

(Rheumatism 2008; 23: 129-34)

Key words: Carpal tunnel syndrome, diagnostic testing, maneuver, likelihood ratio

Received: 21.09.2008 Accepted: 11.11.2008

Özet

Amaç: Amacımız kadın hastalarda karpal tünel sendromu (KTS) tanısında, rahatlatıcı manevra (KTS-RM) ve Flick belirtisinin tansal değerini ortaya koymaktı.


Bulgular: Elektrofizyolojik değerlendirmeye sonucunda 58 hastada KTS tanısı konuldu (%67). KTS-RM ve Flick manevraları için duyarlılık ve özgüllük oranları sırasıyla %81,86 ve %69,79 olarak bulundu. KTS-RM ve Flick belirtilinin aynı anda pozitif olmasının durumunda, özgüllük oranının %93’e yükseldiği görüldü. KTS-RM ve Flick manevraları için PLR oranları 3.3 ve 5.9, NLR oranları ise .39 ve .22 olarak bulundu. KTS-RM ve Flick belirtilinin aynı anda pozitif ol- masının durumunda, PLR oranı 9.5, NLR oranı 0.37 olarak hesaplandı. KTS tanısı konulan denekler değerlendirildiğinde, KTS-RM’nin Flick manevrasına göre anlamlı oranda daha fazla KTS saptayabilidir. Elektrodiyagnostik açıdan hastalığın jözdeti altıksız manevra- ların duyarlılığının artına eğitiminde olduğu görüldü.

Sonuç: Bu çalışma Flick belirtilinin KTS tanısı koymada yetersiz olduğunu göstermektedir. KTS-RM, tipik semptomları olan hastalarda KTS tanısını doğrulamak için tek başına yeterlidir. Flick belirtisi ile beraber ele alındığında doğrulama gücü daha da artmaktadır.

(Romatizma 2008; 23: 129-34)

Anahtar sözcükler: Karpal tünel sendromu, tansal test, manevra, olabilirlik oranı

Introduction

Carpal tunnel syndrome (CTS) is the most common nerve compression disorder of the upper extremity. According to population-based studies, about 3% of adults have electrodiagnostically confirmed CTS and women are affected 3 times more than men (1). While clinicians use electrodiagnosis frequently to confirm the diagnosis of CTS and some third-party payers require it before compensating claims, there has been a growing interest in developing a useful predictive diagnostic test based solely on physical examination.

Various clinical tests that rely on provoking the symptoms have been proposed for the evaluation of patients suspected to have CTS and some of them have become a part of the routine investigative procedure (e.g. Phalen, tunnel). However, clinical tests based on relieving the symptoms are limited. Among these is the carpal tunnel syndrome relief maneuver (CTS-RM), a new technique developed by Manente et al. (2). In an attempt to develop a provocative test by squeezing the distal heads of the metacarpal bones together, they observed relief of symptoms serendipitously. Because of its high sensitivity and specificity, they suggested the CTS-RM to be a useful test in confirming the diagnosis of CTS in patients with positive symptoms. However, the diagnostic importance of this potentially useful maneuver has not been validated by other studies.

Another maneuver that relieves patients’ symptoms is the flicking motion of the hands and wrist called The Flick sign. Typically, patients flick their affected hand at times when symptoms are at their worst especially during sleep. Clinicians frequently look for this sign as part of their clinical evaluation before referring to electrodiagnostic evaluation. In this sense, the Flick sign may be considered as a clinical information based on the patient’s medical history rather than a true clinical test based on physical examination like Phalen or 2-point discrimination (3). In some studies it was also utilized as a bedside evaluation tool in which patients are asked what they do with the affected hand at times when symptoms are at their worst (4). Although, the Flick sign is seen frequently in patients with CTS, it has not been well studied and its clinical utility has been limited due to the varying estimates of diagnostic accuracy in previous studies (4-8).

The aim of this study was to evaluate and compare the diagnostic accuracy of the CTS-RM and the Flick sign in diagnosis of carpal tunnel syndrome in female patients. The secondary purpose of this study was to examine the effects of disease severity and combination of maneuvers on diagnostic accuracy.

Materials and Methods

Data were collected from 100 consecutive female subjects with complaints of paresthesia, numbness and pain in hands consistent with CTS who were referred to electrodiagnostic evaluation. 13 patients with systemic etiological factors including diabetes mellitus, connective tissue diseases, kidney and thyroid diseases, peripheral neuropathy or with onset of symptoms after trauma were excluded from the study. To avoid the effects of dependence between hands, only the involved side was selected in unilateral cases. In cases of bilateral involvement, the more affected side was chosen based on patient’s symptoms. A total of 87 patients were included in the study.

Nerve Conduction Studies

Electrodiagnostic studies were performed with DISA Nueromatic 2000C ENMG device according to a protocol (9) inspired by the recommendations of the American Association of Electrodiagnostic Medicine (10). Normal values of nerve conduction tests were obtained form 50 asymptomatic healthy female volunteers. Median sensory conduction velocity was measured antidromically by a ring electrode on the 2nd digit (normal: ≥47.5 m/s). Median distal motor latency (DML) was measured with a surface electrode placed on abductor pollicis brevis muscle (normal: ≤4.0 ms). When these studies are normal, median-to-ulnar latency difference was calculated by stimulating the ring finger (normal: ≤0.5 ms) (11). If abnormalities were observed in the median and ulnar nerves of the same limb, other limb and one lower limb were examined to rule out a generalized polyneuropathy. Electrophysiological evaluation was normal in all control subjects. Patients diagnosed as CTS were divided into three groups using an arbitrary scale according to the values of DML; 1) 4.1-4.4 ms (mild), 2) 4.5-5.0 ms (moderate), 3) ≥5.1 ms (severe).

Clinical Evaluation

After the electrodiagnostic assessment, all patients were examined by a physician who was blind to the results of nerve conduction study. Physical evaluation included testing of all patients for the CTS-RM and Flick sign. Testing was performed in random order since it might be possible that they may have a cumulative effect that increases positive results for the test that is performed last.

Flick sign was elicited by asking the patients what they do with the affected hand at times when symptoms are at their worst. If the patient showed a flicking motion of the hands and wrists, the maneuver is considered to be positive.

The CTS-RM was performed as prescribed by Manente et al. (2). The affected hand was maintained with palm up and the distal heads of metacarpal bones (excluded first) were gently squeezed inducing a slight adduction of digits II and V. When this was not sufficient to relieve symptoms, palm was turned down and the digits III and IV were stretched simultaneously. Patients were blinded to the possible effects of the maneuvers and asked to indicate whether the maneuver: (1) abolished; (2) improved; (3) worsened; or (4)
did not change their symptoms. Results were dicotomized in such a way that either an improvement or abolition of symptoms was considered as positive.

**Statistical Analysis**

Main outcome measures included the estimates of sensitivity, specificity, positive and negative likelihood ratios (PLR, NLR). The likelihood ratios (LRs) incorporate both the sensitivity and specificity and provide a direct estimate of how much a test result (positive or negative) will change the odds of having a disease. The diagnostic accuracy of any maneuver is considered useful if the PLR is 2.0 (to rule in disease) or greater or if the NLR is 0.50 or less (to rule out disease) (12). Since the prevalence (pre-test probability) in our study (67%) was different from the true prevalence of CTS in the population, likelihood ratios were weighted for prevalence according to the Bayes’ Theorem (\(\text{Odds}_{\text{post}} = \text{Odds}_{\text{pre}} \times \text{likelihood ratio}\)) to calculate post-test probabilities. Posterior probabilities were then compared with the prior probability. According to Jaeschke et al., LRs of 2 to 5 and 0.5 to 2 would generate small shifts, LRs of 5 to 10 and 0.1 to 0.2 would generate moderate shifts, and LRs greater than 10 or less than 0.1 would generate large and often conclusive shifts from pre-test to post-test probability (13). We also determined the diagnostic utility of combining the maneuvers.

It’s been known that the precision of the estimates of diagnostic accuracy varies as a function of both the point estimate itself and the sample size. Hence, to enhance the clinical usefulness of information on estimates, the corresponding 95% confidence intervals were calculated for all outcome measures according to the efficient-score method (corrected for continuity) described by Robert Newcombe (14). As LRs of 1 are diagnostically indeterminate, a clinical test is considered useful only if the 95% confidence interval (CI) about its LRs excludes 1. McNemar \(\chi^2\) results were used to evaluate if one maneuver identified significantly more subjects with CTS than the other. Measurement of agreement was evaluated by using Cohen’s Kappa.

**Results**

The mean age of patients was 48.9 years (range, 23-72 years) and the mean duration of symptoms was 29.7 months (range, 1-140 months). The mean age of subjects in the control group was 47.3 years (range, 19-80 years). 58 of 87 patients had electrodiagnostic evidence of CTS. Hence, the pre-test probability of CTS in our study population was 67%. The CTS-RM was positive in 60% of patients. The basic maneuver was effective in 85% of patients, whereas stretching of digits III and IV was also required to induce improvement in the remaining 15%. The Flick sign was positive in 53% of patients (Table 1).

The sensitivity estimates of the CTS-RM and Flick sign were 81% and 69% and the specificity estimates were 86% and 79% respectively in our population (Table 2). Combining a positive CTS-RM and Flick sign improved the specificity to 93% at the expense of the sensitivity (66%). When evaluating the subjects with CTS, the CTS-RM detected significantly more subjects compared with the Flick sign (McNemar \(\chi^2 = 4.5\), with \(p=0.03\)). We found a moderate agreement between them (Kappa=0.50 CI, 0.26-0.75). The positive LRs of the CTS-RM and Flick sign were 3.3 and 5.9 and the negative LRs were 0.39 and 0.22 respectively. Combining a positive CTS-RM and Flick sign had the PLR of 9.5 and the NLR of 0.37. With increasing electrodiagnostic severity, the sensitivity of the two maneuvers showed a consistently increasing trend.

**Discussion**

Our study shows that under the conditions where the pre-test probability of CTS is intermediate or high, the accuracy of Flick sign is low in the diagnosis of CTS. The CTS-RM alone is helpful in confirming the diagnosis in female patients with typical symptoms and combination with the Flick sign further improves its prediction.

The majority of studies on the accuracy of clinical tests in the diagnosis of CTS were based on provoking symptoms

<table>
<thead>
<tr>
<th>Table 1. Results of the clinical evaluation according to disease severity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Results of evaluation</strong></td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Flick maneuver</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Not observed</td>
</tr>
<tr>
<td>CTS-RM</td>
</tr>
<tr>
<td>Symptoms abolished</td>
</tr>
<tr>
<td>Symptoms improved</td>
</tr>
<tr>
<td>Symptoms worsened</td>
</tr>
<tr>
<td>No change</td>
</tr>
</tbody>
</table>
Although widely used, sensitivity and specificity have some deficiencies in clinical use. They are usually not helpful to clinicians trying to revise the probability of disease at the individual level since they are population measures summarising the characteristics of a test over a population (19). However, likelihood ratios are independent of disease prevalence and can be used directly at the individual patient level allowing clinicians to quantitate the probability of disease at the individual level since they are population measures summarising the characteristics of a test over a population (19).

Table 2. Validity of clinical maneuvers with corresponding 95% confidence intervals (CI)

| Maneuver (+/-) | CTS (+)* (n=58) | CTS (-)* (n=29) | Sensitivity %, CI | Specificity %, CI | PLR CI | NLR CI | Posterior probability % M+ % M- |
|---------------|-----------------|-----------------|-----------------|-----------------|-------|-------|-----------------|----------------|
| Flick sign    | 40/18           | 6/23            | 69, 55-80       | 79, 60-91       | 3.3, 1.6-6.9 | 0.39 | 0.26-0.58      | 87, 44         |
| CTS-RM        | 47/11           | 4/25            | 81, 68-89       | 86, 67-95       | 5.9, 2.3-14.7 | 0.22 | 0.13-0.38      | 92, 31         |
| CTS-RM and    | 38/20           | 2/27            | 66, 52-77       | 93, 76-99       | 9.5, 2.5-36.7 | 0.37 | 0.26-0.53      | 95, 43         |
| Flick sign    |                 |                 |                 |                 |       |       |                 |               |

* Maneuver (+/-)

M+/M- = posterior probability of a positive/negative maneuver

Prior probability is 67%
According to Jaeschke et al. (13), when patients with a certain disorder all have severe disease, sensitivity for a test increases; if patients are mildly affected sensitivity decreases and LRs move toward a value of one. In accordance with this, we found a substantial and consistent increase in the sensitivity of maneuvers (especially CTS-RM) with increasing electrodiagnostic severity (Figure 1). Similarly, Hansen et al. (4) found an increasing trend in the sensitivity of Flick sign. In contradiction with these results, Mondelli et al. (18) found a decreasing trend in the sensitivity of Flick sign in advanced clinical and electrophysiological stages of disease. The discrepancy in these results can probably be explained by differences in study populations, study design, definition of CTS and disease severity. For example, in our study the disease severity was determined by the DML alone, while the other authors (2, 18) took both distal motor latency and sensory conduction velocity into account as well as clinical severity.

Several factors limit the generalizability of our results, including sample size, lack of a challenging control group, disease severity, reliability of the tests, and recruitment source. Our relatively small sample size resulted in wide 95% CIs of our point-estimates. Therefore further validation of these results is needed in larger patient populations. The control group consisted of patients with typical symptoms for CTS who had normal electrophysiological findings. If it included the patient with symptomatic hand pathology other than CTS, then the estimate of specificity would have been lower with resultant changes in the point estimates of likelihood ratios. While some suggest use of a healthy, asymptomatic control group to assess true sensitivity and specificity, others emphasize use of symptomatic control group relevant with the disease in question.

Currently electrodiagnostic studies are the only truly objective ancillary test available for diagnosis of CTS. However, they are still not considered as an ideal gold standard. Since we used a set of electrodiagnostic criteria for diagnosis, the possibility that some of the patients in our control group might have had CTS remains unanswered. If so far a that is the case, our estimates of LRs could be incorrectly affected by this spectrum bias. Assessment of disease severity according to an arbitrary electrophysiological scale was another potential source of bias which might have influenced our results. It should be noted that there was an over-representation of mild disease (69%) in our study sample reflecting a spectrum bias. Therefore the point estimates of sensitivity obtained in our study actually would have been higher. For example, in neurology practice where study population would have more pronounced symptoms and more advanced disease the CTS-RM and the Flick sign would have higher estimates of sensitivity.

The reliability of clinical tests is a major concern in determining diagnostic accuracy. The Flick sign has usually been considered as a simple, operator-independent and reproducible test. It has been reported to have an excellent reliability (90%) in a recent study by Wainner et al. (8). On the other hand, the CTS-RM may be considered as an operator-dependent clinical test. Although we did not observe any worsening by CTS-RM in our patients, the question of its test-retest reliability still remains unanswered and needs to be determined.

### Table 3. Results of the diagnostic accuracy for the Flick sign in different study populations

<table>
<thead>
<tr>
<th>Study</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PLR*</th>
<th>NLR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pryse Phillips⁵</td>
<td>0.93</td>
<td>0.96</td>
<td>21.4</td>
<td>0.10</td>
</tr>
<tr>
<td>Roquer and Herraiz⁶</td>
<td>0.36</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kendall et al.⁷</td>
<td>0.25</td>
<td>0.61</td>
<td>0.6</td>
<td>1.23</td>
</tr>
<tr>
<td>De Krom et al.³</td>
<td>0.50</td>
<td>0.61</td>
<td>1.3</td>
<td>0.82</td>
</tr>
<tr>
<td>Hansen et al.⁴</td>
<td>0.37</td>
<td>0.74</td>
<td>1.4</td>
<td>0.85</td>
</tr>
<tr>
<td>Wainner et al.⁸</td>
<td>0.81</td>
<td>0.57</td>
<td>1.9</td>
<td>0.34</td>
</tr>
<tr>
<td>Our results</td>
<td>0.69</td>
<td>0.79</td>
<td>3.3</td>
<td>0.39</td>
</tr>
</tbody>
</table>

PLR: Positive likelihood ratio
NLR: Negative likelihood ratio
* Calculated from the clinical data of those studies

![Figure 1. Sensitivity of the CTS-RM and Flick sign with prolonging motor latency values](image-url)
The predictive value of diagnostic tests is dependent on the prevalence of disease in the population being tested (pre-test probability). They are mostly valuable as complementary information to clinical assessment, particularly when pre-test probability of a disease is intermediate (i.e. between 20% and 80%) (20). According to Jaschke et al. (13), each item of history, or each finding on physical examination represents a diagnostic test that either increases or decreases the probability of a target disorder. Thus, the pre-test probability is likely to be lowest with screening tests and greatest with tests performed in referred patients (20). Since our study population was derived from symptomatic patients referred to electrodiagnostic assessment after a clinical evaluation by a physiatrist, the pre-test probability was fairly high. Therefore, our results are most applicable to female patients with severe enough symptoms to warrant such a referral in a secondary care setting such as neurology or physiatry where the pre-test probability of CTS is expected to be intermediate or high.

References