Comparison of Clinical Diagnoses Versus Computerized Test Diagnoses using the Maryland Clinical Diagnostics Diagnostic Paradigm (Expert System) for Diagnosing Chronic Pain in the Neck, Back and Limbs

Abstract

Past research found that 40%-71% of chronic pain patients had overlooked diagnoses, underscoring the need for more accurate diagnostic methodology. In this study, 28 diagnoses were made by the senior author during the initial evaluation of 5 chronic pain patients. Of these diagnoses, 27 diagnoses made by the senior author were also made by the computer scored and interpreted Maryland Clinical Diagnostics (MCD) “Diagnostic Paradigm.” The MCD Diagnostic Paradigm matched the senior author’s evaluation 96.37 % of the time, and the Diagnostic Paradigm missed 1 diagnoses made by the senior author, for a 3.63% “missed diagnosis” rate, based on the initial clinic evaluation.

Overall there were 4 “Test Only Positives,” i.e. diagnoses made by the Diagnostic Paradigm, but not made by the senior author. Therefore, these patients received the benefit of additional diagnostics studies, which helped the chiropractor obtain an additional diagnoses 16% of the time, which he otherwise might have overlooked, as the results of the “Test Only Positive” diagnosis.

The Pain Validity Test could predict which of the patients would have moderate or severe abnormalities on objective medical tests with 100% accuracy. The Diagnostic Paradigm predicted the specific abnormality on objective medical testing with 98% accuracy.

Keywords: Expert System; Chronic Pain; Misdiagnosis; Validating Pain; Accurate Diagnosis

Introduction

Past research reports from Mensana Clinic indicate that 40% to 67% of chronic pain patients involved in litigation are misdiagnosed [1,2]. When evaluating just the diagnosis of complex regional pain syndrome, type I, (CRPS I), formerly called reflex sympathetic dystrophy (RSD), Hendler found that 71% of the patients who were told they had only CRPS I, actually had nerve entrapment syndromes, and 26% had a combination of both nerve entrapment syndrome and CRPS I [3]. Therefore, 97% of patients diagnosed by other physicians as having CRPS I, were misdiagnosed, or only partially diagnosed. In specialized diagnostic situations, the overlooked diagnosis rate for people who survived lightning strikes was 93 %, and for people who survived electrical injury, the rate was 98% [4]. These errors in diagnoses are costly to the patient and the medical system alike, since they prolong or result in inappropriate treatment.

Psychiatric problems arise as the result of chronic pain [5,6]. Hendler reported that 77% of patients seen at Mensana Clinic had coexisting depression and chronic pain, but when questioned about pre-existing depression, 89% of the patients had never had significant depression before the onset of their pain [6]. The presence of psychiatric problems, even though a normal response to chronic pain, biases many physicians, which results in a less extensive evaluation [7,8]. This physician bias is often compounded by factors such as litigation, an additional negative bias against women with pain complaints, and a bias against men by female physicians [9-13]. These biases also influence the length of time, and the extent of an evaluation. Some physicians spend less than 15 minutes with a patient [14], while other “high volume” physicians have reduced by 30% the amount of time they spend with their patients [15]. Since some physicians have reduced the length of time spent with patients, an automated history is a desirable efficiency, and may improve the accuracy of diagnosis and treatment, since a comprehensive questionnaire can ask questions overlooked by the time conscious physician.

Training physicians and transfer of knowledge is a cumbersome process. Hansen and his coauthors divide knowledge transfer into computer based systems, which they call a “codification strategy,” and direct person to person contact which they call the “personalization strategy” [16]. The use of the “codification strategy” “…allows many people to search for and retrieve
Comparison of Clinical Diagnoses Versus Computerized Test Diagnoses using the Maryland Clinical Diagnostics Diagnostic Paradigm (Expert System) for Diagnosing Chronic Pain in the Neck, Back and Limbs


In order to see if the diagnostic methods could be duplicated, Maryland Clinical Diagnostics (MCD) developed the “Diagnostic Paradigm” described in the following report. The following research report tests the reliability, specificity and sensitivity of these methods.

Subjects
All subjects who received the Maryland Clinical Diagnostics (MCD) tests were patients at a private chiropractor’s office. Of the 5 patients included in this study, all were involved in active litigation. The patients included in this study represent new evaluations seen between November 2011, through March 2012. The average age of the patients was 41.2 years. Three subjects were females, and two were males.

Methods
For this study, an evaluation at the senior author’s office consisted of the administration of the MCD Diagnostic Paradigm, and The MCD Pain Validity Test, and immediately afterwards, a one hour clinical evaluation with the chiropractor, who was the senior author. The diagnoses from the initial evaluation of senior author were compared to the diagnoses generated by the MCD Diagnostic Paradigm, which had been administered just prior to the clinical evaluation, and scored within 5 minutes of completion.

The diagnoses were considered a match only if the type of pathology and the spinal level diagnosed by the senior author corresponded with the type of pathology and spinal level generated by the diagnostic paradigm.

The MCD Diagnostic Paradigm was designed to detect 60 diagnoses and 44 differential diagnosis, for a total of 104 diagnoses commonly seen in post traumatic auto accident or workers compensation injuries. A list of the possible diagnoses, and differential diagnoses is shown in appendix A. These diagnoses were selected for programming for computer analysis and scoring, since these were the most frequent ones seen in post-traumatic injuries.

Sources of errors
Two patients had “Patient Input Error.” A “Patient Input Error” is defined as an error in completing the questionnaire. Two of the patients initially neglected to complete sections of the questionnaire dealing with pain the shoulder, because the questions about shoulder pain were included under arm pain. When this oversight was shared with the patients, they then were able to complete the questionnaire accurately.

Another potential source of errors was a Patient Communication Error: Included in this report were two patients who had marked questions for areas of the body that they had not discussed with the senior author during the initial evaluation. When questioned after the clinical evaluation and MCD Diagnostic Paradigm results were both available, the patients said “I forgot to tell the doctor” and “I was too anxious or scared during the clinical interview so I neglected to mention the problem “and “I didn’t think it was important.” This error was considered a “Patient Communication Error” This error demonstrated the value of a self

...
administered questionnaire, which reduces the chance of missed or incomplete history, and therefore missed diagnoses, through patient communication errors.

Results

There were 28 diagnoses made by the senior author during the initial evaluation on the 5 patients included in the study. Of these diagnoses, 27 diagnoses made by the senior author were also made by the MCD Diagnostic Paradigm. Therefore, the comparison reported in this article is between initial clinical evaluations and Diagnostic Paradigm. The MCD Diagnostic Paradigm matched the senior author’s evaluation 96.37% of the time, and the Diagnostic Paradigm missed 1 diagnosis made by the senior author, for a 3.63% “missed diagnosis” rate, based on the initial clinic evaluation.

Overall there were four Test Only Positives, i.e. diagnoses made by the Diagnostic Paradigm, but not made by the senior author. Finding these four diagnoses, in addition to the 28 diagnoses made (16%), is of great benefit to the treating chiropractor, since it improves his or her level of care, and assures that the patients get a comprehensive evaluation.

The only medical laboratory studies performed in this study population were CTs, MRIs and X-rays. Every abnormality found on these tests was predicted by diagnoses generated by the Diagnostic Paradigm. In other words, the tests confirmed diagnoses made by the Diagnostic Paradigm 100% of the time. Other diagnoses made by the Diagnostic Paradigm were not confirmed, since the objective testing needed to confirm the diagnoses was never performed, i.e. EMG/Nerve conduction testing to confirm radiculopathy, provocative discograms to confirm damaged discs, etc. With more extensive physiological testing, provocative discogram, facet blocks, EMG, etc., a clinician should expect to confirm many more diagnoses, in addition to those confirmed by CT, MRI and X-ray, which are anatomical tests.

In all 5 patients, the Pain Validity Test predicted that the patients would have moderate or severe abnormalities on objective medical testing (Objective Pain Patients). Indeed, all 5 patients had at least one moderate or severe abnormality on objective medical testing. This 100% accuracy is compatible with previously published reports indicating a 94%- 95% accuracy level in a larger number of patients [7,11].

Discussion

A number of deficits exist with expert systems. In the absurd extreme, if the computerized expert system lists all the possible diagnoses, there is 100% sensitivity, but the specificity is very low. Conversely, if the specificity is tightened to such a degree that the computerized expert system always gets a specific diagnosis, but misses other associated diagnoses, the sensitivity of the system is reduced to a level of inaccuracy that approaches or exceeds the false positive rate, i.e. the MRI says there is pathology in a disc, and lumbar spine. However, in 98 patients, the MRI has a 29% false positive rate, i.e. the MRI says there is pathology in a disc, in patients who are asymptomatic [32], and a 69%–77% there is a false negative rate, i.e. the MRI says there is no abnormality, in patients who are symptomatic, and have positive provocative findings.

The accuracy of any computer scored and interpreted expert systems is a major issue. Those expert systems that seem to have the best results are the ones that focus on a narrow and highly specialized area of medicine. One questionnaire, consists of 60 questions, to cover 32 rheumatologic diseases, for 358 patients [20]. The correlation rate was 74.4%, and an error rate of 25.6%, with the 44% of the errors attributed to “information deficits of the computer using standardized questions,” [20], but in a later version “RHEUMA” was studied prospectively in 51 outpatients, and achieved a 90% correlation with clinical experts [21]. Several groups have approached the diagnosis of jaundice. ICTERUS produced a 70% accuracy rate [22], while ‘jaundice’ also had a 70% overall accuracy rate [23]. An expert system for vertigo was reported, and it generated and accuracy rate of 65%, [24]. This later was reported as OtoNeurological Expert (ONE), which generated the exact same results reported in the earlier article [25]. There was a 76% agreement for diagnosis of depression, between an expert system and a clinician [26]. When a Computer Assisted Diagnostic Interview (CADI) was used to diagnosis a broad range of psychiatric disorders, there was an 85.7% agreement level with three clinicians [27]. In a review of twenty charts by a computerized analysis of treatment for hypertension, using HyperCritic, a panel of 18 family practitioners felt the treatment suggested by the computer system was erroneous or possibly erroneous 16% of the time [28]. The panel accepted HyperCritic critiques equally as beneficial as critiques from 8 human reviewers [28]. Others have developed a “to do” list to remind and alert treating physicians about tests they should order, based on input into electronic patient records [29]. In the narrow area of managing lipid levels, there was a 93% agreement between management advice given by the expert system, and the specialist, after interpretation of laboratory and clinical data [30]. However, physicians have a 65% level of accepting comments from expert systems regarding diagnosis of a patient, and are resistant to comments about prescriptions for patients, with only a 35% acceptance level [31]. Therefore, there may be more resistance from untrained physician to the use of the diagnostic studies recommended by the Report of the MCD Diagnostic Paradigm, than there might be to accepting the diagnoses generated by the report of the MCD Diagnostic Paradigm. This premise needs to be tested in future research.

Many of the recommended diagnostic studies from the Treatment Algorithm are not commonly used in community medical centers, but have been used for years by major teaching hospitals in the United States. A classic example of this is the wide spread use of the MRI for detecting disc damage in the cervical and lumbar spine. However, in 98 patients, the MRI has a 29% false positive rate, i.e. the MRI says there is pathology in a disc, in patients who are asymptomatic [32], and a 69%–77% there is a false negative rate, i.e. the MRI says there is no abnormality, in patients who are symptomatic, and have positive provocative findings.
discogram [33-35]. The value of the provocative discogram is clearly demonstrated by the groundbreaking work by Bogduk, who clearly demonstrated pain fibers in the posterior portion of the annulus of an inter-vertebral disc, which can be damaged, and produce pain, without any anatomical distortion of the disc [36]. He terms this condition “internal disc disruption” [37]. Central to understanding the value of the provocative discogram the concept that pain is a physiological condition, not an anatomical event. While the use of an MRI can detect only anatomical distortions, the use of the provocative discogram, which is a physiological test, is more reliable for diagnosing chronic pain. The same rationale applies to the use of other physiological tests, used to make diagnoses in chronic pain patients, such as root blocks, nerve blocks, facet blocks, bone scans, gallium scans, Indium 111 scans, neurometer studies, somatosensory evoked potentials, and flexion-extension X-rays with obliques. With more extensive physiological testing, (provocative discogram, facet blocks, EMG, etc.), a clinician should expect to confirm many more diagnoses, in addition to those confirmed by CT, MRI and X-ray, which are anatomical tests. This is why the majority of the recommended tests in the Treatment Algorithm are physiological ones (Tables 1-3).

Table 1: Comparison of Chiropractor to Diagnostic Paradigm Diagnoses.

<table>
<thead>
<tr>
<th>Patient #1</th>
<th>Diagnosis by Chiropractor</th>
<th>Diagnosis by Diagnostic Paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Lumbar facet damage L2-4</td>
<td>Facet syndrome L2-4</td>
</tr>
<tr>
<td>2</td>
<td>Lumbar disc damage L3-4</td>
<td>L3-4 disrupted disc</td>
</tr>
<tr>
<td>3</td>
<td>Subluxation C3 on C4</td>
<td>C3-4 unstable spinal segment</td>
</tr>
<tr>
<td>4</td>
<td>--</td>
<td>Occult fracture navicular bone</td>
</tr>
<tr>
<td>Patient #2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Herniated disc L4-L5</td>
<td>Herniated disc L4-L5</td>
</tr>
<tr>
<td>2</td>
<td>L4-L5 Radiculopathy</td>
<td>L4-L5 Radiculopathy</td>
</tr>
<tr>
<td>3</td>
<td>Lumbar Strain</td>
<td>---</td>
</tr>
<tr>
<td>4</td>
<td>--</td>
<td>L3-S1 Lumbar facet syndrome</td>
</tr>
</tbody>
</table>

The purpose of an “expert system” is to improve the level of the reliability and accuracy of diagnosis, and enhance medical care. In our research, we had essentially the same results with the www.MarylandClinicalDiagnostics.com Diagnostic Paradigm and Treatment Algorithm as published in an earlier article, which reported that the test generated diagnoses which had a 96% correlation with diagnoses of Johns Hopkins Hospital staff members [38]. Likewise, the fact that in this research, the Pain Validity Test could predict with 100% accuracy who would have moderate or severe abnormalities on objective medical testing corresponds with the published reported accuracy level of 95% [7,11,39,40]. Therefore, we conclude that the tests offered by www.MarylandClinicalDiagnostics.com are a most useful adjunct to evaluating and treating a patient in a chiropractic setting.

Table 2: Predictive Results of Pain Validity Test, where “objective pain patients” are predicted to have abnormal medical test results with 95% accuracy [39].

<table>
<thead>
<tr>
<th>Result of Pain Validity Test</th>
<th>Medical Test Abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient #1</td>
<td></td>
</tr>
<tr>
<td>Objective Pain Patient</td>
<td>X-ray-uncovertebral osteophytes, C5-6</td>
</tr>
<tr>
<td>LF</td>
<td>Subluxation C3 on C4</td>
</tr>
<tr>
<td>MRI-L2-3 facet degenerative changes</td>
<td>L3-4 facet degenerative changes</td>
</tr>
<tr>
<td>L3-4 hip within normal limits</td>
<td>X-ray-shoulder within normal limits</td>
</tr>
<tr>
<td>X-ray-shoulder within normal limits</td>
<td>X-ray-shoulder within normal limits</td>
</tr>
</tbody>
</table>

Table 3: Confirmation of Diagnostic Paradigm Diagnosis with Objective Testing.

<table>
<thead>
<tr>
<th>Result of Diagnostic Paradigm</th>
<th>Medical Test Abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient #1</td>
<td></td>
</tr>
<tr>
<td>C5-6 radiculopathy, facet syndrome</td>
<td>X-ray-uncovertebral osteophytes, C5-6</td>
</tr>
<tr>
<td>LF</td>
<td>Subluxation C3 on C4</td>
</tr>
<tr>
<td>MRI-L2-3 facet degenerative changes</td>
<td>L3-4 facet degenerative changes</td>
</tr>
<tr>
<td>L3-4 hip within normal limits</td>
<td>X-ray-shoulder within normal limits</td>
</tr>
<tr>
<td>X-ray-shoulder within normal limits</td>
<td>X-ray-shoulder within normal limits</td>
</tr>
</tbody>
</table>

References

addition of MRI, nerve blocks, 3-D CT, and qualitative flow meter. Psychosomatics 37(6): 509-517.


