Research Article

Journal of Anesthesia & Pain Medicine

Diagnoses from an On-Line Expert System for Chronic Pain Confirmed by Intra-Operative Findings

Alessandro Landi*[1], Reginald Davis[2], Nelson Hendler[3] and Al-Rahim Abbasali Tailor[4]

[1]*Department of Neurology and Psychiatry, Division of Neurosurgery, University of Rome.*

[2]*Assistant Professor of Neurosurgery, Johns Hopkins University School of Medicine and Chief of Neurosurgery, Greater Baltimore Medical Center.*

[3]*Former Assistant Professor of Neurosurgery, Johns Hopkins University School of Medicine.*

[4]*Department of Neurological Surgery, Wexner Medical Center, Ohio State University.*

***Corresponding author**

Nelson Hendler, Former Assistant Professor of Neurosurgery, Johns Hopkins University School of Medicine, University in Baltimore, Maryland, US, Tel: 443-277-0306; E-mail: DocNelse@aol.com

Submitted: 03 Aug 2016; **Accepted**: 15 Aug 2016; **Published**: 19 Aug 2016

Abstract

A number of researchers from Johns Hopkins Hospital report that 40%-80% of chronic pain patients are misdiagnosed. Previous reports indicate an on-line questionnaire, The Diagnostic Paradigm and Treatment Algorithm, provides diagnoses with a 96.3% correlation with diagnoses of Johns Hopkins Hospital staff members, in patients with chronic back, neck or limb pain. This research was undertaken to determine if diagnoses generated by the Diagnostic Paradigm and Treatment Algorithm could be confirmed by irrefutable indications of pathology, i.e. intra-operative findings. Prior to surgery, the Diagnostic Paradigm and Treatment Algorithm was administered to ten patients. The Diagnostic Paradigm predicted 61/61 (100%) diagnoses which were confirmed intra-operatively. The Diagnostic Paradigm had 71 false positive diagnoses, but these were part of the differential diagnoses of the correct diagnoses. These differential diagnoses were refined by medical testing.

Keywords: Intra-operative findings, Expert system for chronic pain, Back pain, Neck pain, Predicting organic pathology, Misdiagnosis of chronic pain, Arm pain, Leg pain, On-line diagnoses.

Introduction

A surgeon is often faced with multi-factorial challenges when evaluating a patient with chronic pain problems. Chronic pain is defined as a constant pain lasting 6 months or longer and often causes psychological problems, which interferes with accurate medical assessment [1]. X-ray studies, electromyelograms (EMG), nerve conduction velocity studies and EMG may document an organic basis of chronic back pain, but some pain problems cannot be identified by objective tests, since there are many false negative and false positive results on "objective" medical testing [2-5].

Physician prejudice against woman patients can result in a significantly less extensive evaluation of their complaints of back pain [6]. Litigation may influence symptoms and the type of litigation may influence outcomes [7,8]. For that reason, there is a need to differentiate between "organic" (valid) and "functional" (negative physical and laboratory examination) back, neck and limb pain, before undertaking an extensive medical evaluation, prescribing narcotic medication, or performing surgery [9]. Patients often have difficulty describing the location of their complaint of pain. The combination of these factors leads to a misdiagnosis rate of 40%-80% in chronic pain patients while for specific diagnoses; this rate may reach as high as 97%, as is the case in the overuse of the term fibromyalgia [10-14]. In order to improve, surgery was needed on 50%-80% of the misdiagnosed patients [10-13]. The patient improvement, documented in published outcome studies, establishes the benefit of these surgeries [12,13,15].

The Diagnostic Paradigm and Treatment Algorithm is a 72 question questionnaire, which asked about patient symptoms, with 2008 possible multiple choice answers, about conditions which improve or worsen symptoms. It is available in English and Spanish over the Internet at www.MarylandClinicalDiagnostics.com. It was designed to evaluate 104 of the most common post-traumatic injuries, resulting in chronic pain. Based on the diagnoses and differential diagnoses, a Treatment Algorithm is generated. Results are emailed back to physicians in 5 minutes after completion of the test. The diagnoses from the Diagnostic Paradigm have a 96.3% correlation with diagnoses of Johns Hopkins Hospital staff members [16].

The present study is designed to investigate the usefulness of the Diagnostic Paradigm and Treatment Algorithm for predicting the presence or absence of intra-operatively documented organic pathological conditions in patients with chronic back, neck and/or limb pain. Rather than compare "expert system" diagnosis to clinical diagnosis, or abnormal medical tests, this research is an attempt to determine if a properly designed "expert system" questionnaire and Bayesian analysis of the answers gave diagnoses which could identify the actual intra-operative findings, using predictive analytic techniques. In this fashion, diagnoses from the "expert system" were confirmed by intra-operative findings, which is a much more powerful validation of the accuracy of diagnoses of the expert system than previous comparisons.

The Diagnostic Paradigm and Treatment Algorithm

Clinical symptoms (what the patient reports to the doctor) in medicine very often are the result of a convergence of medical conditions. This means a single symptom, such as pain and numbness in the last two fingers of a hand, may have multiple origins, such as a C6-7 radiculopathy, ulnar nerve entrapment, or thoracic outlet syndrome. Given a single clinical symptom, all three diagnoses need to be considered, and are rank ordered from most likely to least likely, as part of the diagnosis and differential diagnosis. The cause of a single symptom can be rank ordered based on clinical experience.

The Diagnostic Paradigm was constructed in this fashion, so the most likely cause for the symptom was considered the working diagnosis, and in declining order, the other causes for the symptom were considered. This type of thinking is called Bayesian logic, and is the basis of the scoring and interpretation of the Diagnostic Paradigm. The rank order of causes for a symptom was based on a retrospective chart review of 10,000 charts over a 17 year period of time. In this review, the origins of a single symptom were tabulated, and assigned a weight, in terms of percentage of likelihood. This leads to the diagnoses and differential diagnoses generated by the Diagnostic Paradigm, followed by the percentage likelihood of a cause for the symptom following each diagnosis or differential diagnosis. This format has two consequences. All causes for a single symptom are included, so no diagnosis is ever missed, and the diagnoses and differential diagnoses are intentionally overly inclusive.

In a prospective study to determine the accuracy of retrospectively derived diagnoses, the diagnoses from the Diagnostic Paradigm were found to have a 96.3% correlation with diagnoses of Johns Hopkins Hospital staff members [16]. The output from the Diagnostic Paradigm lists diagnoses and differential diagnoses, rank ordered from most likely to least likely, and assigned a "percentage of likelihood" of the diagnosis, based on the number of symptoms a patient has divided by the total number of symptoms a physicians would expect a patient to report for a certain disorder. Diagnoses are clustered into 18 groups of similar diagnoses, which are further differentiated by conducting the medical testing recommended in the Treatment Algorithm. As an example, if a patient had symptoms compatible with L4-L5 radiculopathy, the symptoms could be caused by a herniated disc at L4-L5 compressing the nerve root, or by neural foraminal stenosis of L4-L5. The two diagnoses would be clustered as part of a group with similar clinical symptoms, which would require the same set of diagnostics tests to differentiate the cause, i.e. 3D-CT, MRI, facet block, root block, and provocative discogram. The ultimate confirmation of the cause of the problem would be intra-operative findings.

Medical tests very often have false positive and false negative results, which confounds the decision to perform surgery. However, the real concern is not whether a patient has an abnormal test. This real issue is the presence or absence of intra-operative pathology, i.e. was the surgery warranted. Therefore, a verbal test which could predict intra-operative findings would be a valuable screening tool for non-medical professionals, such as psychologists, insurance carriers, or attorneys. It would help them decide if extensive medical tests should be ordered, and the Treatment Algorithm portion of the test provides the surgeon with suggestions for interventional testing which is employed at Johns Hopkins Hospital, in addition to the common tests, such as the MRI of discs, in which the false negative rate could reach as high as a 78% [17].

Methods

Patients

Ten patient charts were reviewed. These consecutively chosen patients had been selected by the senior author for spinal surgery, based on his clinical assessment and laboratory studies. Prior to surgery, each of these ten patients were administered the Diagnostic Paradigm and Treatment Algorithm, from www.marylandclinicaldiagnostics.com.

Results

Analysis of Intra-Operative Findings

The operative note from each patient who received surgery was blindly reviewed by a researcher who did not perform the surgery. Findings were considered normal if pathology reports and intra-operative notes indicated no pathology. Findings were considered mild if pathology reports and/or intra-operative notes found "mild scaring of a nerve root," or "mild scarring of a peripheral nerve" or "mild neural foraminal stenosis," or "mild compression of a vessel." Likewise, if the reports or note contained the word moderate or severe, then the pathology was considered moderate or severe. Various surgeries were performed, including fusions, laminectomies, discectomies, removal of arachnoiditis, foraminotomy, and others.

Diagnostic Paradigm Diagnoses		Intra-operative Findings
L5-S1 Herniated or Disputed Disc Score=1.000000		
L5-S1 Radiculopathy Score=0.937500	1	hard bony stenosis and a soft stenosis
L4-L5 Radiculopathy Score=0.770833	1	ligamentum flavum hypertrophy.

L4-L5 Herniated or Disrupted Disc Score=0.750000		decompressed the dural sac
L3-L4 Radiculopathy Score=0.650000	1	
L3-L4 Herniated or Disrupted Disc Score=0.500000		
Lumbar Facet Syndrome L3-S1 Score=0.500000	1	
Neural Foraminal Stenosis L1-S1 Score=0.500000	1	opening of foramen of L3,L4,L5 and S1 roots.
Lumbar Facet Syndrome L3-S1 Score=0.500000		
Neural Foraminal Stenosis L1-S1 Score=0.500000	1	
Retrolysthesis L1-S1 - Score=0.500000		
L5-S1 Radiculopathy Score=0.946429	1	L5 and S1 roots are decompressed bilaterally
L5-S1 Herniated or Disputed Disc Score=0.937500		
Spinal Stenosis of the Lumbar Spine Score=0.750000	1	decompressed the dural sac
L4-L5 Radiculopathy Score=0.730769	1	L4-L5 with an extensive scar tissue
L4-L5 Herniated or Disrupted Disc Score=0.687500		medial arthrectomy of L4-L5-S1
Arachnoiditis L5-S1 Score=0.687500	1	
L3-L4 Radiculopathy Score=0.678571		
Lumbar Facet Syndrome L3-S1 Score=0.625000	1	hyperthrophy of the facet joints
Neural Foraminal Stenosis L1-S1 Score=0.625000	1	opening of the right foramen of L5 and S1
Unstable Spinal Segment at L3-L4 Score=0.522727		
Spondylolysis/Spondylolythesis/Antero-Lysthesis		
Unstable Lumbar Spinal Segment Score=0.522727	1	
Unstable Spinal Segment at L4-L5 Score=0.500000	1	Stabilization with pedicular screws at L4- S1
Retrolysthesis L1-S1 - Score=0.500000		
Unstable Spinal Segment at L5-S1 Score=0.500000	1	Posterolateral arthrodesis
L5-S1 Radiculopathy Score=0.946429	1	ligamentum flavum hypertrophy
L5-S1 Herniated or Disputed Disc Score=0.916667		Neurolysis of the L5 right root adherence.
Spinal Stenosis of the Lumbar Spine Score=0.750000	1	
L4-L5 Radiculopathy Score=0.704545		Decompressive laminectomy at L3-L4 and L4-L5.
Arachnoiditis L5-S1 Score=0.687500	1	
L4-L5 Herniated or Disrupted Disc Score=0.666667	1	Neurolysis of the dural adherence
Unstable Spinal Segment at L3-L4 Score=0.550000		lumbar disc herniation at L4-L5.

Spondylolysis/Spondylolythesis/Antero-Lysthesis		
s/Unstable Lumbar Spinal Segment Score=0.550000		
L3-L4 Radiculopathy Score=0.550000		
Lumbar Facet Syndrome L3-S1 Score=0.541667		
Unstable Spinal Segment at L4-L5 Score=0.522727		
Unstable Spinal Segment at L5-S1 Score=0.522727		
Retrolysthesis L1-S1 - Score=0.500000		
L5-S1 Radiculopathy Score=1.000000	1	
L5-S1 Herniated or Disputed Disc Score=0.900000	1	Neurolysis of the L5 right root adherence
Spinal Stenosis of the Lumbar Spine Score=0.750000	1	Microdiscectomy L5-S1.
L4-L5 Radiculopathy Score=0.727273		Neurolysis of the dural adherence and scar tissue.
Arachnoiditis L5-S1 - Score=0.678571	1	
L4-L5 Herniated or Disrupted Disc Score=0.650000		
Lumbar Facet Syndrome L3-S1 Score=0.625000		
Neural Foraminal Stenosis L1-S1 Score=0.625000	1	Severe foraminal stenosis l5-s1
Unstable Spinal Segment at L3-L4 Score=0.562500		
Spondylolysis/Spondylolythesis/Antero-Lysthesis		
Unstable Lumbar Spinal Segment Score=0.562500		
Unstable Spinal Segment at L4-L5 Score=0.538462	1	Severe spinal l4-l5-s1 instability
Unstable Spinal Segment at L5-S1 Score=0.538462	1	Stabilization with screws and rods L4-L5-S1.
L3-L4 Radiculopathy Score=0.500000		
Retrolysthesis L1-S1 - Score=0.500000		
L5-S1 Herniated or Disputed Disc Score=1.000000		Severe spinal l4-l5 instability
L5-S1 Radiculopathy Score=0.906250	1	Severe right lumbar disc herniation l4-l5.
L4-L5 Radiculopathy Score=0.781250		Stabilization with screws and rods L4-L5.
L4-L5 Herniated or Disrupted Disc Score=0.750000	1	Decompressive laminectomy of L4-L5 with HARD stenosis.
Unstable Spinal Segment at L3-L4 Score=0.541667		Foraminotomy of L5 with SEVERE FORAMINAL STENOSIS.
Lumbar Facet Syndrome L3-S1 Score=0.541667		Microdiscectomy L4-L5 with removal
Spondylolysis/Spondylolythesis/Antero-Lysthesis/		of a SEVERE lumbar disc herniation.

Unstable Lumbar Spinal Segment Score=0.541667	1	
Unstable Spinal Segment at L4-L5 Score=0.500000	1	
L3-L4 Herniated or Disrupted Disc Score=0.500000		
L3-L4 Radiculopathy Score=0.500000		
Retrolysthesis L1-S1 - Score=0.500000		
Neural Foraminal Stenosis L1-S1 Score=0.500000	1	
Unstable Spinal Segment at L5-S1 Score=0.500000		
L5-S1 Radiculopathy Score=0.953125	1	Decompressive laminectomy of L3-L4-L5
L5-S1 Herniated or Disputed Disc Score=0.850000		SEVERE SPINAL STENOSIS L3-L4-L5
L4-L5 Radiculopathy Score=0.750000	1	L3, L4 and L5 SEVERE FORAMINAL STENOSIS
L4-L5 Herniated or Disrupted Disc Score=0.600000		due to facet joint hypertrophy
Unstable Spinal Segment at L3-L4 Score=0.583333		
Spondylolysis/Spondylolythesis/Anterio-Lysthesis/		
Unstable Lumbar Spinal Segment Score=0.583333		
Unstable Spinal Segment at L4-L5 Score=0.550000		
L3-L4 Radiculopathy Score=0.550000	1	
Unstable Spinal Segment at L5-S1 Score=0.550000		
Lumbar Facet Syndrome L3-S1 Score=0.500000	1	
Retrolysthesis L1-S1 - Score=0.500000		
Unstable Spinal Segment at L5-S1 Score=0.656250	1	L3, L4 & L5 SEVERE FORAMINAL STENOSIS
Unstable Spinal Segment at L3-L4 Score=0.714286	1	facet joint hypertrophy L3. L4 L5
Unstable Spinal Segment at L4-L5 Score=0.656250	1	dural adherence.
L5-S1 Radiculopathy Score=1.000000	1	L3-L4-L5 WITH INSTABILITY
L5-S1 Herniated or Disputed Disc Score=0.850000	1	
L3-L4 Radiculopathy Score=0.833333	1	
L4-L5 Radiculopathy Score=0.750000	1	
Lumbar Facet Syndrome L3-S1 Score=0.607143	1	
L3-S1 Facet Break - Score=0.875000		
Lumbar Facet Syndrome L3-S1 Score=0.607143	1	
L5-S1 Radiculopathy Score=0.931818	1	SEVERE SPINAL STENOSIS L4-L5

L5-S1 Herniated or Disputed Disc Score=0.900000		Decompressive laminectomy of L4-L5
L4-L5 Radiculopathy Score=0.750000		with HARD and SOFT stenosis.
Spinal Stenosis of the Lumbar Spine Score=0.750000	1	Foraminotomy of L5
Arachnoiditis L5-S1 - Score=0.666667	1	SEVERE FORAMINAL STENOSIS
L4-L5 Herniated or Disrupted Disc Score=0.650000		due to facet joint hypertrophy
Lumbar Facet Syndrome L3-S1 Score=0.583333	1	Neurolysis of dural adherence.
Unstable Spinal Segment at L3-L4 Score=0.562500		
Unstable Spinal Segment at L4-L5 Score=0.562500		
Spondylolysis/Spondylolythesis/Anterio-Lysthesis/		
Unstable Lumbar Spinal Segment Score=0.583333		
Unstable Spinal Segment at L5-S1 Score=0.562500		
L3-L4 Radiculopathy Score=0.500000		
Retrolysthesis L1-S1 Score=0.500000		
L3-S1 Facet Break Score=1.000000		
Lumbar Facet Syndrome L3-S1 Score=0.583333	1	
Retrolysthesis L1-S1 Score=0.500000		
L5-S1 Radiculopathy Score=1.000000	1	SEVERE SPINAL STENOSIS L4-L5
L5-S1 Herniated or Disputed Disc Score=0.892857		Decompressive laminectomy of L4-L5 with
Spinal Stenosis of the Lumbar Spine Score=0.750000	1	HARD and SOFT stenosis.
L4-L5 Radiculopathy - Score=0.708333	1	Foraminotomy of L4 and L5
Arachnoiditis L5-S1 - Score=0.666667	1	SEVERE FORAMINAL STENOSIS
L4-L5 Herniated or Disrupted Disc Score=0.642857		due to facet joint hypertrophy
Lumbar Facet Syndrome L3-S1 Score=0.562500		Neurolysis of dural adherence
Unstable Spinal Segment at L3-L4 Score=0.550000		
Spondylolysis/Spondylolythesis/Anterio-Lysthesis/		
Unstable Lumbar Spinal Segment Score=0.583333		
Unstable Spinal Segment at L4-L5 Score=0.531250		
Unstable Spinal Segment at L5-S1 Score=0.531250		
L3-L4 Radiculopathy Score=0.500000		
Retrolysthesis L1-S1 - Score=0.500000		

Neural Foraminal Stenosis L1-S1 Score=0.500000	1	
L3-S1 Facet Break - Score=0.875000		
Lumbar Facet Syndrome L3-S1 Score=0.562500	1	
Neural Foraminal Stenosis L1-S1 Score=0.500000	1	
L5-S1 Radiculopathy Score=1.000000		extraforaminal disc herniation l4-l5
L5-S1 Herniated or Disupted Disc Score=0.900000		removal of medial articular mass of l4,
L4-L5 Radiculopathy Score=0.750000	1	removal of the hernia which compress and
L3-L4 Radiculopathy Score=0.666667		disloged the l4 root outside the l4 right foramen,
L4-L5 Herniated or Disrupted Disc Score=0.650000	1	stabilization with interspinous fusion device ASPEN
Unstable Spinal Segment at L3-L4 Score=0.625000		
Unstable Spinal Segment at L4-L5 Score=0.625000	1	
Spondylolysis/Spondylolythesis/Anterio-Lysthesis/		
Unstable Lumbar Spinal Segment Score=0.583333	1	
Unstable Spinal Segment at L5-S1 Score=0.625000	1	
Retrolysthesis L1-S1 - Score=0.500000		
TOTAL		

Table 1: Lists the various surgeries for which intra-operative findings were reviewed.

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 lack of accuracy of current physician diagnostic skills and no benefit accrues from the use of the computerized expert system [1-4].

After 30 years of work in this area, some authors feel only limited progress has been made in expert systems [18]. Engelbrecht feels that the quality of knowledge used to create the system, and the availability of patient data are the two main problems confronting any developer of an expert system, and advocates an electronic medical record system to correct one component of the problem [19]. Babic concurs with the value of the longitudinal collection of clinical data, and data mining to develop expert systems [20].

The accuracy of any computer scored and interpreted expert systems are a major issue. One of the major sources of error seems to be the use of Boolean logic in programming the expert system. The other problem is selecting too broad a topic of medicine, such as "internal medicine." As an example, think of the differential diagnoses associated with the symptom of "fever." Even if this is broken down into "fever below 100 F," fever between 100 and 102 F," and fever greater than 102 F," the task of determine a diagnosis for a symptom such as fever becomes daunting.

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 [21]. 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," [21]. However, a later version called "RHEUMA" was used prospectively in 51 outpatients, and achieved a 90% correlation with clinical experts [22]. Several groups have approached the diagnosis of jaundice. ICTERUS produced a 70% accuracy rate while 'Jaundice' also had a 70% overall accuracy rate [23,24]. An expert system for vertigo was reported, and it generated and accuracy rate of 65%, [25]. This later was reported as OtoNeurological Expert (ONE), which generated the exact same results reported in the earlier article [26]. There was a 76% agreement for diagnosis of depression, between an expert system and a clinician [27]. 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 [28]. In a review of twenty charts by a computerized analysis of treatment for hypertension, using Hyper Critic, 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 Hyper Critics critiques equally as beneficial as critiques from 8 human reviewers [29]. 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 [30]. 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 [31]. 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 [32]. Therefore, there may be more resistance from untrained physician to the use of the diagnostic studies recommended by the Report of the Diagnostic Paradigm, than there might be to accepting the diagnoses generated by the Report of the Diagnostic Paradigm. This premise needs to be tested in future research.

The rationale for the output of the Diagnostic Paradigm was to have a high degree of sensitivity, i.e. to be as inclusive as possible with diagnoses and differential diagnoses, and then use the recommended diagnostic studies and laboratory tests in the Treatment Algorithm to increase the specificity of the diagnoses. This led to generating a large number of false positive results, which then would require refinement using objective testing. In this fashion, the chance of missing a possible diagnosis is reduced. Moreover, 100% of the false positive results were within the same cluster of diagnostic considerations or the Diagnostic Group as the diagnosis which predicted a positive intra-operative finding. As an example, L4-5 retrolysthesis, in the absence of neural foraminal stenosis, and L3-S1 facet syndrome will have very similar clinical manifestations, which would be impossible to differentiate on the basis of symptoms alone, i.e. worse pain in the lower back when leaning backwards, and improvement with flexion, and can be differentiated only by testing recommended in the Treatment Algorithm.

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 with no complaint of back pain, the MRI has a 29% false positive rate, i.e. the MRI says there is pathology in a disc, in patients who are asymptomatic and a 69%-79% false negative rate, i.e. the MRI says there is no abnormality, in patients who are symptomatic, and have positive provocative 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, peripheral nerve blocks, bone scans, gallium scans, Indium 111 scans, neurometer studies, EMG/nerve conduction velocity studies, somatosensory evoked potentials, and flexion-extension X-rays with oblique's. This is why the majority of the recommended tests in the Treatment Algorithm are physiological ones.

Additionally, there were 61 pathological conditions found intra-operatively by the senior author on the 10 patients included in the study, or 6.1 diagnoses per patient on the average. This indicates the complex nature of the type of patients included in the study. The higher than normal level of medical diagnoses is further complicated by the average IQ of 93 found in workers compensation patients with active cases, which comprised 35% of the Mensana Clinic population as well as 6% of the population that was functionally illiterate [10,11]. Therefore, 41% of the patient population would have some difficulty reading and understanding a written questionnaire. Since patients do not accurately complete paper and pencil questionnaires, this results in faulty information being conveyed and analyzed. This underscores the necessity of developing an input methodology that forces the patient to complete the questionnaire properly, such as an automated entry mechanism, that notes inconsistencies, i.e. if a patient marks he has pain in the leg, then he must complete the section on the symptoms of pain, or else the system will not let the patient continue. Conversely, if a patient does not mark that he has leg pain in the verbal section of the tests, and then completes the symptoms in the pictorial section of the test, he should be instructed to return to the verbal section. This potential source of errors has been addressed in a computerized version of the Diagnostic Paradigm and Treatment Algorithm, which is now available over the Internet, at www.mensanadiagnostics.com.

The purpose of an "expert system" is to improve the level of the reliability and accuracy of diagnosis, and enhance medical care. While the Diagnostic Paradigm is a first step to help diagnosis chronic pain patients, further research is needed to refine the value of the Diagnosis Paradigm. Work needs to be done by reducing the number of False Positive results, and by expanding the number of diagnoses covered by Diagnostic Paradigm. Moreover, the Treatment Algorithm can be further refined to make testing more specific. Finally, the Diagnostic Paradigm needs testing at other medical centers for further validation with other clinicians.

References

1. Hendler N (1982) The four stages of pain. Diagnosis and Treatment of Chronic Pain. Wright-PSG Publishing Co: 1-8.
2. Hendler N, Uematesu S, Long D (1982) Thermographic validation of physical complaints in 'psychogenic pain' patients. Psychosomatics 23: 283-287.
3. Hender N, Zinreich J, Kozikowski J (1993) Three-dimensional CT validation of physical complaints in "psychogenic pain" patients. Psychosomatics 34 :90-96.
4. Uematsu S, Hendler N, Hungerford D, Long D, Ono N (1981) Thermography and electromyography in the differential diagnosis of chronic pain syndromes and reflex sympathetic dystrophy. Electromyogr Clin Neurophysiol 21: 165-182.
5. Brown BR Jr (1978) Diagnosis and therapy of common myofascial syndromes. JAMA 239: 646-648.
6. Armitage KJ, Schneiderman LJ, Bass RA (1979) Response of physicians to medical complaints in men and women. JAMA 241: 2186-2187.
7. Daus AT, Freeman WW, Wilson J (1984) Psychological variable and treatment outcome of compensation and auto accident patients in a multidisciplinary chronic spinal pain clinic. Orthop Rev 13: 596-605.
8. Talo S, Hendler N, Brodie J (1989) Effects of active and completed litigation on treatment results: workers' compensation patients compared with other litigation patients. J Occup Med 31: 265-269.
9. Southwick SM, White AA (1983) The use of psychological tests in the evaluation of low-back pain. J Bone Joint Surg Am 65: 560-565.
10. Hendler N, Kozikowski J (1993) Overlooked Physical Diagnoses in Chronic Pain Patients Involved in Litigation. Psychosomatics 34: 494-501.
11. Hendler N, Bergson C, Morrison C (1996) Overlooked

physical diagnoses in chronic pain patients involved in litigation, Part 2. The addition of MRI, nerve blocks, 3-D CT, and qualitative flow meter. Psychosomatics 37: 509-517.
12. Long D, Davis R, Speed W, Hendler N (2006) Fusion For Occult Post-Traumatic Cervical Facet Injury. Neurosurg 16: 129-135.
13. Dellon AL, Andonian E, Rosson GD (2009) CRPS of the upper or lower extremity: surgical treatment outcomes. J Brachial Plex Peripher Nerve Inj 4: 1.
14. Hendler N, Murphy ME, Romano T (2010) Chronic Pain Due to Thoracic Syndrome, Acromo-Clavicular Joint Syndrome, Disrupted Disc, Nerve Entrapments, Facet Syndrome, and Other Disorders Misdiagnosed as Fibromyalgia: 10-13.
15. Hendler N (1988) Validating and Treating the Complaint of Chronic Back Pain: The Mensana Clinic Approach. Clinical Neurosurgery 35: 385-397.
16. Hendler N, Berzoksky C, Davis RJ (2007) Comparison of Clinical Diagnoses Versus Computerized Test Diagnoses Using the Mensana Clinic Diagnostic Paradigm (Expert System) for Diagnosing Chronic Pain in the Neck, Back and Limbs. Pan Arab Journal of Neurosurgery: 8-17.
17. Sandhu HS, Sanchez-Caso LP, Parvataneni HK, Cammisa FP Jr, Girardi FP, et al. (2000) Association between findings of provocative discography and vertebral endplate signal changes as seen on MRI. J Spinal Disord 13: 438-443.
18. Metaxiotis KS, Samouilidis JE (2000) Expert systems in medicine: academic exercise or practical tool? J Med Eng Technol 24: 68-72.
19. Engelbrecht R (1997) [Expert systems for medicine--functions and developments]. Zentralbl Gynakol 119: 428-434.
20. Babic A (1999) Knowledge discovery for advanced clinical data management and analysis. Stud Health Technol Inform 68: 409-413.
21. Schewe S, Herzer P, Krüger K (1990) Prospective application of an expert system for the medical history of joint pain. Klin Wochenschr 68: 466-471.
22. Schewe S, Schreiber MA (1993) Stepwise development of a clinical expert system in rheumatology. Clin Investig 71: 139-144.
23. Molino G, Marzuoli M, Molino F, Battista S, Bar F, et al. (2000) Validation of ICTERUS, a knowledge-based expert system for Jaundice diagnosis. Methods Inf Med 39: 311-318.
24. Cammà C, Garofalo G, Almasio P, Tinè F, Craxì A, et al. (1991) A performance evaluation of the expert system 'Jaundice' in comparison with that of three hepatologists. J Hepatol 13: 279-285.
25. Kentala E, Auramo Y, Juhola M, Pyykkö I (1998) Comparison between diagnoses of human experts and a neurotologic expert system. Ann Otol Rhinol Laryngol 107: 135-140.
26. Kentala EL, Laurikkala JP, Viikki K, Auramo Y, Juhola M, et al. (2001) Experiences of otoneurological expert system for vertigo. Scand Audiol Suppl : 90-91.
27. Cawthorpe D (2001) An evaluation of a computer-based psychiatric assessment: evidence for expanded use. Cyberpsychol Behav 4: 503-510.
28. Miller PR, Dasher R, Collins R, Griffiths P, Brown F (2001) Inpatient Diagnostic Assessments: 1. Accuracy of Structured vs. Unstructured Interviews. Psychiatry Res. 105: 255-264.
29. van der Lei J, van der Does E, Man in 't Veld AJ, Musen MA, van Bemmel JH (1993) Response of general practitioners to computer-generated critiques of hypertension therapy. Methods Inf Med 32: 146-153.
30. Silverman BG, Andonyadis C, Morales A (1998) Web-based health care agents; the case of reminders and todos, too (R2Do2). Artif Intell Med 14: 295-316.
31. Sinnott MM, Carr B, Markey J, Brosnan P, Boran G, et al. (1993) Knowledge Based Lipid Management System for General Practitioners. Clin Chim Acta 222: 71-77.
32. Kuilboer MM, van der Lei J, de Jongste JC, Overbeek SE, Ponsioen B, et al. (1998) Simulating an integrated critiquing system. J Am Med Inform Assoc 5: 194-202.
33. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, et al. (1994) Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 331: 69-73.
34. Braithwaite I, White J, Saifuddin A, Renton P, Taylor BA (1998) Vertebral end-plate (Modic) changes on lumbar spine MRI: correlation with pain reproduction at lumbar discography. Eur Spine J 7: 363-368.
35. Sandhu HS, Sanchez-Caso LP, Parvataneni HK, Cammisa FP Jr, Girardi FP, et al. (2000) Association between findings of provocative discography and vertebral endplate signal changes as seen on MRI. J Spinal Disord 13: 438-443.
36. Carragee EJ, Chen Y, Tanner CM, Truong T, Lau E, et al. (2000) Provocative discography in patients after limited lumbar discectomy: A controlled, randomized study of pain response in symptomatic and asymptomatic subjects. Spine 25: 3065-3071.
37. Bogduk, McGuirk (2002) Pain Research and Clinical Management. Elsevier 13: 119-122.