Comparison of clinical diagnoses versus computerized test diagnoses

Nelson Hendler[1], Craig Berzoksky[1], Reginald Davis[2,3]

Abstract: Past research from Mensana Clinic found that 40 - 71% of chronic pain patients had overlooked diagnoses, underscoring the need for more accurate diagnostic methodology. In this study, 937 diagnoses were made by the senior author during the initial evaluation of 87 chronic pain patients. Of these diagnoses, 903 diagnoses made by the senior author were also made by the computer scored and interpreted Mensana Clinic Diagnostics (MCD) "Diagnostic Paradigm." The MCD Diagnostic Paradigm matched the senior author's evaluation 96.37% of the time, and the Diagnostic Paradigm missed 34 diagnoses made by the senior author, for a 3.63% "missed diagnosis" rate, based on the initial clinic evaluation.

Overall, there were 2764 "false positives", i.e. diagnoses made by the Diagnostic Paradigm, but not made by the senior author. However, when analyzed by diagnostic groups, 2639 of the 2764 or 95.5% of the "false positive" diagnoses were in the same diagnostic group as the diagnoses made by the senior author. Therefore, these patients would receive the same diagnostic studies, thereby helping differentiate the correct diagnoses from the "false positive" diagnosis.

Key words: Radiculopathy, chronic pain, complex regional pain syndrome and reflex sympathetic dystrophy

Introduction

Past research reports from Mensana Clinic indicate that 40 to 67% of chronic pain patients involved in litigation are misdiagnosed.[18,19] 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.[14] Therefore, 97% of patients diagnosed by other physicians as having CPRS 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%.[15] 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.[12,16] 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.[16] The presence of psychiatric problems, even though a normal response to chronic pain, biases many physicians, which results in a less extensive evaluation.[12,17] 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.[1,10,13,20,28] These biases also influence the length of time and the extent of an evaluation. Some physicians spend less than 15 minutes with a patient, while other "high volume" physicians have reduced by 30% the amount of time they spend with their patients.[9,35] 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

[1]Mensana Clinic
Stevenson, Maryland 21153
USA

[2]Greater Baltimore Medical Center
Towson, Maryland 21204
USA

[3]Johns Hopkins University School of Medicine
Baltimore, Maryland 21205
USA

Correspondence:
Dr. Nelson Hendler, MD, MS
Mensana Clinic
Stevenson, Maryland 21153
USA

they call a "codification strategy," and direct person to person contact which they call the "personalization strategy".[11] The use of the "codification strategy" allows many people to search for and retrieve codified knowledge without having to contact the person who originally developed it.[11] Medicine has traditionally relied upon the tutorial or "personalization strategy" to provide clinical training. However, some aspects of clinical training do lend themselves to the more efficient and less expensive "codification strategy". All physicians should recognize the need for an accurate, thorough history to establish diagnoses. History taking does lend itself to computerization, as a way of sharing the knowledge of an experienced physician with a trainee and of assuring the thoroughness of an evaluation. However, an expert system is more than just an automatic history-taking tool. The interpretation of the answers to the questions, the integration of the answers and the ability to formulate diagnoses based on the integration and interpretation of the answers has traditionally been primarily a person to person training process in medicine. This is labour intensive and subject to personal variations, which generates uneven quality of care. In the past, this process had defied duplication.

Moreover, the selection and interpretation of appropriate laboratory studies is a highly individualized phenomenon. Physicians are poorly trained to recognizing the sensitivity and specificity of a laboratory study, and have a tendency to rely on the results of a laboratory test rather than their own clinical judgment. The fact that a laboratory test lends a degree of objectivity to diagnosis is very appealing in the uncertain world of medicine and provides a "linga Franca" that is universally understood and is less subjective than "clinical judgment". However, if there are false positive and false negative errors in the results of the laboratory studies, these errors and the elimination of them for diagnostic considerations, require clinical judgment. Furthermore, physicians fail to recognize the distinction between anatomical tests and physiological tests, which is a critical issue, since each category of testing provides a different answer to the same question, and the degree of correlation is very poor.

One method of eliminating the subjective component to diagnosis is to review outcome results. In this fashion, the efficacy of accurate diagnosis, and by extension, proper treatment, can be impartially analyzed. This "results based" data is being used by a number of hospital systems to market their institution, especially when compared to another institution offering the same services. Insurance carriers have used outcome studies from various hospitals to determine if the insurance company will reimburse for a procedure.

In the area of the diagnosis and treatment of chronic pain patients, most insurance carriers have abysmal results with claimants involved in workers compensation claims. The reported return to work rates, for claimants out of work for two years or more is less than 1%. For a comparable group of patients, Mensana Clinic has a return to work rate of 19.5% for workers compensation claimants, and 62.5% for auto accident cases.[12,28]

In order to see if the diagnostic methods could be duplicated, Mensana Clinic 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 customized version of the MCD tests were patients at Mensana Clinic. Mensana Clinic was an inpatient and outpatient multidisciplinary diagnostic and treatment centre for diagnosing and treating chronic pain problems. Seventy-five percent of the inpatients came from 43 states and 8 foreign countries. Of the 133 patients included in this study, 35% are involved in active workers compensation cases, 11% are involved in active automobile litigation and the remaining 54% are covered by commercial insurance or Medicare for disabled people. This corresponds with the statistics derived from a review of 554 distinct patients seen in the 1997-2000 period at Mensana Clinic. The patients included in this study represent new evaluations seen between January 2000 through July 2001. The average age of the patients was 41.2 years. Forty-six-percent of the subjects were males and 54% females. Previously reported average IQ of the workers compensation patients was 93, the average IQ of the auto accident patients was 99 and the remainder of the patients had an average IQ of 101.[18] The range for all patients was 75 - 121. Six-percent were functionally or totally illiterate. Earlier reports indicate that 77% were depressed at the time of evaluation but 89% had never been depressed before the onset of pain.[17]

Methods

For this study, an evaluation at Mensana Clinic 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 clinical director of Mensana Clinic, who was the senior author. A review of 133 charts from Mensana Clinic, which underwent the previously described process, were conducted. 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 several days later.

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, i.e. the clinical diagnosis of C5 - 6 radiculopathy was considered a match only if the tests generated the diagnosis of C5 - 6 radiculopathy, but was not a match if the tests generated a diagnosis of C4 - 5 radiculopathy or C6 - 7 radiculopathy, or C5 - 6 disrupted disc. Likewise, if the senior author diagnosed L3 - S1 facet syndrome and the test generated the diagnosis of retrolysthesis of L4 on L5, this was not considered a match, even though one of the components of retrolysthesis is the production of a lumbar facet syndrome. In this fashion, the test results were biased against a favourable outcome for the research, by requiring 1) a specific diagnosis and 2) a specific spinal level of pathology.

At the time the research was conducted, Version 7 of the MCD Diagnostic Paradigm contained 76 multi-part questions and Version 8 contained 95 multi-part questions ranging from 2 to 32 possible responses to a question, which was computer scored and interpreted. There were 1966 possible answers on Version 7 and 2068 possible answers to Version 8 of the MCD Diagnostic Paradigm. Depending on how many limbs or what part of the body is injured, the patient may take between 10 to 50 minutes to complete a questionnaire. At the time this research was conducted, these questionnaires were "paper-pencil" questionnaires, requiring that the patient read and complete the questionnaires by filling in a "bubble" answer on the forms with a #2 pencil. These questionnaires were then electronically scanned and scored by computer. These questionnaires asked location of the pain, the type of pain, what made the pain worse or better and then repeated the questions with pictures and associated symptoms, for an internal control. The MCD Diagnostic Paradigm was designed to detect 60 diagnoses and 44 differential diagnoses, 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.

The diagnoses generated by the MCD Diagnostic Paradigm can be grouped into 17 clinical categories or "diagnostic groups". These "diagnostic groups" are so grouped because any one of the diagnoses that appear in a diagnostic group very often have the same clinical symptoms and will require the same diagnostic procedures as any other diagnosis in the group, to further refine the diagnosis. As an example, the symptoms of ulnar nerve entrapment and thoracic outlet syndrome, both of which may manifest as numbness and pain in the ulnar forearm and last two fingers, are sometimes difficult to differentiate clinically. To differentiate these diagnoses requires the use of EMG/ nerve conduction velocity studies, vascular flow studies and a nerve block of the suspected affected nerve, to differentiate the diagnosis of ulnar nerve from thoracic outlet syndrome. Of course, the patient may have both of these medical conditions co-existing. The logical progressions through a diagnostic scheme are represented by the MCD Treatment Algorithm, which starts with the least expensive, least invasive tests and progresses through to the most complex, invasive and expensive testing, using branching logic, rank ordering the probability that certain tests would have higher yields than others. No research into the accuracy or correlation with outcome or results was ever done to test this premise, and this was a highly subjective, single physician determination.

The computer program for the analysis of the answers to the questions is proprietary. Initial programming of Version 7 and Version 8 of the MCD Diagnostic Paradigm was done by a computing firm no longer in existence and then revised and consolidated by Cape Computing Corporation, in Baltimore, Maryland, based on a scoring system developed by the senior author, which offered diagnosis and differential diagnosis for symptoms.

Sources of errors

***1) Filing errors*:** Of the 133 charts reviewed, 8 were not included, even though they had printouts of the diagnoses from the MCD Diagnostic Paradigm, since no actual questionnaire completed by the patient was in the file, which prevented verification of the actual answers to the questionnaire. This was termed a "filing error."

***2) Patient input error*:** Another 19 charts were excluded from analysis because of "patient input error." A "patient input error" is defined as an error in completing the questionnaire. There were four types of patient input error:

i) The patient marked pain in a body part then never completed the questions in the symptom section pertaining to that body part, i.e. they marked "pain in the leg" but then never marked any symptoms pertaining to pain in the leg, in either the drawings, showing the location of pain in the leg, or the verbal section of the test, defining the type of pain, what made it better or worse, and/or describing the location of the pain.

ii) The patient marked that they had no pain in a body part and then went on to complete the symptom section.

iii) The patient told the senior author they had pain in a particular body part during the clinical evaluation, but then never completed that section on the questionnaire.

iv) The verbal answers on the questionnaire did not correspond with the answers on the pictorial part of the questionnaire, i.e. the patient would mark they had pain in the outside of the ankle, but then on the pictorial part of questionnaire they would mark the inside of the ankle (internal discrepancy).

3) Inappropriate patient selection error: Another 13 patients were eliminated from the data, after the senior author performed an initial evaluation and none of the diagnoses for a patient that he made clinically were diagnoses designed to be covered by the MCD Diagnostic Paradigm, i.e. organic brain syndrome, post-concussion syndrome, severe reactive depression, amorosa fugax, partial complex seizures, torticollis, etc. Patients who were clinically diagnosed to have at least one of the diagnoses for which the Diagnostic Paradigm was programmed were included in the study.

4) Programming error: Six patients marked symptoms associated with a diagnosis, but the report of the MCD Diagnostic Paradigm did not report the diagnosis.[6] This was termed a "programming error." This was verified by checking each of the 133 questionnaires for every diagnosis that was reported in the initial evaluation, but not generated as a diagnosis by the report of the MCD Diagnostic Paradigm. Three of these errors were associated with tempro-mandibular joint syndrome and three were associated with acromo-clavicular joint syndrome. Once these errors were identified, the programming errors for these two diagnoses were detected and corrected, in a revised version, which is now available.

5) Interview oversight error: Included in the analysis were those patients who had marked questions for areas of the body that they had not discussed with the senior author during the initial evaluation. This type of error occurred when the senior author was interrupted during his evaluation of the patient, or he was either late, and rushed the evaluation, or the patient was late, and the evaluation time had to be reduced, or the patient was particularly inarticulate, and vague in their answers. When questioned after both the clinical evaluation and MCD Diagnostic Paradigm results were both available, the patient would often say "I forgot to tell the doctor," or "I was too anxious or scared during the clinical interview so I neglected to mention the problem," or "I felt rushed during the interview, so I didn't include these symptoms," or "I didn't think it was important, or "The doctor never asked me". This error was considered a "interview oversight error" and these represented ??? of the 133 charts. These charts demonstrated the value of a self administered questionnaire, which reduces the chance of missed or incomplete history, and therefore missed diagnoses, through physician oversight or patient communication errors. Patients who had at least one of the diagnoses for which the Diagnostic Paradigm was programmed were included in the study.

Author - How many?

As a result of the exclusion of charts, due to various errors, 87 charts were then available for review.

6) Diagnostic groups: The "diagnostic groups" represents a group of diagnoses from each of the diagnoses generated from the Report of the MCD Diagnostic Paradigm. There are 17 corresponding diagnostics groups, and each group is defined as a cluster of diagnoses that have similar aetiologies, and require the same medical diagnostic studies to differentiate one diagnosis from another. As an example, differentiating a C5 - 6 herniated disc from a C6 - 7 herniated disc requires the use of magnetic resonance images (MRI), 3D-CT, EMG-nerve conduction velocity studies, nerve root blocks, and provocative discometry. Using the Diagnostic Paradigm, each diagnosis generated by MCD Diagnostic Paradigm was placed into its corresponding group. The diagnosis groups, and the diagnoses subsumed in each group are shown in Appendix B. Also shown are the tests recommended for each diagnostic group. This is a critical concept for understanding the value of the MCD Diagnostic Paradigm and treatment algorithm. The MCD Diagnostic Paradigm offer diagnoses and differential diagnoses, that cluster within a group, and the tests within the treatment algorithm group confirm or eliminate various diagnoses within the group. Therefore, a positive diagnosis of a C5 - 6 radiculopathy would fall into diagnostic group 3, as would a "false positive" diagnoses of C4 - 5 radiculopathy and a C6 - 7 radiculopathy. However, any diagnosis in this group would still have the same diagnostic studies, which would help differentiate whether or not the diagnosis or false diagnosis was actually the correct diagnosis.

Appendix A - Diagnostic Paradigm was programmed to diagnose the following diagnoses

Diagnosis number for computer analysis	Name of diagnosis	Principal diagnosis or not	Diagnostic group based on cluster of tests recommended
1	Arachnoiditis L5 - S1	x	2
2	Carpel tunnel syndrome	x	4
3	Common perioneal nerve entrapment	x	1
4	C2 - 3 herniated or disrupted disc	x	6
5	C3 - 4 herniated or disrupted disc	x	6

Appendix A/ cont'd

6	C4 - 5 herniated or disrupted disc	x	5
7	C5 - 6 herniated or disrupted disc	x	5
8	C6 - 7 herniated or disrupted disc	x	5
9	Deep perioneal nerve entrapment	x	1
10	Femoral nerve entrapment	x	1
11	Lateral femoral cutaneous nerve entrapment	x	1
12	Lumbar facet syndrome L3 - S1	x	2, 14
13	L2 - 3 radiculopathy	x	2
14	L3 - 4 herniated or disrupted disc	x	2
15	L3 - 4 radiculopathy	x	2
16	L4 - 5 herniated or disrupted disc	x	2
17	L4 - 5 radiculopathy	x	2
18	L5 - S1 herniated or disupted disc	x	2
19	L5 -S1 radiculopathy	x	2
20	Neural foraminal stenosis L3 - 4	x	14
21	Neural foraminal stenosis L4 - 5	x	14
22	Neural foraminal stenosis L5-S1	x	14
23	Occipital neuralgia	x	6
24	Radial nerve entrapment	x	4
25	Reflex sympathetic dystrophy of the arm	x	4
26	Reflex sympathetic dystrophy of the leg	x	1
27	Retrolysthesis of L3 - 4	x	2, 14
28	Retrolysthesis of L4 - 5	x	2, 14
29	Retrolysthesis of L5 - S1	x	2, 14
30	Rib tip syndrome	x	8
31	Rotator cuff tear	x	12
32	Sacroiliac joint instability	x	3
33	Saphenous nerve entrapment	x	1
34	Superficial perioneal nerve entrapment	x	1
35	Supraspinatus tendonitis	x	12
36	Sural nerve entrapment	x	1
37	Syringomyelia-lower cervical spine	x	11
38	Temporo-manibular joint syndrome	x	9
39	Thoracic outlet syndrome	x	10
40	Tibial nerve entrapment	x	1
41	Tietze's syndrome	x	7
42	Ulnar nerve entrapment	x	4
43	Unstable spinal segment at L5 - S1	x	2
44	Unstable spinal segment at L4 - 5	x	2
45	Unstable spinal segment at L3 - 4	x	2
46	Acromo-clavicular joint impingement or sclerosis	x	12
47	Aseptic necrosis of hip	x	16
48	Brain stem lesion		99
49	Chondromalacia		99
50	C2 root compression		6
51	C2 - 3 facet syndrome		6
52	C2 - 3 radiculopathy		6
53	C2 - 3 unstable spinal segment		6
54	C3 - 4 facet syndrome		6
55	C3 - 4 radiculopathy		6
56	C4 - 5 cervical facet syndrome		5
57	C4 - 5 radiculopathy		5
58	C4 - 5 unstable spinal segment		5

Appendix A/ cont'd

59	C5 - 6 facet syndrome		5
60	C5 - 6 neural foraminal stenosis		5
61	C5 - 6 radiculopathy		5
62	C5 - 6 unstable spinal segment		5
63	C6 - 7 facet syndrome		5
64	C6 - 7 neural foraminal stenosis		5
65	C6 - 7 radiculopathy		5
66	Compression of posterior column of spinal cord		6
67	Deep vein thrombosis		99
68	Discitis		99
69	Genitofemoral nerve entrapment		1
70	Iliohypogastric nerve entrapment		1
71	Ilioinguinal nerve entrapment		1
72	Damage to the liver		8
73	Long thoracic nerve entrapment		17
74	L1 - 5 facet break		14
75	L3 - 4 herniated or disruped disc and unstable L3 - 4 segment		2
76	L3 - S1 facet break		14
77	L4 - 5 herniated or disrupted disc and unstable L4 - 5 segment		2
78	L5 - S1 herniated or disrupted disc and unstable L5 - S1 segment		2
79	Medial meniscus tear		99
80	Occult fracture of lateral malleolus		99
81	Occult fracture of medial malleolus		99
82	Occult fracture of navicular		99
83	Osteomyelitis		2, 13
84	Peripheral neuropathy		1, 4
85	Sacroiliac joint displacement		99
86	Spinal accessory nerve entrapment		17
87	Spinal stenosis		2
88	Spinal stenosis of cervical spine		5
89	Damage to the spleen		6
90	Spondylolysis/ spondylolythesis/ anterio-lysthesis/ unstable lumbar spinal segment		8
91	Unstable lumbar spine		2
92	L3 - S1 disc		2
93	L3 - S1 plexopathy		2
94	Intercostal neuralgia	x	2
95	Thoracic syrinx	x	15
96	C7 - T1 herniated or disrupted disc	x	15
97	T1 - 5 herniated or disrupted disc	x	5
98	T5 - 9 herniated or disrupted disc	x	15
99	T9 - 12 herniated or disrupted disc	x	15
100	T1 - 5 facet syndrome	x	15
101	T5 - 9 facet syndrome	x	15
102	T9 - 12 facet syndrome	x	15
103	Costovertebral joint displacement	x	15
104	Fibromyalgia		99

Author ? *Author ?*

Appendix B - Diagnostic groups and associated tests

No. of pts	Group	Diagnoses	Tests
2	Low back and leg pain due to disc, root damage, or motion	Arachnoiditis L5 - S1, lumbar facet syndrome L3 - S1, L2 - 3 radiculopathy, L2 - 3 herniated or disrupted disc, L3 - 4 radiculopathy, lumbar facet syndrome L3 - S1, L3 - 4 herniated or disrupted disc, L4 - 5 radiculopathy, L4 - 5 herniated or disrupted disc, L5 - S1 radiculopathy, L5 - S1 herniated or disrupted disc, retrolysthesis of L3 - 4, retrolysthesis of L4 - 5, retrolysthesis of L5 - S1, unstable spinal segment at L5-S1, unstable spinal segment at L4 - 5, unstable spinal segment at L3 - 4, L3 - 4 herniated or disrupted disc and unstable L3 - 4 segment, L4 - 5 herniated or disrupted disc and unstable L4 - 5 segment, L5 - S1 herniated or disrupted disc and unstable L5 - S1 segment, spinal stenosis of lumbar spine, unstable lumbar spine, L3 - S1 plexopathy,	MRI L1 - S1, 3D-CT L1 - S1, facet blocks L1 - 4, facet block L3 - S1, root block L1 - 2, root block L2 - 3, root block L3 - 4, root block L4 - 5, root block L5 - S1, EMG-nerve conduction velocity studies of the legs, neurometer studies of the legs, flexion-extension x-ray with obliques L1 - S1, provocative discogram L1 - S1, trial with body jacket with thigh spike, bone scan,
4	Arm pain - non-radicular	Ulnar nerve entrapment, radial nerve entrapment, median nerve entrapment, carpel tunnel syndrome, reflex sympathetic dystrophy of the arm (RSD) (CRPS I),	EMG-nerve conduction velocity studies of arms, neurometer studies of arms, nerve block of ulnar nerve, nerve block of radial nerve, nerve block of **median** nerve, bone scan, stellate ganglion block
1	Leg pain- non-radicular	Common perioneal nerve entrapment, deep perioneal nerve entrapment, entrapment, femoral nerve entrapment, lateral femoral cutaneous nerve entrapment, saphenous nerve entrapment, superficial perioneal nerve entrapment, sural nerve entrapment, tibial nerve entrapment, genitofemoral nerve entrapment, ilioinguinal nerve entrapment, ilionypogastric nerve entrapment, peripheral neuropathy of feet, reflex sympathetic dystrophy of leg (RSD) (CRPS I)	EMG-nerve conduction velocity studies of legs, neurometer studies of legs, nerve block of the common perioneal nerve, nerve block of superficial perioneal nerve, nerve block of deep perioneal nerve, bone scan, lumbar sympathetic block, saphenous nerve block, sural nerve block, tibial nerve block, blood for diabetes, T3, T4, thyroid antibodies, B12, folic acid, HIV, RPR, STS, genitofemoral nerve block, ilioinginal nerve block,
6	Upper cervical pain with or without associated headache	C2 - 3 herniated or disrupted disc, C3 - 4 herniated or disrupted disc, occipital neuralgia, compression of posterior column of spinal cord, **C2** - 2 root compression, C2 - 3 facet syndrome, C2 - 3 unstable spinal segment, C3 - 4 facet syndrome, C3 - 4 radiculopathy, neural foraminal stenosis C3 - 4, neural foraminal stenosis C2 - 3,	MRI ~~of~~ ~~c1~~ - 7, 3D-CT C1 - 7, occipital nerve block, facet blocks C2 - 5, root block C2 - 3, root block C3 - 4, root block C4 - 5, EMG-nerve conduction velocity studies of the legs, neurometer studies of the legs, flexion-extension x-rays with obliques C1 - 7 with open mouth odontoid views , provocative discogram C2 - 5, trial with 2 poster brace, bone scan,
5	Lower cervical pain with or without radiculopathy	C4 - 5 cervical facet syndrome, C4 - 5 neural foraminal stenosis, C4 - 5 radiculopathy, C4 - 5 unstable spinal segment, C5 - 6 facet syndrome C5 - 6, neural foraminal stenosis, C5 - 6 radiculopathy, C5 - 6 unstable spinal segment, C6 - 7 facet syndrome, C6 - 7 neural foraminal stenosis, C6 - 7 radiculopathy, compression of posterior column of spinal cord, due to excessive motion or spinal stenosis	MRI C1 - 7, 3D-CT C1 - 7, facet blocks C4 - 7, root block C5 - 6, root block C6 - 7, root block C7 - T1, EMG-nerve conduction velocity studies of the arms, neurometer studies of the arms, flexion-extension x-rays with obliques C1 - 7, provocative discogram C4 - T1, trial with 2 poster brace, bone scan,
12	Shoulder pain	Rotator cuff tear, supraspinatus tendonitis, acromio-clavicular joint impingement or sclerosis	MRI of shoulder, bone scan, arthrogram of shoulder
99	Miscellaneous	Occult fracture of lateral malleolus, occult fracture of medial malleolus, medial meniscus tear, chondromalacia, medial compartment syndrome, medial meniscus tear, medial compartment syndrome, deep vein thrombosis, discitis, occult fracture of the navicular,	Various tests
9	Jaw pain	Temporo-manibular joint syndrome	MRI of temporo-mandibular joint with jaw open and shut.
3	Low back pain	Sacroiliac joint instability	SI joint blocks, bone scan
8	Chest pain	Rib tip syndrome, slipping rib syndrome,	Blocks of the rib tip
7	Sternal pain	Tietze's syndrome	Blocks of the costo-sternal junction
10	Non-radicular arm pain	Thoracic outlet syndrome	Vascular flow studies, arms up and down, neurometer studies, somato-sensory evoked potentials, EMG-nerve conduction velocity studies, blocks of scalene muscles,
11	Numbness in chest/ arms	Syringomyella of lower cervical spine	MRI with gadolinium of C1 - T1, EMG-nerve conduction velocity of arms
12	Subcondral pain	Liver, pancreas, and spleen damage	CT of abdomen, MRI of abdomen, LDH, AST, ALT, SGOT, amylase, CBC with differential
13	Bone pain	Osteomyelitis	bone scan, indium 111, gallium, CBC with differential, C3, C4, total compliment, CPK, CRP, alkaline phosphate

Appendix B/ cont'd

14	Pain with movement	L3 - S1 facet break, L1 - 3 facet break, neural foraminal stenosis L3 - 4, neural foraminal stenosis L4 - 5, neural foraminal stenosis L – S1	Facet blocks at pars, 3D-CT, flexion-extension x-rays with obliques, trial with body jacket with thigh spike, root block L3 - 4, root block L4 - 5, root block L - S1
15	Thoracic pain	T1 - 5 herniated or disrupted disc, T5 - 9 herniated or disrupted disc, T9 - 12 herniated or disrupted disc, T1 - 5 facet syndrome, T5 - 9 facet syndrome, T9 - 12 facet syndrome, costovertebral joint displacement, thoracic syrinx	MRI T1 - 12, 3D-CT T1 - 12, facet blocks T1 - 4 facet block T7 - 12, root block T1 - 2, root block T2 - 3, root block T3 - 4, root block T4 - 5, root block T9 - 10, root block T10 - 11, root block T11 - 12, flexion-extension x-ray with obliques T1 - 12, provocative discogram T1 - 4, provocative discogram T8 - 12, trial with body jacket with thigh spike, bone scan, blocks of costo-vertebral joint.
16	Hip pain	Aseptic necrosis of hip	Bone scan, MRI of hip
17	Shoulder pain	Long thoracic nerve entrapment, spinal accessory nerve entrapment	Block of long thoracic nerve, spinal accessory nerve block

Results

There were 937 diagnoses made by the senior author during the initial evaluation on the 87 patients included in the study. Of these diagnoses, 903 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 34 diagnoses made by the senior author, for a 3.63% "missed diagnosis" rate, based on the initial clinical evaluation.

Overall there were 2764 false positives, i.e. diagnoses made by the Diagnostic Paradigm, but not made by the senior author. However, when analyzed by diagnostic groups, 2639 of the 2764 or 95.5% of the "false positive" diagnoses were in the same diagnostic group as the diagnoses made by the senior author, and therefore would receive the same diagnostic studies, thereby helping differentiating the correct diagnoses from the "false positive" diagnosis. As an example, if a patient has a C5 - 6 disc and radiculopathy diagnosed by the MCD Diagnostic Paradigm, and one of the differential diagnosis generated by the MCD Diagnostic Paradigm is C6 - 7, either of these diagnoses would require flexion-extension x-rays with obliques of the cervical spine, EMG/ nerve conduction velocity studies of the arms, an MRI of C2 - 7, a trial with a two poster brace, provocative discograms, etc.

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.[14,15,18,19]

After 30 years of work in this area, some authors feel only limited progress has been made in expert systems.[25] 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.[8] Babic concurs with the value of the longitudinal collection of clinical data and data mining to develop expert systems.[2]

The accuracy of any computer scored and interpreted expert system 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.[30] The correlation rate was 74.4%, and an error rate of 25.6%, with 44% of the errors attributed to "information deficits of the computer using standardized questions", but in a later version "RHEUMA" studied prospectively in 51 outpatients, achieved a 90% correlation with clinical experts.[30,31] Several groups have approached the diagnosis of jaundice. ICTERUS produced a 70% accuracy rate, while 'Jaundice' also had a 70% overall accuracy rate.[5,27] An expert system for vertigo was reported, and it generated an accuracy rate of 65%.[22] This later was reported as OtoNeurological Expert (ONE), which generated the exact same results reported in the earlier article.[23] There was a 76% agreement for diagnosis of depression, between an expert system and a clinician.[7] When a Computer Assisted Diagnostic Interview (CADI) was used to diagnose a broad range of psychiatric disorders, there was an 85.7% agreement level with three clinicians.[26] 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.[34] The panel accepted HyperCritic's critiques equally as beneficial as critiques from 8 human

reviewers.[34] 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.[32] 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.[33] 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.[24] Therefore, there may be more resistance from untrained physicians 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.

The rationale for the report of the MCD 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 would then require refinement using objective testing. In this fashion, the chance of missing a possible diagnosis is reduced. Moreover, 2639/2764 or 95.5% of the false positive results were those within the same cluster of diagnostic considerations or diagnostic group as the diagnosis generated by the MCD Diagnostic Paradigm requiring the use of objective testing to further refine and differentiate the specificity of diagnoses. 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 forward, and improvement with flexion.

Many of the recommended diagnostic studies from the treatment algorithm are not commonly used in community medical centres, but have been used for years by major teaching hospitals in the United States. A classic example of this is the widespread use of MRI for detecting disc damage in the cervical and lumbar spine. However, in 98 patients, 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.[4,6,21,29] The value of the provocative discogram is clearly demonstrated by the groundbreaking work by Bogduk, who clearly demonstrated pain fibres 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.[3] He terms this condition "internal disc disruption".[3] 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, nerve or facet blocks, bone, gallium or Indium 111 scans, neurometre studies, somatosensory evoked potentials, and flexion-extension x-rays with obliques. This is why the majority of the recommended tests in the Treatment Algorithm are physio-logical ones.

Additionally, there were 937 diagnoses made by the senior author on the 87 patients included in the study, or 10.7 diagnoses per patient on 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.[18] 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 error/s has been addressed in a computerized version of the MCD 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 MCD Diagnostic Paradigm is a first step to help diagnose 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 Paradigms. Moreover, the treatment algorithm can be further refined to make testing more specific. Finally, the MCD Diagnostic Paradigm needs testing at other medical centres for further validation

with other clinicians.

References

1. Armitage KJ, Schneiderman LJ, Bass RA: Response of physicians to medical complaints in men and women. J Amer Med Assoc 1979, 241: 2186-2187
2. Babic A: Knowledge discovery for advanced clinical data management Stud Health Technol Inform 1999, 68: 409-413
3. Bogduk **Initials?** McGuirk **Initials?**: Pain Research and Clinica Management. **Place Published Please?**, Elsevier 2002, 13 pp119-122
4. Braithwaite I, White J, Saifuddin A, Renton P, Taylor BA: Vertebral end-plate (Modic) changes on lumbar spine MRI: Correlation with pair reproduction at lumbar discography. Eur Spine J 1998, 7(5): 363-8
5. Camma C, Garofalo G, Almasio P, Tine F, Craxi A, Palazzo U, Pinzello G, Fiorello F, Angelo PM, Pagliaro L: A performance evaluation of the expert system 'Jaundice' in comparison with that of three hepatologists. J Hepatol 1991, 13(3): 279-85
6. Carragee EJ, Chen Y, Tanner CM, Truong T, Lau E, Brito JL: Provocative discography in patients after limited lumbar discectomy: A controlled, randomized study of pain response in symptomatic and asymptomatic subjects. Spine 2000, 25(23): 3065-71
7. Cawthorpe D: An evaluation of a computer based psychiatric assessment Evidence of expanded use. Cyberpyschol Behav 2001, 4 (4): **Page Numbers Please?**
8. Engelbrecht R: Expert systems for medicine-functions and development Zentralbl Gynakol 1997, 119(9): 428-34
9. Gross DA, Zyzanski SJ, Borawski EA, Cebul RD, Stange KC: Patient satisfaction with time spent with their physician. J Fam Pract 1998, 47 (2): 133-137
10. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH: Satisfaction, gender and communication in medical visits. Med Care 1994, 32(12): 1216-31 1994
11. Hansen M, Nohria N, Tierney T: What's your strategy for managing knowledge? Harvard Business Review, 1999, pp 106-116
12. Hendler N: Depression caused by chronic pain. J of Clin Psych 1984, 45: 30-36
13. Hendler N: Diagnosis and non-surgical management of chronic pain. In: Hendler N (ed). New York, Raven Press, 1981
14. Hendler N: Differential diagnosis of complex regional pain syndrome Pan Arab Journal of Neurosurgery 2002, 6(2): 1-9
15. Hendler N: Overlooked diagnosis in electric shock and lightning strike survivors. J of Occup and Environ Med 2005, 47(8): 796-805
16. Hendler N: The four stages of pain. In: Hendler N, Long D, Wise T (eds): Diagnosis and Treatment of Chronic Pain. John Wright/PSG, Littleton Mass 1982, Chapter 1, pp 1-8
17. Hendler N: Validating and treating the complaint of chronic pain: the Mensana Clinic Approach. In: P Black (Ed), Clinical Neurosurgery, Vo 35. Williams & Wilkens, Baltimore, Chapter 20 ,1989, pp 385-397
18. Hendler N, Bergson C, Morrison C: Overlooked physical diagnoses ir chronic pain patients involved in litigation, Part 2. Psychosomatics 1996 37(6): 509-517
19. Hendler N, Kozikowski J: Overlooked physical diagnoses in chronic pair patients involved in litigation. Psychosomatics 1993, 34(6): 494-501
20. Hendler N, Mollett A, Talo S: A comparison between the MMPI and the "Mensana Clinic Back Pain Test" for validating the complaint of pain. J of Occup Med 1988, 30: 98-102
21. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS: Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994, 331(2): 69-73
22. Kentala E, Auramo Y, Juhola M, Pyykko I: Comparison between diagnoses of human experts and a otoneurological expert system. Ann Otol Rhinol Laryngol 1998, 107(2): 135-40
23. Kentala EL, Laurikkala JP, Viikki K, Auramo Y, Juhola M, Pyykko I: Experiences of otoneurological expert system for vertigo. Sand Audio (Suppl) 2001, 52: 90-91
24. Kuilboer MM, van der Lei J, de Jongste JC, Overbeek SE, Ponsioen B van Bemmel JH: Simulating an integrated critiquing system. J Am Med Inform Assoc 1998, 5(2): 194-202
25. Metaxiotis KS, Samouilidis JE: **Name of Article?** J Med Eng Techno 2000, 24(2): 68-72
26. Miller PR, Dasher R, Collins R, Griffiths P, Brown F: Inpatient diagnostic assessments: 1. Accuracy of structured *vs.* unstructured interviews Psychiatry Res 2001, 105(3): 255-264
27. Molino G, Marzuoli M, Molino F, Battista S, Bar F, Torchio M, Lavelle SM Corless G, Cappello N: Validation of ICTERUS, a knowledge-based expert system for jaundice diagnosis. Methods Inf Med 2000, 39(4-5): 311-8
28. Talo S, Hendler N, Brodie J: Effects of active and completed litigation on treatment results: Workers compensation patients compared with other litigation patients. J of Occup Med 1989, 31: 265-270
29. Zyzanski SJ, Stange KC, Langa D, Flocke SA: Trade-offs in high volume primary care practice. J Fam Prac 1998, 46(5): 397-402
30. Sandhu HS, Sanchez-Caso LP, Parvataneni HK, Cammisa FP Jr Girardi FP, Ghelman B: Association between findings of provocative discography and vertebral endplate signal changes as seen on MRI. J Spinal Disord 2000,13(5): 438-43
31. Schewe S, Herzer P, Kruger K: Prospective application of an expert system for the medical history of joint pain. Klin Wochenschr 1990, 68 (9): 466-471
32. Schewe S, Schreiber MA: Stepwise development of a clinical expert system in rheumatology. Clin Investig 1993, 71(2): 139-44
33. Silverman BG, Andonyadis C, Morales A: Web based health care agents: The case of reminders and todos, too (R2Do2). Artificial Intel Med 1998, 14(3): 295-316
34. Sinnott MM, Carr B, Markey J, Brosnan P, Boran G, Grimson J, McGill T Feely J, O'Moore R: Knowledge based lipd management system for general practitioners. Clin Chim Acta 1993, 222(1-2): 71-77
35. van der Lei J, van der Does E, Man in 't Veld AJ, Musen MA, van Bemmel JH: Response of general practitioners to computer generated critiques of hypertension therapy. Methods of Info Med 1993, 32(2): 146-153

Author - Initials + Place Pub. please?

Author - Page No. Please?

Author - Name of article please