The Influence of an Expert System for Test Ordering and Interpretation on Laboratory Investigations

Beverly J. Smith1* and Michael D.D. McNeely2

Background: The Laboratory Advisory System (LAS) is an expert system interface that works interactively with clinicians to assist them with test selection and result interpretation throughout the laboratory investigation of a patient.

Methods: To study the influence of the LAS on laboratory investigations, a repeated-measures experiment using clinical vignettes was conducted. To collect baseline data on how laboratory investigations are currently conducted, clinicians investigated one-half of the vignettes using a conventional (noncomputer) approach. To determine the influence of the LAS on clinicians’ behavior, the other half of the vignettes were investigated using the LAS.

Results: Clinicians using the LAS (compared with conventional practice) ordered fewer laboratory tests during the diagnostic process (mean, 17.8 vs 32.7), completed the diagnostic workup with fewer sample collections (mean, 5.8 vs 7.5), generated lower laboratory costs (mean, $194 vs $232), shortened the time required to reach a diagnosis (mean, 1 day vs 3.2 days), showed closer adherence to established clinical practice guidelines, and exhibited a more uniform and diagnostically successful investigation.

Conclusion: The LAS enhances the outcome of the investigation and improves laboratory utilization.

© 1999 American Association for Clinical Chemistry

Over the past decade, the cost of laboratory testing has increased considerably around the world (1). Although the unit cost of performing each test has declined relative to inflation, the number of tests ordered has increased substantially. Various initiatives have been introduced to control this increase in “utilization”, including feedback, participation, education, cost-awareness, financial incentives, penalties or risk sharing, administrative change, and rationing (2–4). One of the most effective approaches is the use of ordering protocols (i.e., mandated clinical practice guidelines) (5). However, to have an effect on costs and to be workable, such protocols (a) must address high-volume ordering areas; (b) must be amenable to a few, simple rules that can be easily remembered by clinicians, conveniently expressed in the test order, and easily carried out by the laboratory; and (c) require general agreement among clinicians, laboratories, and payment agencies. Thus, although ordering protocols and clinical practice guidelines offer a proven approach to utilization management, the number of areas in which they can be implemented successfully has been limited because of these constraints. It has even been suggested that this development cease until appropriate methods of implementation become available (6).

A further drawback to the use of guidelines for utilization control is that guideline dissemination alone does not sustain clinician behavior change even when clinicians are in agreement (7–9). Guideline implementation is more likely to succeed when accompanied by changes in the practice environment that make it easier for physicians to maintain a new desired behavior (8). The work environment must provide opportunities to practice the new behavior and provide continuing reinforcement and feedback. In their discussion paper, Elson and Connelly (8) stated that: “Physicians’ behavior cannot be systematically corrected merely by changing knowledge, attitudes and/or motivation . . . they need new tools that will make it easier for them to do the right thing”. An example of this is a recent study investigating the laboratory testing of 969 newly diagnosed cases of essential hypertension (10). The authors found that 24% of the cases did not receive the laboratory testing recommended by the guidelines of the Canadian Hypertension Society, whereas >50% received tests that were not specifically recommended. The Canadian Hypertension Society guidelines

---

1 Acquired Intelligence Inc., No. 205-1095 McKenzie Ave., Victoria BC V8P 2L5, Canada.
2 MDS Metro Laboratory Services, 4489 Viewmont Ave., Victoria BC V8Z 5K8, Canada.
*Author for correspondence. Fax 250-479-0764; e-mail bsmith@aiinc.ca.

Received December 9, 1998; accepted May 10, 1999.
were widely disseminated: they were mailed to every licensed physician in the country and published in the Canadian Medical Association Journal. The authors noted that “it is unlikely that any guideline could be more widely disseminated without extraordinary expense” (10).

An expert system implemented as an interface between the clinician and the laboratory offers the possibility of addressing some of the difficult issues of developing, disseminating, and adhering to guidelines. The major benefits of an expert system interface are (a) the ability to represent more sophisticated and widely applicable protocols than can currently be implemented with traditional approaches, (b) the ability to make that information available to clinicians at the time of ordering and viewing the results, (c) the ability to make the information specific to the clinical circumstances of the patient, and (d) to provide a complete record for retrospective review of the clinical problem, test orders, and results. Expert systems have been applied to specific clinical problems for 15 years (11). Recently, these techniques have been directed toward more generalized clinical assistance (12, 13). Advances in office computerization and in expert systems technology (in both integration and representation schemes) have made the development of an expert system interface timely.

Today, many physicians use computers as an integral part of their office practices. Some physicians work interactively with a computer during their office activities to review their patients’ medical records. It is expected that this trend will continue. Peters (14) argues that computer technology is now able to facilitate the routine use of clinical guidelines and should contribute proactively to the delivery of medical care.

The Laboratory Advisory System (LAS) is an expert system program that works interactively with clinicians to assist them with test selection and result interpretation throughout the laboratory investigation of a patient. More than simply a tool for electronic ordering and results receipt, the LAS uses its underlying knowledge bases to optimize the laboratory investigations for better care and lower costs by (a) optimizing patient-specific test ordering strategies, (b) providing patient-specific result interpretation, and (c) offering context sensitive assistance throughout the process.

To study the influence of the LAS on clinicians’ pursuit of laboratory investigations, a repeated measures experiment using clinical vignettes was conducted. To collect baseline data on how laboratory investigations are currently conducted, physicians investigated one-half of the vignettes using a conventional (noncomputer) approach. The other half of the vignettes were investigated using the LAS. Each vignette appeared an equal number of times in the conventional condition as in the LAS condition. The influence of the LAS on clinicians’ pursuit of laboratory investigations was measured by the number of tests ordered and the number of sample collections required. Additional observations were made of the laboratory costs, the turnaround time, the adherence to established clinical practice guidelines, and the success of the investigation.

**Materials and Methods**

The LAS has been described previously (15). It has two operating modes: test ordering and result receipt. In the ordering mode (see Fig. 1), tests can be requested by name or through the pursuit of a clinical problem (e.g., thyroid function). The diagnostic approach to each clinical problem is guided by the expert system, which may use several steps in which clinical information is elicited, testing approaches are presented, and further information is available. For example, if the clinician pursues a thyroid function investigation, clinical information is collected to make the appropriate recommendation (see Fig. 2).

In the result receipt mode (see Fig. 3), the test results are displayed in a standard fashion. Below the numerical values an interpretation is presented. The knowledge base of the expert system is able to combine the clinical information elicited during the ordering phase with the numerical value(s) of the tests to derive a specific interpretation. In addition, an option for further testing (on the same sample) may be provided. This feature allows additional testing to be carried out when it will enhance the diagnosis. Because the second round of testing is conducted on the same sample, it is possible to reduce costs and turnaround time by avoiding a second venipuncture. Interpretation of the results of additional tests takes into account the clinical information provided and the test results from earlier rounds. Fig. 4 shows examples of a recommended follow-up test and of an interpretation of results based on information from all rounds of testing. An additional feature of the system is context-sensitive “help”, which provides explanatory information about each test, diagnostic pursuit, and clinical problem.

The LAS acts as an intelligent interface between the clinician and the laboratory information system. It is built...
using the ACQUIRE® suite of expert system tools (16–19). This includes a development environment that consists of a knowledge acquisition system, an inference engine, and an application programmer’s interface. The knowledge acquisition system provides a structured approach to developing the expert system application. In addition to the standard production rule representation of knowledge, ACQUIRE offers a knowledge representation scheme based on a pattern recognition approach that conceptualizes expertise as a highly developed pattern-recognition skill. Following from this approach, the specialist’s knowledge is captured in “pattern-consequence” relationships. In the LAS knowledge bases, patterns are made of relevant clinical information, and the consequences are testing recommendations and interpretations. Representing expertise as pattern-consequence relationships provides greater specificity in the knowledge base and is particularly useful for representing the exceptional cases not handled well by production rules. The LAS user interface is built using Visual Basic, and the underlying relational database is implemented in Microsoft Access. The LAS provides a well-defined programmable interface to integrate with other compatible software systems such as patient management systems and laboratory information systems. In this experiment, the LAS ran under Microsoft Windows 95 on a Pentium laptop with 16 megabyte memory.

THE TRIAL

Subjects. Physicians with between 10 and 20 years of experience, working in private, general practice were contacted. Six physicians with varying familiarity with computers were able to volunteer their time during the experimental period. (Procedures using human subjects were in accordance with the current revision of the Helsinki Declaration of 1975.)

Vignettes. Vignettes of 14 cases representing standard clinical problems found in ambulatory practice were prepared (i.e., 4 thyroid, 4 hepatitis, and 6 autoimmune). For example, one example states: “The patient has been on thyroid replacement therapy for 3 months. The dosage was increased 4 weeks ago. She continues to complain of occasional palpitations. You wish to monitor the status of the therapy. What do you order?” Test results returned to the physician were indicative of excessive replacement therapy, i.e., a raised free thyroxine result (26 pmol/L; reference range, 10.5–20.0 pmol/L) and a suppressed thyrotropin (TSH; 0.02 mIU/L; reference range, 0.5–3.59 mIU/L). Another example states: “A previously well 37 years old high school teacher with no previous history of drug or alcohol abuse or transfusion, attends your office with a seven day history of malaise. One week ago he noted his urine to be ‘very dark’ On examination he is slightly jaundiced, and his liver is slightly enlarged, tender and smooth. You wish to pursue the possibility of hepatitis. What do you order?” The test results rule out hepatitis, i.e., the aspartate aminotransferase is within the reference range (30 U/L; reference range, ≤40 U/L). A final example states: “A 45 years old woman has a several year history of gradual thinning of the skin on her fingers. She has had Raynaud’s phenomena since she was a teenager but these symptoms have become occasionally debilitating in the past three years. She has also experienced episodes of malaise and in the past 6 months has experienced shortness-of-breath on exertion. You wish to pursue a diagnosis of scleroderma (disseminated sclerosis). What do you order?” The laboratory results confirm the diagnosis of scleroderma, i.e., the antinuclear antibody titer was a high positive, the antinuclear antibody pattern was nucleolar, and the Scl70 was positive. In each case a specific diagnostic direction was suggested. This
was done because it was the test ordering and result interpretation for a given diagnosis that was of interest, not the diagnostic ability of the clinician. In each case, the laboratory results were unambiguous in confirming or rejecting the diagnosis the physician was asked to investigate.

Procedure. In this repeated measures experiment, each clinician was given one-half of the vignettes to investigate using the conventional (noncomputer) approach, using paper requisitions and paper laboratory reports. After the investigation of these cases was completed, each clinician was given the other half of the vignettes to investigate using the LAS. It is important to note that the advice provided by the LAS can be completely or partially ignored and that additional tests can be added. To emulate the usual office environment, the initial laboratory orders were obtained on all the cases as though they were consecutive patients in an office. Only when all the orders had been placed were the results returned. The clinicians were then interviewed to find out what diagnostic conclusions they had been able to reach from the test results and whether there were any further tests that they required. These steps were repeated until the clinicians had completed the investigation. Each vignette appeared an equal number of times in the two investigative modes. The presentation of cases within a session was randomized.

The investigative approaches, conventional and LAS, were compared to determine whether LAS could influence physician behavior. The criteria studied were:

(a) The number of tests ordered. The term “test” refers to the common ordering practice in our jurisdiction, where panel testing is generally discouraged. In some instances a test such as hematology profile or urinalysis will involve several “results”, but these have been considered single tests.

(b) The number of venipunctures (or other collections) required. It was assumed that with the LAS, the initial sample would be used if any additional test (or tests) was recommended by the system. In the conventional version, a second or follow-up collection was assumed to be needed for a second set of tests. Although samples currently are held and calling to request further tests is acceptable, it is not standard practice. In the instructions to subjects, physicians were asked to handle the vignettes as much like a real investigation as possible, and none of the physicians suggested that they would call the laboratory to order further tests on a sample.

(c) The cost of the testing strategies. It was not the goal of this study to determine the precise financial impact of the LAS. However, it is possible to apply in-house costing estimates (i.e., for collection, analysis, and reporting) to the investigations physicians conducted in both the conventional and LAS mode. The results offer some indication of the direction in which the LAS will influence costs.

(d) Diagnostic accuracy. The criterion for diagnostic accuracy was whether the physician correctly confirmed or rejected the diagnosis that he or she was asked to investigate, e.g., hypothyroid, hepatitis B immunity, or scleroderma.

(e) Turnaround time. It was assumed that the laboratory collections were made on the same day as the office visit and that laboratory test results were returned at 24 h (1 day). Other assumptions in the analysis were:

(i) Second-round tests with LAS are reported back in 1 day,

(ii) Second-round tests without LAS take 1 day to notify the patient and achieve a second collection plus an additional 1 day for analysis and reporting.
(iii) Tests that generally take longer than 1 day were assigned “reasonable” timing (for example, to book and complete an electrocardiogram or an x-ray of the hands was 3 days).

Results
All physicians were readily able to use the LAS and were accepting and approving of the LAS recommendations and interpretations. This acceptance was not unquestioning. There were times when physicians challenged the LAS interpretations, and there were often times when they chose to add more tests to those recommended by the LAS. In comparing the physicians’ test ordering strategies for a given case, we found that the test ordering in the conventional approach was well reasoned, varied more among the cases, was less “optimal” in terms of adherence to existing guidelines, and required more re-

Fig. 4. Recommended follow-up test (top) and an interpretation of results based on information from all rounds of testing (bottom).
In this example, the clinician is pursuing thyroid function. After the physician indicated pursuit of thyroid “diagnosis” (rather than monitoring or investigation of special cases), a TSH was recommended. The top screen illustrates the result of the TSH test, the interpretation, and the follow-up recommendation. The bottom screen illustrates the results of all tests and the interpretation of the tests based on all preceding clinical information and test results. In addition, the clinician is given the option to pursue the clinical problem further with an additional test that can be conducted on the original sample.
ferrals to specialists than with the LAS. In addition, there was one case misdiagnosed without the LAS.

**NUMBER OF TESTS**

A comparison of the mean number of tests showed that the six clinicians ordered a mean of 32.7 tests with the conventional approach vs 17.8 when they used the LAS ($t = 5.4; P < 0.01$). This is a highly significant difference, especially because the physicians using the conventional approach referred 34% of the cases to specialists with little or no laboratory investigation.

**NUMBER OF SAMPLE COLLECTIONS**

All first-round testing was judged to involve one collection. If a second round of testing was requested, a second collection was attributed to the conventional approach, whereas a follow-up collection was not required when the LAS was used, provided the original sample was appropriate. Significantly more sample collections were required with the conventional approach (mean, 7.5) vs the LAS (mean, 5.8; $t = 3.4; P < 0.02$). Again the difference is striking given that so many of the cases investigated with the conventional approach were referred to specialists rather than pursued with laboratory tests. Consideration of the number of sample collections is important for two reasons: First, 30% of the overall cost of performing laboratory testing on an outpatient is attributable to the collection process (20). Any methods that would reduce patient encounters would potentially reduce overall costs. Second, additional testing without requiring a second patient visit will significantly reduce turnaround time (see below). These results and the following observations are summarized in Table 1.

**COST OF TESTING**

The total cost of each investigation, as if it were a patient workup, was determined by adding the internal costing estimates of the tests ordered and the collections required. The results show significantly higher costs for investigations pursued using the conventional approach (mean, $232$) vs the LAS (mean, $194; t = 3.3; P < 0.05$). This is partially attributable to fewer tests, but as Winkelman (21) points out, the relationship between costs and number of tests performed is nonlinear. The greatest reduction in costs is attributable to fewer sample collections. As mentioned above, the sample collection process accounts for a large proportion of the overall costs. This difference is again striking because the cases that were referred to specialists rather than pursued with laboratory testing (all in the conventional approach) contributed zero costs. It should be noted that because the vignettes do not represent the usual mixture of ambulatory clinical problems, these results cannot be extrapolated to gauge the total financial impact of the system.

**DIAGNOSTIC ACCURACY**

Physicians using the LAS were more likely to arrive at correct diagnosis (100% of the time vs 66%). In the investigations using the conventional approach, physicians were correct in confirming or rejecting the diagnosis 66% of the time because physicians investigating autoimmune disorders referred many of the cases to specialists before reaching a correct diagnosis. In addition, one physician did not order a key test (Scl70 for scleroderma) and incorrectly concluded the presence of a disorder “on clinical grounds” when the result was negative.

**TURNAROUND TIME**

In all cases using the LAS, the cases could be completely turned around in 1 day. In contrast, when the conventional approach was used, the following mean times could be estimated: thyroid cases, 4.5 days; hepatitis cases, 1.4 days; autoimmune cases, 5.3 days (mean of means, 3.2 days). Reducing the turnaround time of a laboratory investigation is beneficial because it reduces the number of “cases in progress”, thereby streamlining the physician’s office. It also allows the diagnosis to be reached sooner.

Given the significant differences in the number of tests ordered, the number of samples collected, and the cost of testing, it is possible to examine the impact of LAS under different circumstances: (a) in cases for which guidelines exist, e.g., thyroid testing; (b) in cases for which guidelines are not available, e.g., hepatitis and autoimmune; and (c) in cases for which the results are negative.

<table>
<thead>
<tr>
<th>Table 1. Comparison of laboratory investigations using the conventional approach and the LAS.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Results</strong></td>
</tr>
<tr>
<td>Mean number of tests ordered</td>
</tr>
<tr>
<td>Mean number of samples collected</td>
</tr>
<tr>
<td>Mean costs, $</td>
</tr>
<tr>
<td>Turnaround time (mean no. of days)</td>
</tr>
<tr>
<td>Diagnostic accuracy</td>
</tr>
<tr>
<td>Number of referrals</td>
</tr>
<tr>
<td>Established guidelines: thyroid testing</td>
</tr>
<tr>
<td>Total number of sample collections</td>
</tr>
<tr>
<td>No established guidelines: hepatitis and autoimmune testing</td>
</tr>
<tr>
<td>Mean number of tests ordered</td>
</tr>
<tr>
<td>Results returned negative</td>
</tr>
<tr>
<td>Total number of sample collections</td>
</tr>
</tbody>
</table>
ESTABLISHED GUIDELINES

Given that making physicians aware of the guidelines even if they agree with them is not sufficient to sustain behavior change (8), it is of interest to investigate the influence of the LAS in an area where there is a well-established, legislated protocol, i.e., thyroid testing. In addition to being well established and commonly used in our jurisdiction, the implementation of guidelines in this area was documented in 1990 to save more than $1 million per year in a region with a population of 2.3 million (20).

In the current study, a subset of four cases pertaining to thyroid testing was examined to investigate the influence of LAS in a clinical area in which established guidelines exist.

The mean number of tests ordered was the same for physicians using the conventional approach and the LAS approach, but fewer sample collections were required to complete the investigation using the LAS. We looked at the individual test cases and the physicians’ interpretations and found that better compliance with the established guidelines was achieved when the LAS was used and that more specific diagnoses were made with the LAS. For example, a physician using the conventional approach ordered a test that is no longer recommended by the current guideline. Two physicians using the conventional approach interpreted results differently from the guidelines (they used different reference ranges), and in the two cases investigating thyroid monitoring, physicians used shorter retest periods than the guidelines suggest. Physicians using the LAS were more uniform in their test ordering, more consistent with the guidelines, and by choosing to clarify the presence of thyroiditis as offered by the LAS, achieved more specific diagnoses. Although the analysis of this subset of data is limited, there is indication that even with well-established, generally agreed upon, legislated guidelines, there are better outcomes when the LAS approach is used.

NO ESTABLISHED GUIDELINES

A similar investigation was conducted on the cases in clinical areas that did not have established guidelines, i.e., hepatitis and autoimmune disease testing. In these cases, there were twice as many tests ordered by physicians reviewing the cases using the conventional approach than by physicians using LAS, 27% more sample collections, and many more referrals to specialists. The differences in tests ordered and samples collected are conservative given that clinicians in the conventional mode referred 34% of the cases to specialists with little or no investigation.

NEGATIVE RESULTS

Given that the most common outcome of outpatient laboratory testing is negative or normal results, it is of interest to investigate the influence of LAS in the subset of cases for which the results are returned to the physician as negative. There were four such cases (two thyroid, one hepatitis, and one autoimmune). In these cases, there were more than twice as many tests ordered by physicians using the conventional approach than by physicians using the LAS and 55% more sample collections.

Discussion

The results of this study indicate that the use of the LAS reduces the number of tests, the number of sample collections, the laboratory costs, and the turnaround time required to reach a diagnosis. When physicians used the LAS, they were more likely to arrive at a correct diagnosis and to be more consistent with established guidelines than when they used the conventional approach. The greatest difference between pursuing laboratory investigations using the LAS vs the conventional approach was in cases for which the test results were negative and those for which established guidelines were not available.

One of the important features of the LAS is its ability to derive and reason with clinical information during the ordering process and to store this information in a database. This information is notoriously difficult for laboratories to obtain. Having such information allows several possibilities:

(a) The development of test ordering strategies can be enhanced.
(b) The interpretation of the test results can be enhanced.
(c) A statistical database of diagnosis, clinical information, test orders, and results can be readily derived. Such information is unique and is available for optimizing and developing testing strategies and for laboratory management (1).
(d) An appropriate search of the database would enable clinician-targeted education and utilization feedback to be derived.
(e) Examination of the database at the time of ordering would enable the development of a module to identify unnecessary, duplicate testing.
(f) With appropriate additions to the ordering module, a sophisticated “front end” to a compliance-checking program could be developed.

In conclusion, an expert system-based laboratory investigation tool (LAS) has been developed to assist physicians with ordering, interpreting, and understanding laboratory investigations. In this trial, the LAS enhanced the outcome of the investigation while reducing test utilization, turnaround time, and cost. As clinical practice becomes more electronically connected, such knowledge-based tools will be available for every physician to use. This study demonstrates that they may have a manifest effect on improving laboratory investigations.

Partial funding of the development of LAS was provided by the National Research Council of Canada, and partial funding of this experimental trial was provided by Ac-
quired Intelligence Inc. and MDS Metro Laboratory Services.

References